Ep 7 - Fundamentals of Probability

Dec 2, 2023

Key Concepts

Probability Space: A probability space consists of three components: a sample space (S), a set of events (E), and a probability measure (P).

- **Sample Space (S):** It's the *set of all possible outcomes of an experiment*. For example, when rolling a fair six-sided die, the sample space is {1, 2, 3, 4, 5, 6}.
- **Event (E):** An event is a subset of the sample space. It *represents a specific outcome or a set of outcomes*.

Probability of an Event: The probability of an event is denoted as $P(E)$, where E is the event. Probability values range from *0 (impossible)* to *1 (certain)*.

$$P(E) = \frac{\text{Number of favourable outcomes}}{\text{Total number of possible outcomes}}$$

Complement of an Event: The complement of an event E, denoted as E', represents *all outcomes not in E*.

$$P(E') = 1 - P(E)$$

Union of Events: The union of events A and B ($A \cup B$) represents all outcomes that *belong to either A or B or both*.
Intersection of Events: The intersection of events A and B ($A \cap B$) represents outcomes that *belong to both A and B*.

Mutually Exclusive Events: Events A and B are mutually exclusive if they cannot occur at the same time.

Addition Rule: For any two events A and B, the probability of their union is given by:

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

$P(A \cap B) \rightarrow$ to account for double-counting if A and B are not mutually exclusive.

Dependant and Independent Events: *Dependent events* are events where the outcome of one event affects the probability of the other event occurring. *Independent events* are events where the outcome of one event has no influence on the probability of the other event.

Axioms of Probability

- **Non-negativity**: The probability of an event is never negative, P(A) $\geq$ 0 for any event A.
- **Additivity**: For mutually exclusive events, the probability of their union is the sum of their individual probabilities.
- **Total Probability**: The probability of the entire sample space is 1. P(S) = 1, where S is the sample space.

Types of Probability

Marginal Probability: This is the probability of a single event occurring without considering other events.

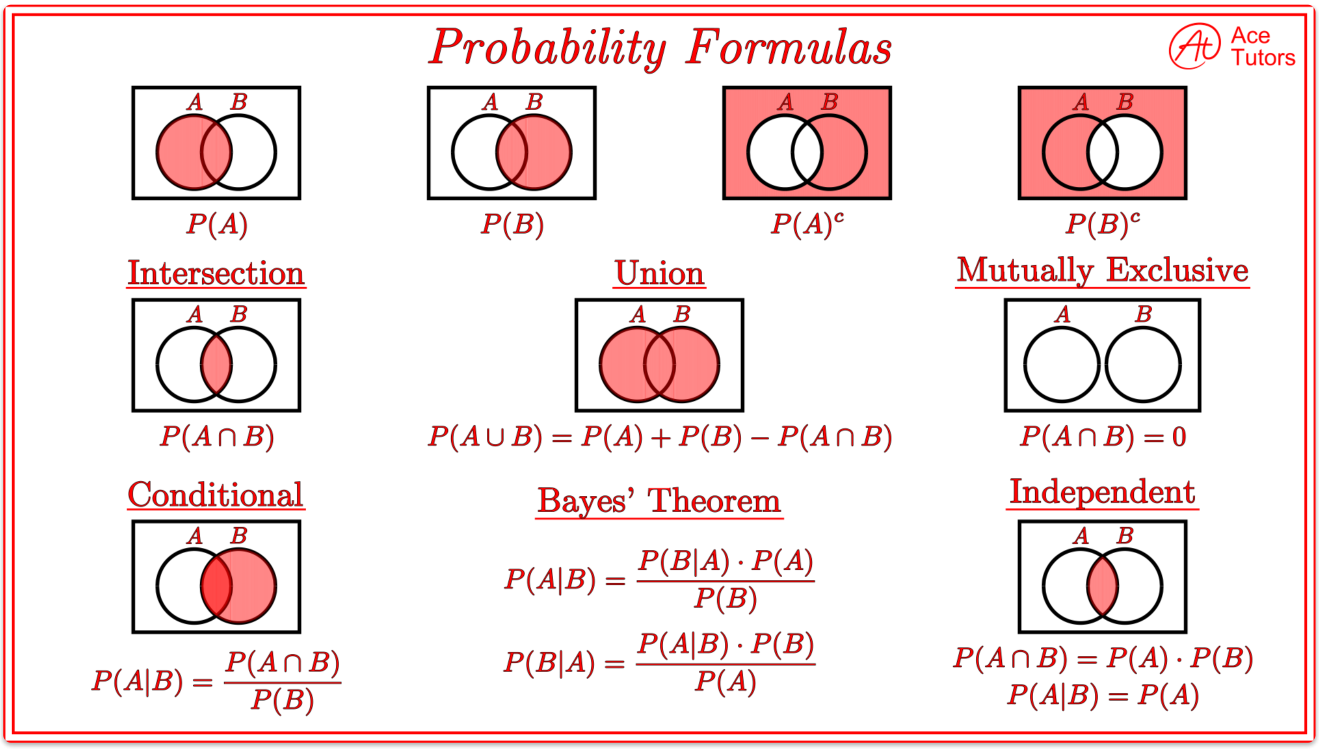
Joint Probability: This refers to the probability of two (or more) events occurring together. It's denoted as $P(A \text{ and } B)$ or $P(AB)$ or $P(A \cap B)$, where A and B are events. It is also known as *intersection event probability*.

