Basic Probability

Applications of Probability in Data Science

What is Probability?

Experimental Probability v/s Theoretical Probability

Sets and Set Theory (Union and Intersection)

Random Experiment, Sample Space, Event, Trial

Types of Events

- Equally Likely
- Mutually Exclusive
- Collectively Exhaustive
- Independent Event
- Dependent Event

Additive And Multiplicative Rule of Probability

Conditional Probability

Application of Probability in Data Science

Data scientists use probability to estimate the likelihood of different outcomes. For instance, you might calculate the probability of a user converting (making a purchase) on your website after clicking an ad.
In data science, you might use probability distributions and statistical tests to determine the probability that a marketing campaign will lead to a specific level of user engagement

Statistical Inference: Probability theory is the foundation of statistical inference. It's used to estimate parameters, test hypotheses, and make predictions about populations based on sample data.

A/B Testing: When conducting A/B tests to evaluate the impact of changes or interventions, probability helps in determining if the observed differences are statistically significant or due to chance.

Risk Assessment: In fields like finance and insurance, probability is used to assess risk and make decisions related to investments, loans, insurance policies, and more.

Predictive Modeling: Probability is used to create predictive models for various applications, such as predicting customer churn, stock price movements, weather conditions, and more.

Recommendation Systems: Probability is used in recommendation systems to predict user preferences and suggest products, services, or content.

`SQL`

What is Probability?

Layman Explanation : Probability is a way of dealing with uncertainty or randomness. Probability is a measure of the chances of an event.
For example, if you're told there's a 30% chance of rain tomorrow, it means it might rain, but it's not certain.

Technical Explanation : Probability is a measure of how often an event is expected to occur in the long run, basically long term frequency of an event. Example if a coin is flipped a very large number of times, half of the flips will be Heads and half will be Tails.

Calculating Probability : To find the probability of an event, divide the number of ways it can happen by the total possible outcomes (called the sample space).

$$\text{Probability} = \frac{\text{Outcome of Interest}}{\text{All possible outcomes}} = \frac{m}{n}$$

Range and Interpretation:

Range -> [0-1]

- 0: Impossible (Event won't happen)
  - Getting 6 on flipping a coin
- 1: Certain (Event will definitely happen)
  - Getting a Head or a Tail on flipping a coin
- 0.5 or 50%: Equally likely (Event has an even chance of happening or not happening)
  - Getting a head on flipping a coin

DATAPLAY

Experimental Probability vs. Theoretical Probability

Experimental Probability is found by repeating an experiment and observing the outcomes.

$$P(\text{event}) = \frac{\text{number of times event occurs}}{\text{total number of trials}}$$

Example:
A coin is tossed 10 times:
A head is recorded 7 times and a tail 3 times.

$$P(\text{head}) = \frac{7}{10}$$

$$P(\text{tail}) = \frac{3}{10}$$

Theoretical Probability is what is expected to happen based on mathematics

$$P(\text{event}) = \frac{\text{number of favorable outcomes}}{\text{total number of possible outcomes}}$$

Example:
A coin is tossed.

$$P(\text{head}) = \frac{1}{2}$$

$$P(\text{tail}) = \frac{1}{2}$$

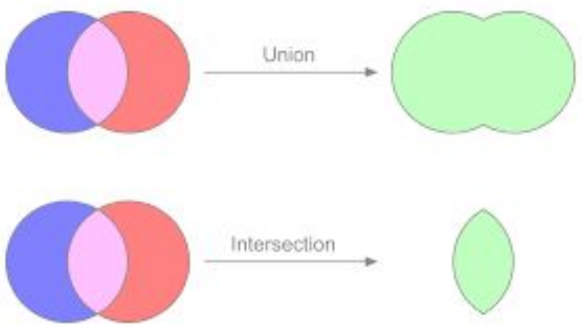
Theoretical probability (classical or a priori probability) is calculated based on the assumption of equally likely outcomes in a well-defined theoretical or mathematical model., while experimental probability (empirical or observed probability) relies on observed data. Both types of probability have their place in solving practical problems, depending on the available information and the nature of the events being studied.

