**Unit 1:**

|  |
| --- |
| * Introduction of Probability Concept * Types of Probability * Permutation and Combination concept * Addition and Multiplication Theorem * Condition Probability * Bayes’s Theorem |

**Introduction of Probability Concept**

**History of Probability**

Probability theory, as a formal branch of mathematics, traces its roots back to the 16th and 17th centuries. However, the concepts underpinning probability have existed much earlier in the form of gambling and games of chance. Ancient civilizations like the Egyptians, Greeks, and Romans used early forms of probability in divination and decision-making processes. Despite these early instances, it wasn’t until the Renaissance that probability began to be studied systematically.

**The Beginnings: The Renaissance and the Birth of Probability Theory**

The formal study of probability is often attributed to the correspondence between French mathematicians Blaise Pascal and Pierre de Fermat in the 1650s. Their discussions were centered around problems in gambling, such as the “problem of points,” which concerned dividing stakes in a game of chance that was interrupted before its conclusion. This collaboration laid the groundwork for the mathematical theory of probability.

During the same period, the Italian mathematician Gerolamo Cardano, in his work \*Liber de Ludo Aleae\* (The Book on Games of Chance), was one of the first to formalize the calculation of odds and outcomes in gambling. Although Cardano's work was published posthumously, it demonstrated a clear understanding of the principles of probability and laid a foundation for future developments**.**

**The Development of Classical Probability: 17th to 18th Centuries**

The 17th century saw the formalization of probability theory, which continued into the 18th century. One of the key figures during this period was the Dutch mathematician Christiaan Huygens, who, in 1657, published the first book on probability theory titled \*De Ratiociniis in Ludo Aleae\* (On Reasoning in Games of Chance). Huygens' work built upon the ideas of Pascal and Fermat, further formalizing the concepts of expected value and fair games.

In the 18th century, probability theory was further developed by figures such as Jakob Bernoulli and Pierre-Simon Laplace. Bernoulli’s work, \*Ars Conjectandi\* (The Art of Conjecture), published posthumously in 1713, introduced the law of large numbers, a fundamental theorem in probability theory. This theorem states that as the number of trials of a random event increases, the average of the results will converge to the expected value.

Laplace, in his seminal work \*Théorie Analytique des Probabilités\* (Analytical Theory of Probability) published in 1812, provided a comprehensive framework for probability theory and applied it to various fields, including astronomy, statistics, and social sciences. Laplace's definition of probability as the ratio of favorable outcomes to the total number of equally likely outcomes became the foundation of classical probability theory.

**The Expansion of Probability: 19th to Early 20th Centuries**

The 19th century saw probability theory expanding beyond gambling and games of chance into broader applications. The development of statistics and the theory of errors in measurement contributed significantly to the evolution of probability. Karl Friedrich Gauss, in the early 19th century, introduced the concept of the normal distribution, also known as the Gaussian distribution, which became central to probability theory and statistics.

Another major development during this period was the concept of the random walk, introduced by Karl Pearson in 1905, and the notion of Brownian motion, studied by Albert Einstein in 1905. These concepts laid the foundation for the theory of stochastic processes, which would become a significant area of research in the 20th century.

**The Formalization of Probability Theory: 20th Century**

The 20th century marked a significant shift in the formalization and abstraction of probability theory. The Russian mathematician Andrey Kolmogorov played a crucial role in this process. In 1933, Kolmogorov published \*Grundbegriffe der Wahrscheinlichkeitsrechnung\* (Foundations of the Theory of Probability), which established a rigorous axiomatic foundation for probability theory. Kolmogorov’s axioms provided a formal mathematical structure for probability, defining it as a measure on a sigma-algebra of events.

The mid-20th century also saw the application of probability theory in various scientific fields, including quantum mechanics, genetics, economics, and computer science. The development of Bayesian probability, named after the Reverend Thomas Bayes, who introduced Bayes' Theorem in the 18th century, gained significant attention. Bayesian probability provided a framework for updating probabilities based on new evidence, becoming widely used in statistics, decision theory, and machine learning.

**Modern Applications and Ongoing Developments**

Today, probability theory is a cornerstone of modern mathematics and is applied in a wide range of disciplines. From predicting stock market trends to understanding the behavior of subatomic particles, probability theory continues to evolve and find new applications. The development of computational methods and algorithms has further expanded the scope of probability theory, allowing for the analysis of complex systems and large datasets.

The history of probability is a testament to the power of mathematical abstraction and its ability to provide insights into the uncertain and unpredictable aspects of the world. As probability theory continues to evolve, it will undoubtedly play an increasingly important role in shaping our understanding of the world and the decisions we make.

##### Probability Concept : What is Probability?

Probability is a mathematical concept that measures the likelihood or chance of an event occurring. It is a fundamental tool for dealing with uncertainty and is widely used in various fields such as mathematics, statistics, finance, science, engineering, and everyday life. The concept of probability helps us quantify the uncertainty surrounding the outcomes of random processes and make informed predictions about future events.

##### Basic Definition

In its simplest form, probability is defined as the ratio of the number of favorable outcomes to the total number of possible outcomes in a given experiment. It is typically expressed as a fraction, decimal, or percentage, and always falls within the range of 0 to 1, where:

##### - 0 indicates that the event will not occur (impossible event).

##### - 1 indicates that the event will certainly occur (certain event).

##### - Any value between 0 and 1 represents the likelihood of the event occurring, with higher values indicating greater likelihood.

##### Life is full of uncertainties. We don’t know the outcomes of a particular situation until it happens. Will it rain today? Will I pass the next math test? Will my favorite team win the toss? Will I get a promotion in next 6 months? All these questions are examples of uncertain situations we live in. Let us map them to few common terminology which we will use going forward.

##### Experiment – are the uncertain situations, which could have multiple outcomes. Whether it rains on a daily basis is an experiment.

##### Outcome is the result of a single trial. So, if it rains today, the outcome of today’s trial from the experiment is “It rained”

##### Event is one or more outcome from an experiment. “It rained” is one of the possible event for this experiment.

##### Probability is a measure of how likely an event is. So, if it is 60% chance that it will rain tomorrow, the probability of Outcome “it rained” for tomorrow is 0.6.