Union Probability: It relates to the probability of at least one of two or more events occurring. The probability of the union of events A and B, denoted as $P(A \text{ or } B)$ or $P(A + B)$ or $P(A \cup B)$ It is also known as *union event probability*.

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

Conditional Probability: This is the probability of an event occurring given that another event has already occurred. It's denoted as $P(A|B)$, where A is the event of interest, and B is the condition. For instance, the probability of drawing a red card from a deck given that the first card drawn was a face card.

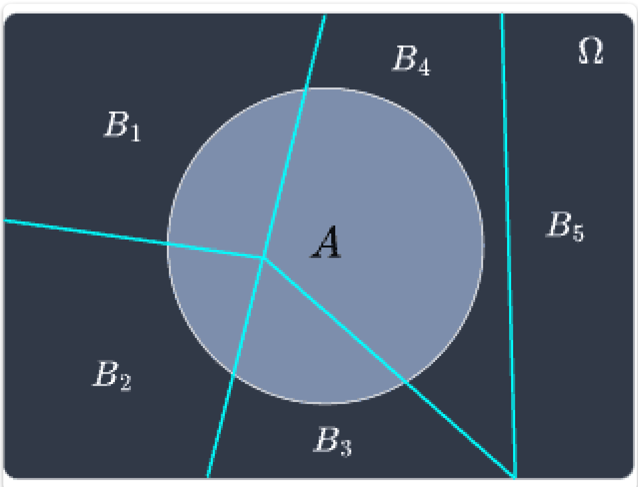
$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$



A Beautiful Explanation on Conditional Probabilities

Total Probability

The concept of *Total Probability* is an important idea in probability theory and statistics. It's often used to find the probability of an event by considering all possible ways that event can occur. The total probability theorem is commonly applied when you have several disjoint cases or conditions that collectively cover all possibilities.



Total Probability Theorem: The total probability theorem states that for any event A, you can find its probability P(A) by summing the probabilities of A occurring within each of several mutually exclusive and exhaustive cases or conditions.

In mathematical terms, the total probability theorem can be expressed as:

$$P(A) = \sum_i [P(A|B_i) \cdot P(B_i)]$$

Where:

- P(A) is the probability of event A.
- B_i represents a set of mutually exclusive conditions or cases.
- $P(A \mid B_i)$ is the probability of event A occurring given condition B_i.
- $P(B_i)$ is the probability of condition B_i occurring.

Example

Let's say you want to find the probability of getting a job offer (A) after a job interview. The outcome depends on three possible scenarios:

- B_1: Interview goes well with a 60% chance $(P(B_1)$ = 0.6
- B_2: Interview goes moderately with a 30% chance $(P(B_2)$ = 0.3
- B_3: Interview goes poorly with a 10% chance $(P(B_3)$= 0.1

Given these probabilities, you can use the total probability theorem to find P(A):

$$P(A) = P(A|B_1) \cdot P(B_1) + P(A|B_2) \cdot P(B_2) + P(A|B_3) \cdot P(B_3)$$

This computes the overall probability of getting a job offer by considering how the interview's outcome depends on the three possible scenarios and the probability of each scenario occurring.

The total probability theorem is valuable in solving problems where the event of interest is influenced by multiple, mutually exclusive conditions or scenarios.

Bayes Theorem

Bayes' Theorem is a fundamental concept in probability theory and statistics that allows you to update the probability of an event based on new evidence or information. It's particularly important in Bayesian statistics, machine learning, and many real-world applications.

Bayes' Theorem: Bayes' Theorem is a way to calculate the conditional probability of an event A, given some new evidence or information B.

It is expressed mathematically as:

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

Where:

- $P(A|B)$ is the conditional probability of event A given evidence B.
- $P(B|A)$ is the conditional probability of evidence B given event A.
- $P(A)$ is the prior probability of event A (the probability of A before considering B).
- $P(B)$ is the prior probability of evidence B (the probability of B before considering A).

Explanation

- $P(A)$ represents our initial belief in the probability of event A before considering any new evidence. This is often called the *prior probability*.
- $P(B)$ is the total probability of evidence B, irrespective of event A. It serves as a normalization factor to make the equation work.
- $P(A|B)$ is the updated probability of A, taking into account the new evidence B. This is often referred to as the *posterior probability*.
- $P(B|A)$ is the probability of observing the evidence B given that event A is true. It's crucial in describing how evidence relates to the event.

Example

Let's say you're in a city where it's known that 1% of the population has a certain rare disease (Event A). There's a test for this disease that's 95% accurate, meaning it correctly detects the disease 95% of the time when it's present and 5% of the time when it's not (Event B). You've just taken the test, and it came back positive. You want to know the probability that you actually have the disease.

- P(A): The prior probability of having the disease is 1% or 0.01.
- P(B | A): The probability that the test is positive when you have the disease is 95% or 0.95.
- P(B | ¬A): The probability that the test is positive when you don't have the disease (false positive rate) is 5% or 0.05.
- P(¬A): The prior probability of not having the disease is 99% or 0.99.

Now, you can use Bayes' Theorem to find P(A | B), the probability of having the disease given a positive test result:

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

Plug in the values:
$P(A \mid B) \approx 0.161$

So, even with a positive test result, your probability of actually having the disease is only about 16.1%. This illustrates how Bayes' Theorem helps you update your probability estimate based on new information, in this case, the test result.

A Nice Explanation of Bayes Theorem
Geometrical Approach to Bayes Theorem
A Quick Proof of Bayes Theorem
Bayes Theorem with Application
Redesigning Bayes Theorem