Theoretical probability is expected to match experimental probability when the theoretical model accurately reflects the real-world situation, especially in simple and idealized scenarios.

For example, if you have a fair coin and you flip it many times, the theoretical probability of getting a head is 1/2, and the experimental probability should approach 1/2 if you flip the coin a large number of times. In this case, the theoretical and experimental probabilities are equal because the coin is fair, and all outcomes (H and T) are equally likely.

Sets and Set Theory

Set : A set is a well-defined collection of distinct elements or objects enclosed in curly braces. These elements are considered as members or elements of the set. C = {2, 4, 6, 8} represents a set with 4 elements.



Properties of Set:

1. Each element has to be unique
2. Elements have no specific order

Key components and concepts in set theory include:

Subset: A set A is considered a subset of another set B if every element of A is also an element of B. The symbol for "is a subset of" is ⊆. For example, {2, 4} ⊆ C

Union (OR): The union of two sets A and B, denoted by A ∪ B, is a set containing all the elements that are in A, in B, or in both. For example, {2, 4, 6} ∪ {6,8} = C

Intersection (AND): The intersection of two sets A and B, denoted by A ∩ B, is a set containing all the elements that are in both A and B. For example, {1, 2} ∩ {3,2} = {2}

Complement: The complement of a set A, denoted by A', is a set containing all elements that are not in A but are in the universal set (the set that contains all relevant elements). For example, {2, 4} and {6,8} are complement of each other with C as universal set.

Cardinality: The cardinality of a set is the number of elements in the set. It is denoted by |A|. For Example, |C|=4

Universal Set: The universal set is the set that contains all relevant elements in a particular context.

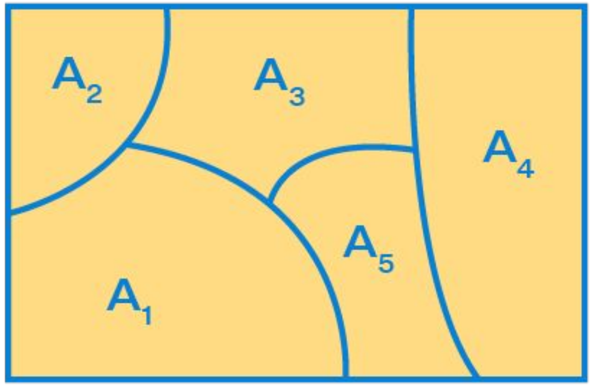
Random Experiment, Sample Space, Event, Trial