##### Sample Space - The *set* of all possible outcomes, e.g. we can roll a one, two, three, four, five or six.

##### Mutually Exclusive - Two events are *mutually exclusive* if both cannot occur at the same time, for example, we cannot roll a six and an odd number at the same time.

##### Independent - Two events are *independent* if the occurrence of one does not affect the probability of the other occurring, e.g. rolling a 6 the first time does not affect the probability of rolling a 6 the next time.

##### Why do we need probability?

##### In an uncertain world, it can be of immense help to know and understand chances of various events. You can plan things accordingly. If it’s likely to rain, I would carry my umbrella. If I am likely to have diabetes on the basis of my food habits, I would get myself tested. If my customer is unlikely to pay me a renewal premium without a reminder, I would remind him about it.

##### The mathematical expression for probability is:

##### A black text on a white background Description automatically generated

##### Example

Consider a simple example of rolling a six-sided die. The die has six faces, numbered from 1 to 6. If we want to calculate the probability of rolling a 4, we can determine the following:

##### - Total number of possible outcomes: There are 6 possible outcomes (1, 2, 3, 4, 5, and 6).

##### - Number of favorable outcomes: There is 1 favorable outcome (rolling a 4).

##### Thus, the probability of rolling a 4 is:

##### A black text on a white background Description automatically generated

##### This means there is approximately a 16.7% chance of rolling a 4.

**Types of Probability**

##### Probability can be interpreted and calculated in several ways depending on the context:

##### 1. Classical Probability: This is the traditional approach where all outcomes are assumed to be equally likely. It is often used in situations involving games of chance, such as flipping a coin or rolling a die.

##### 2. Empirical Probability: This type of probability is based on experimental data or observed frequencies. It is calculated by dividing the number of times an event occurs by the total number of trials. For example, if you flip a coin 100 times and it lands on heads 60 times, the empirical probability of getting heads is 60/100 = 0.6.

##### 3. Subjective Probability: Subjective probability is based on personal judgment, experience, or intuition rather than precise calculations. It is often used when there is no historical data or when the probability is difficult to measure. For instance, a doctor may estimate the probability of a patient recovering from an illness based on their experience with similar cases.

##### 4. Bayesian Probability: Bayesian probability involves updating the probability of an event as new evidence or information becomes available. It is grounded in Bayes' Theorem, which provides a mathematical framework for revising probabilities considering new data.

##### Applications of Probability

##### Probability is a versatile tool with a wide range of applications:

##### - Statistics : Probability forms the basis for many statistical methods, including hypothesis testing, confidence intervals, and regression analysis. It helps statisticians make inferences about populations based on sample data.

##### - Finance : In finance, probability is used to model and assess risks, price financial instruments, and develop investment strategies. Techniques like Monte Carlo simulations rely on probabilistic models to predict market behavior.

##### - Science and Engineering : Probability is essential in scientific research for analyzing experimental data and modeling natural phenomena. Engineers use probability to assess system reliability, manage risks, and optimize processes.

##### - Everyday Life : People use probability, often unconsciously, in everyday decision-making. For example, when deciding whether to carry an umbrella based on a weather forecast, you are considering the probability of rain.

##### Types of Probability – Detailed Explanation with Practical Examples

Probability is a versatile concept in mathematics that can be interpreted and applied in various ways depending on the context. Understanding the different types of probability is crucial for effectively applying probabilistic reasoning in diverse fields such as statistics, finance, engineering, and everyday decision-making. This section delves into the primary types of probability, providing detailed explanations and practical examples for each.

##### 1. Classical Probability

##### Definition: Classical probability, also known as "a priori" or "theoretical probability," is based on the assumption that all possible outcomes of an experiment are equally likely. It is calculated using the ratio of the number of favorable outcomes to the total number of possible outcomes.

##### Formula:

##### A black text on a white background Description automatically generated

##### Key Characteristics:

##### Relies on known and finite sample spaces.

##### Assumes perfect randomness and equal likelihood of all outcomes.

##### Often used in games of chance and combinatorial problems.

##### Practical Examples:

##### 1. Rolling a Fair Die:

##### - Experiment: Rolling a standard six-sided die.

##### - Sample Space: {1, 2, 3, 4, 5, 6}

##### - Event : Rolling an even number (2, 4, 6).

##### - Probability Calculation:

##### A black and white math equation Description automatically generated

##### 

##### - Interpretation: There is a 50% chance of rolling an even number.

##### 2. Flipping a Fair Coin:

##### - Experiment: Flipping a coin.

##### - Sample Space: {Heads, Tails}

##### - Event: Getting Heads.

##### - Probability Calculation:

##### A black text on a white background Description automatically generated

##### - Interpretation: There is a 50% chance of the coin landing on Heads.

##### 3. Drawing a Card from a Standard Deck:

##### - Experiment: Drawing one card from a standard 52-card deck.

##### - Sample Space: 52 unique cards.

##### - Event: Drawing an Ace.

##### - Probability Calculation:

##### A number with numbers and lines Description automatically generated with medium confidence

##### - Interpretation: There is approximately a 7.7% chance of drawing an Ace.

##### 4. Spinning a Fair Spinner:

##### - Experiment: Spinning a spinner divided into 8 equal sectors numbered 1 through 8.

##### - Sample Space: {1, 2, 3, 4, 5, 6, 7, 8}

##### - Event: Spinner lands on a number greater than 5 (i.e., 6, 7, 8).

##### - Probability Calculation:

##### A number and a number Description automatically generated with medium confidence

##### - Interpretation: There is a 37.5% chance of the spinner landing on a number greater than 5.

##### ---

##### 2. Empirical Probability

##### Definition:

##### Empirical probability, also known as "experimental" or "a posteriori" probability, is based on observed data or experiments rather than theoretical calculations. It is determined by conducting experiments or collecting data and calculating the relative frequency of the event occurring.