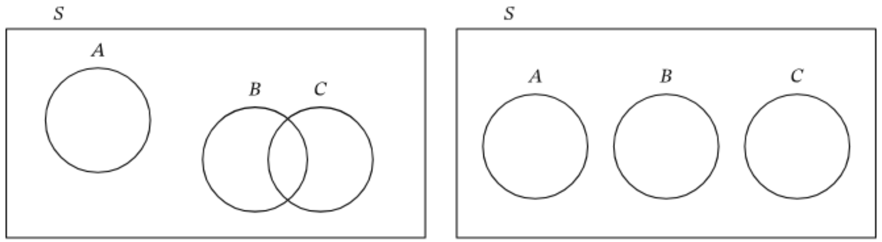
DATAPLAY

Concept	Definition	Example
Random Experiment	A process or procedure with uncertain outcomes i.e. an action or activity where the results or outcomes are not predetermined and can vary each time the process is carried out. In other words, the exact outcome of the process is not known in advance and can differ from one trial to another.	1. Rolling a fair six-sided die 2. Tossing a coin
Sample Space	The set of all possible outcomes of an experiment.	1. Sample space of rolling a die: {1, 2, 3, 4, 5, 6} 2. Sample space of tossing a coin: {H,T}
Event	A subset of the sample space representing a specific outcome or a combination of outcomes.	1. Event "getting an even number": {2, 4, 6} 2. Event "getting a prime number": {2, 3, 5} 3. Event: Getting Heads : {H} 4. Event : Getting a Heads or a Tail : {H,T}
Trial	A trial is a single performance of an experiment, activity, or process that can result in multiple possible outcomes.	1. A single flip of a fair coin represents one trial. The possible outcomes of this trial are "Heads" (H) or "Tails" (T) 2. A single roll of a dice represents one trial. The possible outcomes of this trial are {1, 2, 3, 4, 5, 6}

Type of Events

Event Type	Definition	Example	Formula
Equally Likely Events	Events with equal probabilities of occurring.	Rolling a fair six-sided die, each outcome (1, 2, 3, 4, 5, 6) is equally likely.	$P(E_1) = P(E_2)$ Used in calculating theoretical probabilities of each outcome in various experiments.
Collectively/ Mutually Exhaustive Events	A set of events that together cover the entire sample space i.e. encompass all possible outcomes.	In a deck of playing cards, drawing any card guarantees you will get a card from the deck.	$\bigcup_{i=1 \text{ to } n}(E_i) = S$ The events need not be all mutually exclusive.
Mutually Exclusive Events	Events that cannot occur simultaneously. Occurrence of one event precludes others.	"Heads" and "Tails" are mutually exclusive outcomes while tossing a coin. “K” and “Hearts” while picking a card are not mutually exclusive.	$P(E_1 \text{ and } E_2) = 0$





A is mutually exclusive to *B* and *C*, but *B* and *C* are not mutually exclusive.

A, *B* and *C* are pairwise mutually exclusive.

Type of Events

DATAPLAY

Dependent Events	Events where the outcome of one event affects the outcome of another.	Drawing cards without replacement: The probability of the second card depends on the outcome of the first.	$P(E_1 \mid E_2) = P(E_1 \ \& \ E_2)/\ P(E_2)$ Used in conditional probability to find the probability of one event given the occurrence of another.
Independent Events	Events where the outcome of one event does not affect the outcome of another.	Tossing a coin: Multiple coin tosses are typically independent events.	$P(E_1 \mid E_2) = P(E_1)$ $P(E_1 \ \& \ E_2) = P(E_1)*P(E_2)$

Additive Rule of Probability

The additive and multiplicative rules of probability are fundamental concepts in probability theory that help us calculate the probability of various events. Here's an explanation of these rules and their intuitive interpretations:

Additive Rule of Probability: The additive rule is used to find the probability of the union (either one or both) of two or more events. It allows us to calculate the probability of "either one or both" events occurring. It can be expressed as:

P(A or B) = P(A) + P(B) - P(A and B) {Can be explained by Set theory}

{P(A and B) = 0 in case A and B are mutually exclusive events}

- Intuition: When we calculate P(A or B), we want to find the probability of either event A occurring or event B occurring. To do this, we add the individual probabilities of A and B. However, we must subtract the probability of both A and B happening (the intersection) to avoid double-counting the overlapping region.

  Example : Find probability of getting 2 or even number on a dice roll.

  Answer : P(2 or even) = P(2) + P(even) - P(2 and even) where S = {1,2,3,4,5,6}

  ({2} / S) + ({2,4,6}/ S) - ({2} / S)

  1/6 + 3/6 - 1/6 = 3/6 = ½

  If we had done only P(2) + P(even), we would have counted {2} twice.

Multiplicative Rule of Probability

DATAPLAY

Multiplicative Rule of Probability: The multiplicative rule is used to find the probability of the intersection (both) of two or more independent events. It helps us find the probability of "both" events occurring when they are independent. It can be expressed as:

P(A and B) = P(A) * P(B)   {Can be explained by Set theory}

- Intuition: When we calculate P(A and B), we want to find the probability of both event A and event B occurring simultaneously. To do this, we multiply the individual probabilities of A and B. This makes intuitive sense because if events A and B are independent, the probability of them both happening is the product of their individual probabilities. When events A and B are dependent, you calculate the probability of both events occurring (P(A and B)) using the conditional probability.