##### A black text on a white background Description automatically generated

##### Key Characteristics:

##### - Relies on actual experiments or historical data.

##### - Can accommodate complex and non-uniform sample spaces.

##### - Useful when theoretical probabilities are difficult to determine.

##### Practical Examples:

##### 1. Weather Forecasting:

##### - Experiment: Recording daily occurrences of rain over a year.

##### - Data: Suppose it rained 120 days out of 365.

##### - Probability Calculation:

##### A number and a mathematical equation Description automatically generated with medium confidence

##### - Interpretation: Based on past data, there is approximately a 32.9% chance of rain on any given day.

##### 2. Quality Control in Manufacturing:

##### - Experiment: Inspecting 1,000 units produced by a factory.

##### - Data: Found 50 defective units.

##### - Probability Calculation:

##### A number with a smile Description automatically generated with medium confidence

##### - Interpretation: There is a 5% probability that a randomly selected unit is defective.

##### 3. Sports Performance:

##### - Experiment: Tracking a basketball player's free-throw success rate over 200 attempts.

##### - Data: Successfully made 150 free throws.

##### - Probability Calculation:

##### A number with numbers and symbols Description automatically generated with medium confidence

##### - Interpretation: The player has a 75% probability of making a free throw based on past performance.

##### 4. Epidemiology Studies:

##### - Experiment: Observing the occurrence of a particular disease in a population over a decade.

##### - Data: 300 out of 10,000 individuals developed the disease.

##### - Probability Calculation:

##### A number with numbers and lines Description automatically generated with medium confidence

##### - Interpretation: There is a 3% probability of an individual in the population developing the disease based on historical data.

##### 3. Subjective Probability

##### Definition:

##### Subjective probability is based on personal judgment, intuition, or experience rather than on formal calculations or empirical data. It reflects an individual's degree of belief in the occurrence of an event.

##### Key Characteristics:

##### - Influenced by personal opinions, biases, and experiences.

##### - Not necessarily quantifiable or consistent across different individuals.

##### - Useful in scenarios where objective data is unavailable or incomplete.

##### Practical Examples:

##### 1. Investment Decisions:

##### - Scenario: An investor assesses the likelihood of a stock's price increasing based on their intuition and market experience.

##### - Subjective Probability: The investor believes there is a 70% chance the stock will rise, based on their analysis of market trends and company performance.

##### - Interpretation: The probability is a personal estimate and may differ from objective measures.

##### 2. Medical Diagnoses:

##### - Scenario: A doctor estimates the probability that a patient has a specific disease based on symptoms and medical history.

##### - Subjective Probability: The doctor believes there is an 80% chance the patient has the disease, informed by their clinical experience.

##### - Interpretation: The probability reflects the doctor's judgment and may be adjusted with further tests.

##### 3. Project Management:

##### - Scenario: A project manager assesses the likelihood of a project being completed on time.

##### - Subjective Probability: Based on team performance and project complexity, the manager estimates a 60% probability of on-time completion.

##### - Interpretation: The estimate relies on the manager’s experience and perception of project dynamics.

##### 4. Personal Decision-Making:

##### - Scenario: Deciding whether to carry an umbrella based on the forecast and personal judgment.

##### - Subjective Probability: Believing there is a 40% chance of rain based on the weather forecast and personal observation of cloud patterns.

##### - Interpretation: The decision is influenced by both objective data and personal intuition.

##### ---

##### 4. Bayesian Probability

##### Definition:

##### Bayesian probability is an interpretation of probability that incorporates prior knowledge or beliefs and updates them as new evidence becomes available. It is grounded in Bayes' Theorem, which provides a mathematical framework for revising probabilities.

##### Formula (Bayes' Theorem):

##### A mathematical equation with black text Description automatically generated

##### Where:

##### A white background with black text Description automatically generated

##### Key Characteristics:

##### - Combines prior beliefs with new evidence.

##### - Provides a dynamic approach to probability updating.

##### - Widely used in fields like statistics, machine learning, and decision theory.

##### Practical Examples:

##### 1. Medical Testing:

##### - Scenario: Determining the probability that a patient has a disease given a positive test result.

##### - Prior Probability (P(Disease)): 1% (prevalence of the disease).

##### - Likelihood (P(Positive|Disease)): 99% (test accuracy).

##### - Marginal Probability (P(Positive)): Calculated based on overall prevalence and test accuracy.

##### - Bayesian Calculation:

##### A black text on a white background Description automatically generated

##### - Interpretation: Even with a positive test, the posterior probability may remain low due to the disease's low prevalence.

##### 2. Spam Email Filtering:

##### - Scenario: Updating the probability that an email is spam based on the presence of certain keywords.

##### - Prior Probability (P(Spam)): 20%.

##### - Likelihood (\P(Keyword|Spam)): 80%.

##### - Likelihood (P(Keyword|Not Spam)): 10%.

##### - Bayesian Calculation:

##### A black text with a white background Description automatically generated

##### - Interpretation: The posterior probability of an email being spam increases if it contains specific keywords.

##### 3. Quality Assurance:

##### - Scenario: Estimating the probability that a product is defective after an initial test result.

##### - Prior Probability (P(Defective)): 5%.

##### - Likelihood (P(Passed Test|Defective)): 20%.

##### - Likelihood (P(Passed Test|Not Defective)): 95%.

##### - Bayesian Calculation:

##### A black text on a white background Description automatically generated

##### - Interpretation: Even if a product passes the test, there remains a non-zero probability that it is defective, adjusted based on test characteristics.

##### 4. Legal Proceedings:

##### - Scenario: Assessing the probability of a defendant's guilt based on new evidence.

##### - Prior Probability (P(Guilt)): 10% (based on initial evidence).

##### - Likelihood (P(New Evidence|Guilt)): 90%.

##### - Likelihood (P(New Evidence|Innocent)): 30%.

##### - Bayesian Calculation:

##### A black text on a white background Description automatically generated

##### - Interpretation: The new evidence significantly increases the probability of guilt compared to the prior probability.

##### ---

##### 5. Frequentist Probability (Additional Type)

##### Definition:

##### Frequentist probability defines the probability of an event as the limit of its relative frequency in many trials. It interprets probability strictly in terms of long-run frequencies of events.

##### Key Characteristics:

##### - Objective interpretation based on long-term frequencies.

##### - Does not incorporate prior beliefs or subjective opinions.

##### - Commonly used in classical statistical inference.

##### Practical Examples:

##### 1. Coin Tossing Experiments:

##### - Scenario: Tossing a fair coin 10,000 times.

##### - Frequency Calculation: If heads appear 5,002 times, the frequentist probability of heads is:

##### A number with numbers and a line Description automatically generated with medium confidence

##### - Interpretation: The probability of getting heads is approximately 50%, aligning with the theoretical probability.

##### 2. Manufacturing Defects:

##### - Scenario: Monitoring the defect rate in a production line over time.

##### - Frequency Calculation: Out of 50,000 units produced, 250 are defective.

##### A number and a circle with numbers Description automatically generated with medium confidence

##### - Interpretation: The frequentist probability of producing a defective unit is 0.5%.

##### 3. Elections Polling:

##### - Scenario: Conducting a poll with 1,000 respondents to estimate voter preference.

##### - Frequency Calculation: If 600 respondents favor Candidate A, the frequentist probability is:

##### A black text on a white background Description automatically generated

##### - Interpretation: Based on the poll, there is a 60% probability that a randomly selected voter favors Candidate A.

##### 4. Quality Control Testing:

##### - Scenario: Testing batches of products to determine the failure rate.

##### - Frequency Calculation: In 200 tested units, 4 fail.

##### A number and a number Description automatically generated with medium confidence

##### - Interpretation: The frequentist probability of a unit failing is 2%.

##### Understanding the different types of probability—Classical, Empirical, Subjective, Bayesian, and Frequentist—is essential for accurately modeling and interpreting uncertainty in various contexts. Each type offers unique perspectives and tools for assessing likelihoods, making informed decisions, and conducting rigorous analyses. By applying the appropriate type of probability based on the nature of the problem and the available information, one can effectively navigate and quantify uncertainty in both theoretical and practical scenarios.

##### Example

##### In Newcastle, 70% of small businesses use the internet to advertise new products; 50% of small businesses use flyers to advertise new products and a quarter of small businesses use *both* flyers *and* the internet.

##### (A) What is the probability that a randomly chosen small business in Newcastle uses *either* flyers *or* the internet to advertise new products?

##### (B) What is the proportion of small businesses in Newcastle that use neither the internet *nor* flyers to advertise new products?

##### Solution (A)

##### Let F denote the event that a business advertises new products using flyers and I denote the event that a business uses the internet to advertise new products.

##### We wish to find P(F or I). Using the Addition Law, we have:

##### 

##### There is a 95% probability that a randomly chosen business in Newcastle uses either flyers or the internet to advertise new products

##### Tree Diagrams

##### *Tree Diagrams* can be used to help us to visualize and calculate complex probabilities. When drawing a tree diagram we begin with a dot. From this dot lines (“branches”) are then drawn, extending from the right of the first dot, to represent all possible outcomes for the given situation. The probabilities of each of these outcomes is written just above the corresponding line.

##### To calculate the probability that two events *both* happen, we draw another “branch” extending from the “branch” corresponding to one of these events to represent the second event occurring after the first. Above this line we write the probability (or conditional probability for events which are not independent) of the second event occurring after the first. Multiplying these probabilities “along the branches” gives the required probability.

##### To calculate the probability that one or both of two *independent* events occurs we add the probabilities of the two events “down the columns”.

##### Example

##### 60% of employees at a department store in Newcastle are women. Government research into methods of commuting to city jobs in the North East has shown on average that:

##### 12% of people cycle into work.

##### A quarter of the people drive.

##### 10% of people walk.

##### And the rest use public transport.

##### What is the probability that a randomly selected employee of the department store in Newcastle commutes using public transport and is male? Now calculate the probability that a randomly selected employee is female and drives into work.

##### Solution We can use a tree diagram to present all of the information given to us and calculate the required probability.

##### |center

##### The probabilities in blue are calculated using the multiplication law. So, the probability the employee is female and drives is 0.6×0.25=0.15

##### Tip: To make sure all your calculations are correct, you can check to see that your final probabilities (the blues ones) add up to 11. This must be the case because at least one of all of the possible events *must* (is certain to) occur.

##### Decision Trees - *Decision trees* are very similar to the probability tree diagrams but are used specifically to calculate expected monetary values.

##### Example 4

##### The manager of a small business has the opportunity to buy a fixed quantity of a new product and offer it for sale for a limited time.

##### There will be a fixed cost of £100,000 to buy the product and offer it for sale. The amount of the product that the manager would be able to sell is not certain but market research has suggested that:

##### The probability that sales would be “poor” is 0.25. Selling this quantity would raise an income of £75,000.

##### The probability that sales would be “medium” is 0.6. Selling this quantity would raise an income of £110,000.

##### The probability that sales would be “good” is 0.15. Selling this quantity would raise an income of £145,000.

##### The product can be sold for a trial period before a final decision is made and it costs £18,000 to run the trial. The results of the trial will be “poor” with probability 0.35, “medium” with probability 0.4 or “good” with probability 0.25. Knowing the outcome of the trial changes the probabilities for the main sales project:

##### A table with numbers and text Description automatically generated with medium confidence

##### The manager will make decisions based on expected monetary value.

##### (A) Draw a decision tree for this problem.

##### (B) What is the EMV of a decision to go ahead with the product without a trial?

##### (C) Complete the solution of the decision problem and determine the optimal course of action for the company.

##### Solution (A)

##### |center

##### The values in the blue boxes are the final incomes from buying the new product when sales are “poor”, “medium” or “good” (top to bottom).

##### Solution (B)

##### To calculate the expected monetary value, we need to utilize the formula in the pink box above.

##### For No Trial:

##### EMV=0.25×£75,000+0.6×£110,000+0.15×£145,000=£106,500.

##### The manager has an expected income of £106,500 from selling the new product without a trial.

##### Solution (C)

##### To solve the decision problem it is best to first calculate the seperate EMVEMVs for when a trial is run and when a trial is not run. We must then compare the EMVEMVs for each option (trial or no trial) and choose the option with the highest EMVEMV. This is the optimal course of action for the company.