These rules are essential for combining probabilities in a wide range of real-world scenarios, including decision-making, risk analysis, and statistics.

Example of Additive and Multiplicative Rule

DATAPLAY

Example 1: Let's say you are rolling a standard six-sided die twice. You want to find the probability of rolling either a 3 or a 4 (event A) and then rolling an even number (event B).

- Additive Rule: To find the probability of event A i.e. probability of rolling either a 3 or a 4, you can use the additive rule:
  P(A or B) = P(A) + P(B)
  P(3 or 4) = P(3) + P(4) = 1/6 + 1/6 = 1/3

- Multiplicative Rule: To find the probability of both event A and event B occurring, you can use the multiplicative rule:
  P(A and B) = P(A) * P(B)
  P((3 or 4) and even) = P(3 or 4) * P(even) = 1/3 * 3/6 = ⅙

Using set theory :

P(3 or 4) = P(3) U P(4) = {3,4} / {1,2,3,4,5,6} = ⅓

P(even) = P(2,4,6) = {2,4,6} / {1,2,3,4,5,6} = ½

P(3 or 4) and P(even) = {(3,2),(3,4),(3,6),(4,2),(4,4),(4,6)} / {(1,1),(1,2),(1,3)......(4,1),(4,2),.....(5,5),(5,6)} = 6/36 = ⅙

Example of Additive and Multiplicative Rule

DATAPLAY

Example 2: Drawing Cards from a Deck Suppose you have a standard deck of 52 playing cards. You want to find the probability of drawing a red card (hearts or diamonds, event A) and then drawing a face card (king, queen, or jack, event B) with replacement.

- Additive Rule: To find the probability of event A occurring, you can use the additive rule:
  P(A) = P(hearts or diamonds) = P(hearts) + P(diamonds) - P(hearts and diamonds)

  P(hearts or diamonds) = 13/52 + 13/52 = ½

  P(B) = P(king, queen, or jack) = 12/52

- Multiplicative Rule: To find the probability of both event A and event B occurring (A and B are independent events), you can use the multiplicative rule:
  P(A and B) = P(A) * P(B)
  P(red and face card) = P(red) * P(face card) = (1/2) * (12/52) = 6/52 ≈ 0.115

  Solve this using set theory too.

Practice

DATAPLAY

Q. Suppose you have a box with 3 blue marbles, 2 red marbles, and 4 yellow marbles. You are going to pull out one marble, record its color, put it back in the box and draw another marble. What is the probability of pulling out a red marble followed by a blue marble? (2/27)

Q. Bag A contains 3 Red Marbles and 4 Blue Marbles. Bag B contains 5 Red Marbles and 3 Blue Marbles. A marble is taken from each bag in turn. What is the probability of getting a marble of each colour? (29/56)

Q. A coin is flipped three times. What is the probability of getting heads on the first flip, tails on the second flip, and heads on the third flip? (1/8)

Q. You have a bag containing 5 green candies and 3 red candies. What is the probability of randomly selecting either a green candy or a red candy from the bag? (1)

Q. Suppose you roll two fair six-sided dice. Given that the sum of the two dice is 7, what is the probability that the first die shows a 4? (1/6)

Q. A password consists of 4 digits, each chosen from 0 to 9. If repetitions are allowed, what is the probability that the password contains the digits 2, 5, 7, and 9 in any order? (0.24%)

Conditional Probability

DATAPLAY

Conditional probability is a measure of the likelihood of an event occurring when we have some additional information or knowledge about another event. It quantifies the probability of one event happening under the condition that we know another event has already occurred.