##### When a trial is carried out and has a poor result:

##### EMV=0.75×£75,000+0.15×£110,000+0.1×£145,000=£87,250.

##### When a trial is carried out and has a medium result:

##### EMV=0.25×£75,000+0.5×£110,000+0.25×£145,000=£110,000.

##### When a trial is carried out and has a good result:

##### EMV=0.1×£75,000+0.15×£110,000+0.75×£145,000=£132,750.

##### Now to calculate the overall EMV we multiply each of these by their associated probabilities:

##### EMV=P(Poor result) × £87,250 + P(Medium result) × £110,000 + P(Good result)×£132,750

##### =£107,725.

##### We now need to calculate the expected profit (or loss) the business would make from each option (trial or no trial).

##### No Trial:

##### {Expected Profit} ={EMV} { for no trial} - {Cost of new product}

##### =£106,500 - £100,000

##### =£6,500.

##### So if the manager goes ahead with the product without a trial, the expected profit is £6,500.

##### Trial:

##### {Expected Profit} = {EMV}{ for trial} - {Cost of new product} - {Cost of trial}\

##### =£107,725 - £100,000 - £18,000

##### =-£10,275.

##### With the trial, there will be an expected loss of £10,275.

##### From these results we can see that optimal course of action for the company is to sell the new product but without the trial period as this yields a higher EMV. It is important to note that although the *expected* monetary value is higher when the manager chooses not to run the trial, the realized profit or loss may be or may not be better than it would have been if a trial had been carried out.

Permutation and Combination concept

##### What is a Permutation?

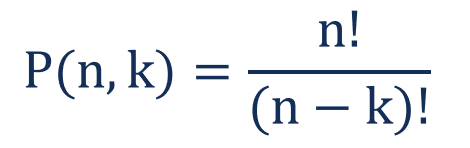
##### A permutation is a mathematical technique that determines the number of possible arrangements in a set when the order of the arrangements matters. Common mathematical problems involve choosing only several items from a set of items in a certain order. Permutations in probability theory and other branches of mathematics refer to sequences of outcomes where the order matters. For example, 9-6-8-4 is a permutation of a four-digit PIN because the order of numbers is crucial. When calculating probabilities, it’s frequently necessary to calculate the number of possible permutations to determine an event’s probability. A permutation is an arrangement of all or part of a set of objects, with consideration of the order of arrangement. The concept of permutation is used when the arrangement order matters. For example, the permutation of letters "ABC" is different from "CAB" because the order of letters differs.



##### Permutations are frequently confused with another mathematical technique called [combinations](https://corporatefinanceinstitute.com/resources/data-science/combination/). However, in combinations, the order of the chosen items does not influence the selection. In other words, the arrangements ab and be in permutations are considered different arrangements, while in combinations, these arrangements are equal.

**Formula for Calculating Permutations**

The general permutation formula is expressed in the following way:



Where:

* **n** – the total number of elements in a set
* **k** – the number of selected elements arranged in a specific order
* **!** – factorial

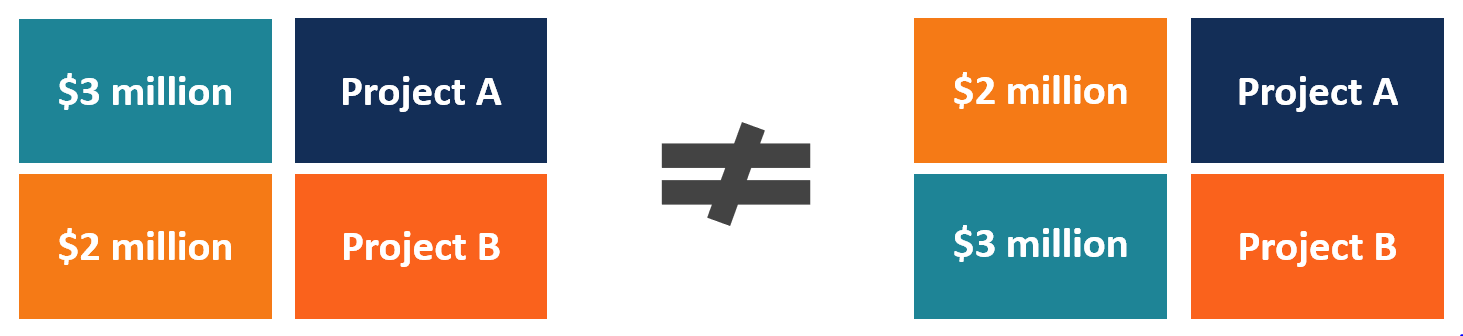
##### Factorial (noted as “!”) is the product of all positive integers less than or equal to the number preceding the factorial sign. For example, 3! = 1 x 2 x 3 = 6.

##### The formula above is used in situations when we want to select only several elements from a set of elements and arrange the selected elements in a special order.

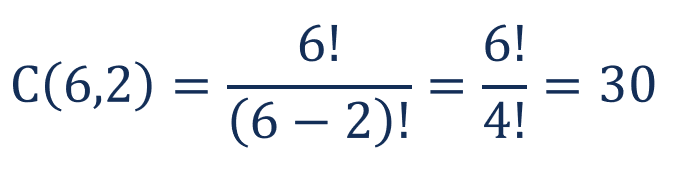
**Example of a Permutation**

##### You are a partner in a [private equity firm](https://corporatefinanceinstitute.com/topic/career/). You want to invest $5 million in two projects. Instead of equal allocation, you decided to invest $3 million in the most promising project and $2 million in the less promising project. Your analysts shortlisted six projects for potential [investment](https://www.thebalancemoney.com/types-of-investments-in-small-business-357246). How many possible arrangements are available for your investment decision?

##### The example above is a permutation problem. Since the allocation of the money for the two [projects](https://corporatefinanceinstitute.com/resources/management/project-evaluation-review-technique-pert/) is not equal, the selection order matters in this problem. For example, consider the following arrangement: invest $3 million in Project A and $2 million in Project B vs. invest $2 million in Project A and $3 million in Project B. The options are not equal to each other. Therefore, we must use the formula above to determine the number of available arrangements:



##### Using the formula above, we can determine the number of available arrangements:



##### Therefore, you can get 30 possible investment arrangements based on the six projects shortlisted by your [analysts](https://corporatefinanceinstitute.com/resources/career-map/corporates/corp-fpa/budget-analyst/).

**Permutations of Distinct Objects**

Permutations of distinct objects refer to the various ways in which a set of distinct (unique) objects can be arranged or ordered. In combinatorial mathematics, a permutation is an arrangement of all the elements in a specific sequence or order. When the objects are distinct, each arrangement is considered unique because the order of objects matters.

**Formula for Permutations of Distinct Objects**

For n distinct objects, the number of possible permutations is given by n! (n factorial). The factorial of a number n is the product of all positive integers from 1 to n.

Mathematically, n! is expressed as: n!=n×(n−1)×(n−2)×⋯×2×1

When all the objects in the set are distinct, the permutation formula is straightforward.

**Example 1:** Consider the set {A, B, C}. The number of ways to arrange all three letters is:

A number with numbers and a line

Description automatically generated with medium confidence

The possible permutations are: ABC, ACB, BAC, BCA, CAB, and CBA.

**Example 2:** For a set of 5 different books, the number of ways to arrange 2 books on a shelf is:

A number and a line

Description automatically generated with medium confidence

**Permutations with Repetition**

Permutations with repetition involve arranging items where some items may be repeated, and we're interested in the number of different sequences that can be formed considering the repetitions.

For example, on a pizza, you might have a combination of three toppings: pepperoni, ham, and mushroom. The order doesn’t matter. For example, using letters for the toppings, you can have PHM, PMH, HPM, and so on. It doesn’t matter for the person who eats the pizza because you have the same combination of three toppings. In other words, the order of these three letters does not matter and they form one combination.

However, imagine we’re using those letters for a weak password. In this case, the order is crucial, making them permutations. PHM, PMH, HPM, etc., are distinct permutations. If the password is PHM, entering HPM will not work. When you have at least two permutations, the number of permutations is greater than the number of combinations. Learn more about the [differences between permutation vs combination](https://statisticsbyjim.com/probability/permutation-vs-combination/).



This type of lock should be known as a permutation lock because the order of digits matters!

**Concept of Permutations with Repetition**

When we allow for repetitions, the problem often involves arranging items where the order is still important, but items are not necessarily unique. The key difference from permutations of distinct objects is that here, we may have multiple arrangements involving the same items.

When the outcomes in a permutation can repeat, [statisticians](https://statisticsbyjim.com/glossary/statistics/) refer to it as permutations with repetition. For example, in a four-digit PIN, you can repeat values, such as 1-1-1-1. Analysts also call this permutations with replacement.

To calculate the number of permutations, take the number of possibilities for each event and then multiply that number by itself X times, where X equals the number of events in the sequence. For example, with four-digit PINs, each digit can range from 0 to 9, giving us 10 possibilities for each digit. We have four digits. Consequently, the number of permutations with repetition for these PINs = 10 \* 10 \* 10 \* 10 = 10,000.

Imagine that a class with 15 children can choose one cookie from five types of cookies: Gingerbread, Sugar, Chocolate Chip, Mint, and Peanut Butter. There are enough cookies that they are free to choose one of any type. How many possible permutations of cookies are there?



In this example,

* n = 5 because there are five possible cookie choices.
* r = 15 because there are 15 students in the class, making it the size of the permutation.

Consequently, the are 515 = 30,517,578,125 permutations with repetition. That’s over 30 billion permutations! If you were to make random guesses for the cookie choice of all 15 children, you’d have a probability of 1/30,517,578,125 of correctly guessing the selections for the entire class! That assumes you don’t have insider knowledge about each child’s cookie preference! I think you’d have better luck in a lottery!

**Formula for Permutations with Repetition**

If we have n types of items, and we are arranging r items (where repetition of items is allowed), the number of possible permutations is given by:

nr

where:

* n is the number of distinct types of items.
* r is the number of positions to fill.

**Examples**

**Example 1: 3 Types of Items, 2 Positions**

Suppose we have 3 types of items (say, A, B, and C) and want to arrange 2 items.

To find the number of permutations:

**Apply the formula:** nr=32=9

**List the permutations:**

* + AA
  + AB
  + AC
  + BA
  + BB
  + BC
  + CA
  + CB
  + CC

There are 9 unique ways to arrange 2 items where each position can be filled with any of the 3 types of items.

**Example 2: 4 Types of Items, 3 Positions**

Consider 4 distinct items (say, 1, 2, 3, and 4), and we want to arrange 3 items.

**Apply the formula:** nr=43=64

**List a few permutations (for illustration):**

* + 111
  + 112
  + 113
  + 121
  + 122
  + 123
  + (and so on...)

There are 64 possible arrangements of 3 items with 4 possible choices for each position.

Permutations with repetition involve:

* **Permutations with Repetition (Formula):** nrn^rnr, where nnn is the number of distinct items, and rrr is the number of positions to be filled.
* **Order Matters:** The arrangement of items is significant, so permutations consider the sequence of items.
* **Repetition Allowed:** The same item can appear in multiple positions.

This concept is widely applicable in scenarios like password generation, where each position can be filled by any of the allowed characters, and in scenarios where choices are repeated multiple times.

**Permutations without Repetition**

When the outcomes cannot repeat, statisticians call them permutations without repetition. This situation frequently occurs when you’re working with unique physical objects that can occur only once in a permutation. Imagine you have 10 different books and want to calculate how many possible ways you can arrange them on a bookshelf. After you place the first book, the second book must be a different book. Consequently, this is an example of permutations without repetition. Analysts also call these permutations without replacement.



For the first book, you have 10 books from which to choose. For the second book, you have nine. There are eight options for the third book, and so on. Like before, this process involves multiplying the number of possible outcomes together. However, we must reduce the number of outcomes for each subsequent event.

Mathematically, we’d calculate the permutations for the book example using the following method:

10 \* 9 \* 8 \* 7 \* 6 \* 5 \* 4 \* 3 \* 2 \* 1 = 3,628,800

There are 3,628,800 permutations for ordering 10 books on a shelf without repeating books.

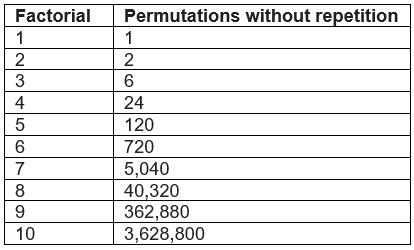
Whew! I bet you didn’t realize there we so many possibilities with 10 books. I’ll stick to alphabetical order!