Generic formula for conditional probability is: $P(A|B) = P(A \cap B) / P(B)$ {Can be explained by Set theory}

Where:

$P(A|B)$ represents the conditional probability of event A occurring given that event B has already occurred.

$P(A \cap B)$ denotes the probability of both events A and B occurring simultaneously.

$P(B)$ represents the probability of event B occurring.

Intuition:

We restrict our attention to the sample space where event B has occurred and then calculate the probability of A within that restricted space. Conditional probability allows us to adjust probabilities based on new information, making it a fundamental concept in probability theory and statistics. It's widely used in various applications, including machine learning, decision making, and real-world problem-solving.

Conditional Probability is the base for Naive Bayes Algorithm.

Example of Conditional Probability

DATAPLAY

Drawing Marbles from a Bag

Suppose you have a bag containing three red marbles (R) and two blue marbles (B). You want to find the probability of drawing a blue marble (B) given that you've already drawn a red marble, without replacement.

Here are the probabilities:

- P(R) is the probability of drawing a red marble first: P(R) = 3/5.
- P(B) is the probability of drawing a blue marble first: P(B) = 2/5.
- Now, you want to calculate the probability of drawing a blue marble (B) given that you've already drawn a red marble (R), denoted as P(B | R).

Using the formula for conditional probability:

P(B | R) = 2/4 = 0.5 (intuitively)

P(B | R) = P(B and R) / P(R)

- P(B and R): The probability of drawing a red marble (R) and then a blue marble (B) in that order. Since you've already drawn the red marble, this is the probability of drawing a blue marble second. P(B and R) = P(B)*P(R|B) = (2/5)*(3/4) = 3/10.
- P(R) is 3/5, as calculated earlier.
- P(B|R) = (3/10)/(3/5) = 0.5

Example

DATAPLAY

Carlos plays college soccer. He makes a goal 65% of the time he shoots. Carlos is going to attempt two goals in a row in the next game. A = the event Carlos is successful on his first attempt. $P(\mathrm{A}) = 0.65$. B = the event Carlos is successful on his second attempt. $P(\mathrm{B}) = 0.65$. Carlos tends to shoot in streaks. The probability that he makes the second goal **GIVEN** that he made the first goal is 0.90.

a. What is the probability that he makes both goals?
b. What is the probability that Carlos makes either the first goal or the second goal?
c. Are A and B independent?
d. Are A and B mutually exclusive?

Solutions

a. The problem is asking you to find $P(\mathrm{A\ AND\ B}) = P(\mathrm{B\ AND\ A})$. Since $P(\mathrm{B}|\mathrm{A}) = 0.90 : P(\mathrm{B\ AND\ A}) = P(\mathrm{B}|\mathrm{A})P(\mathrm{A}) = (0.90)(0.65) = 0.585$

Carlos makes the first and second goals with probability 0.585.

b. The problem is asking you to find $P(\mathrm{A\ OR\ B})$.

$$P(\mathrm{A\ OR\ B}) = P(\mathrm{A}) + P(\mathrm{B}) - P(\mathrm{A\ AND\ B}) = 0.65 + 0.65 - 0.585 = 0.715 \tag{4.3.3}$$

Carlos makes either the first goal or the second goal with probability 0.715.

c. No, they are not, because $P(\mathrm{B\ AND\ A}) = 0.585$.

$$P(\mathrm{B})P(\mathrm{A}) = (0.65)(0.65) = 0.423 \tag{4.3.4}$$

$$0.423 \neq 0.585 = P(\mathrm{B\ AND\ A}) \tag{4.3.5}$$

So, $P(\mathrm{B\ AND\ A})$ is **not** equal to $P(\mathrm{B})P(\mathrm{A})$.

d. No, they are not because $P(\mathrm{A\ and\ B}) = 0.585$.

To be mutually exclusive, $P(\mathrm{A\ AND\ B})$ must equal zero.