**Using Factorials for Permutations**

When you multiply all numbers from 1 to n, it’s a factorial. In the book example, we multiplied all numbers from 1 to 10. Instead of using the long string of multiplication, you can write it as 10! and read it as 10 factorial.

In general, n! equals the product of all numbers up to n. For example, 3! = 3 \* 2 \* 1 = 6. The exception is 0! = 1, which simplifies equations.

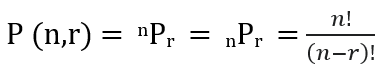
Factorials are crucial concepts for permutations without replication. The number of permutations for n unique objects is n!. This number snowballs as the number of items increases, as the table below shows.



**Partial Permutations without Repetition**

In some cases, you want to consider only a portion of the possible permutations. In the bookshelf example, we wanted to know the total number for 10 books. But what if we could fit only five of the 10 books on the shelf? How many permutations of five books are possible using our 10 books?

Use the following formula to calculate the number of arrangements of r items from n objects. There are several standard methods that statisticians use to notate permutations without repetition, which I show below with the formula.



Where:

* n = the number of unique items. For instance, n = 10 for the book example because there are 10 books.
* r = the size of the permutation. For example, r = 5 for the five books we want to place on the shelf.

This equation works both for the complete and partial sets of permutations without repetitions, depending on the values you enter in the equation. For complete sets, n = r. Additionally, r cannot be greater than n because there are no repetitions.

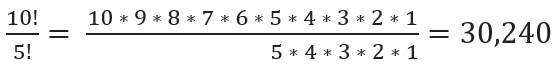
For the book example, we have 10 books, but we can put only five on the shelf. The first book still has 10 options. However, for placing the second book, we have only nine options because we already placed one. We have eight options for the third book and so on until we place the fifth book. Mathematically, we’d write this as the following for the five books:

10 \* 9 \* 8 \* 7 \* 6 = 30,240

There are 30,240 permutations for placing five books out of our 10 books on a shelf.

Using the equation to calculate the number of permutations

Now, we’ll use the formula to calculate this example. Again, we’ll use n=10 and r=5.



Notice how the 5! cancels itself out in the fraction? That leaves us with the 10 \* 9 \* 8 \* 7 \* 6 that we had before.

Here’s how the equation works. The numerator calculates the complete number of permutations for all the unique items. The denominator cancels out the permutations in which we’re not interested. For the book example, the denominator cancels out permutations with more than five books.

Using one form of the notation, we’d write this problem as P (10, 5) = 30,240.

Worked Example of Using Permutations to Calculate Probabilities

When you’re given a probability problem that uses permutations, you need to follow these steps to solve the problem.

1. Set up a ratio to determine the probability.
2. Determine whether the numerator and denominator require combinations, permutations, or a mix? For this post, we’ll stick with permutations.
3. Are these permutations with repetitions, without, or a mix?
4. Both types of repetition require you to identify the n and r to enter into the equations.

**Problem: What is the probability that a four-digit PIN does not have repeated digits?**

This question builds on several of the examples in this post.

Let’s set up our ratio for the probability. In this example, we can use the following ratio for the events of interests and the total number of events.

Probability formula for permutations.

**Numerator**

Let’s tackle the numerator. We need to find the number of four-digit PINs that do not have repeating digits. That’s a permutation because order matters, and it’s without replication because we can’t have repeats. Let’s identify the n and r. We’ll use n=10 because 10 digits are available for the first item and r=4 because we’re discussing four-digit PINs.

Let’s enter that into the equation for permutations without repetition to calculate the numerator:

Permutations without repetition for the PIN example.

**Denominator**

For the denominator, we need to calculate all possible permutations for four-digit PINs with repeats. We need to enter our n and r into the equation for permutations with repeats.

nr = 104 = 10,000

Consequently, the probability of a four-digit PIN with no repeating digits equals the following:

A black numbers on a white background

Description automatically generated

Just over half of all four-digit PINs have repeated values.

The birthday problem is a classic probability problem. What’s the smallest size group that has a greater than 50% chance of people sharing a birthday? Solving this problem uses similar methods. Read my post about [Solving the Birthday Problem](https://statisticsbyjim.com/fun/birthday-problem/) to find out!

**Circular Permutation**

Circular permutations refer to arrangements of objects in a circle where the order of the objects is important, but rotations of the arrangement are considered identical. This concept is useful in problems where the arrangement is cyclic and you want to count distinct configurations that are rotationally unique.

**Formula for Circular Permutations**

For n distinct objects arranged in a circle, the number of distinct circular permutations is given by:

A black math symbol with a white background

Description automatically generated

Here’s why: In a linear permutation, n objects can be arranged in n! ways. However, in a circular arrangement, each unique arrangement can be rotated n different ways (all considered the same circular permutation). Therefore, to find the number of distinct circular permutations, you divide the total number of linear permutations by n.

**Examples**

**Example 1: 3 Distinct Objects**

Consider 3 distinct objects: A, B, and C.

1. **Calculate the number of circular permutations:** (n−1)!=(3−1)!=2!=2
2. **List the permutations:**
   * ABC
   * BCA
   * CAB

Here, rotations of ABC (like BCA and CAB) are considered the same circular permutation. Thus, there are 2 unique circular permutations.

**Example 2: 4 Distinct Objects**

Consider 4 distinct objects: 1, 2, 3, and 4.

1. **Calculate the number of circular permutations:** (n−1)!=(4−1)!=3!=6
2. **List the permutations:**
   * 1234
   * 2341
   * 3412
   * 4123
   * 3412
   * 4123

These are the unique circular permutations, where each arrangement is considered identical under rotation.

**Example 3: 5 Distinct Objects**

Consider 5 distinct objects: A, B, C, D, and E.

1. **Calculate the number of circular permutations:** (n−1)!=(5−1)!=4!=24
2. **List the permutations:**
   * ABCDE
   * BCDEA
   * CDEAB
   * DEABC
   * EABCD
   * (and so forth)

Each permutation is distinct in a circular arrangement, and there are 24 unique ways to arrange 5 objects in a circle.

**Example 4: 6 Distinct Objects**

Consider 6 distinct objects: 1, 2, 3, 4, 5, and 6.

1. **Calculate the number of circular permutations:** (n−1)!=(6−1)!=5!=120
2. **List the permutations:**
   * 123456
   * 234561
   * 345612
   * 456123
   * 561234
   * 612345
   * (and so on)

There are 120 unique circular permutations for 6 distinct objects.

**Example 5: 7 Distinct Objects**

Consider 7 distinct objects: A, B, C, D, E, F, and G.

1. **Calculate the number of circular permutations:** (n−1)!=(7−1)!=6!=720(n - 1)! = (7 - 1)! = 6! = 720(n−1)!=(7−1)!=6!=720
2. **List the permutations:**
   * ABCDEFG
   * BCDEFGH
   * CDEFGAB
   * DEFGABC
   * EFGABCD
   * FGABCDE
   * GABCDEF
   * (and so forth)

With 7 distinct objects, there are 720 unique circular permutations.

* **Circular Permutations Formula:** For n distinct objects, the number of unique circular permutations is (n−1)!.
* **Consider Rotations as Identical:** Each permutation can be rotated nnn ways, so you only count one arrangement per unique rotation.

Circular permutations are particularly useful in scenarios where the arrangement is cyclic, such as in circular tables, clock arrangements, or certain scheduling problems. If you have more questions or need further examples, feel free to ask!

**Basics of Combinations**

**Definition of Combination**

A **combination** is a selection of all or part of a set of objects without regard to the order of arrangement. The concept of combination is used when the order of selection does not matter.

**Formula for Combination**

The number of combinations of nnn distinct objects taken rrr at a time is given by:

A mathematical equation with black text

Description automatically generated

where:

* N! is the factorial of n
* r! is the factorial of r.

**Combinations of Distinct Objects**

When all the objects in the set are distinct, and the order does not matter, the combination formula is used.

**Example 1:** Consider the set {A, B, C, D}. The number of ways to select 2 letters out of 4 is:

A number with numbers and lines

Description automatically generated with medium confidence

The possible combinations are: AB, AC, AD, BC, BD, and CD.

**Example 2:** In a lottery where 6 numbers are chosen out of 49, the number of possible combinations is:

A number with numbers and a line

Description automatically generated with medium confidence

**Combinations with Repetition**

When repetition is allowed in combinations, the formula changes slightly. The number of combinations of n objects taken r at a time with repetition is given by:

A black text on a white background

Description automatically generated

**Example 3:** If you have three types of fruits (apple, banana, cherry), and you want to select 2 fruits with repetition, the number of possible combinations is:

A black text on a white background

Description automatically generated

The possible combinations are AA, AB, AC, BB, BC, and CC.

**Practical Applications of Permutations and Combinations**

**Permutations in Real-Life Scenarios**

**Example 4: Password Generation** Consider a scenario where you need to create a password using 4 letters (where repetition is not allowed) from the alphabet set of 26 distinct letters. The total number of permutations possible is:

A number and a number

Description automatically generated with medium confidence

This indicates there are 358,800 possible unique passwords.

**Example 5: Arranging People** Suppose 5 people need to be arranged in a row for a group photo. The total number of permutations possible is:



Thus, there are 120 different ways to arrange 5 people in a line.

**Combinations in Real-Life Scenarios**

**Example 6: Forming Committees** Imagine you need to form a committee of 3 members from a group of 10 people. The number of possible combinations is: A number and a line

Description automatically generated with medium confidence

This means there are 120 different ways to form a committee of 3 members from 10 people.

**Example 7: Selecting Ingredients** In a recipe, you can choose 4 ingredients out of 8 available. The number of combinations is:

A number and a number with numbers

Description automatically generated with medium confidence

So, there are 70 possible ways to choose 4 ingredients from 8.

**Advanced Permutation and Combination Concepts**

**Permutations of Non-Distinct Objects**

When some objects in a set are identical, the formula for permutations needs to be adjusted. The number of distinct permutations of nnn objects where there are n1, n2 …..nk​ objects of the same type is given by:

A black and white image of a line with letters

Description automatically generated with medium confidence

**Example 8: Arranging Letters** Consider the word "BALLOON". The total number of distinct permutations of these letters is:

A black text with a white background

Description automatically generated

So, there are 1,260 unique ways to arrange the letters in "BALLOON".

**Practical Examples of Permutations**

**Example 1: Arranging Books on a Shelf** Suppose you have 5 different books and you want to arrange them on a shelf. The total number of permutations is:



So, there are 120 different ways to arrange these 5 books.

**Example 2: Forming a 3-Digit Number** Consider forming a 3-digit number using the digits 1, 2, 3, 4, and 5, without repetition. The number of possible permutations is:

A number and a line

Description automatically generated with medium confidence

This means you can form 60 different 3-digit numbers.

**Example 3: Seating Arrangements** If you have 4 people and want to seat them in a row, the number of permutations is:

A black and white image of a number

Description automatically generated

Thus, there are 24 possible ways to arrange these 4 people in a row.

**Example 4: Selecting and Arranging Employees** Suppose you need to select 3 employees from a group of 6 and assign them different positions. The number of permutations is:

A number and a mathematical equation

Description automatically generated with medium confidence

There are 120 different ways to select and assign these 3 employees to the positions.

**Example 5: Creating a Password** If you need to create a 4-character password using the letters A, B, C, and D, with repetition not allowed, the total number of permutations is:



There are 24 possible unique passwords.

**Example 6: Lottery Number Arrangement** Imagine a lottery where you must choose 3 numbers from a set of 5 (1, 2, 3, 4, 5), and the order matters. The number of possible permutations is:

A black number and a equal sign

Description automatically generated with medium confidence

So, there are 60 different possible outcomes in this lottery.

**Example 7: Arranging Letters in a Word** Consider the word "TRAIN". The number of ways to arrange the letters is:

A number and a symbol

Description automatically generated with medium confidence

So, there are 120 different ways to arrange the letters in "TRAIN".

**Example 8: Organizing a Race** Suppose 7 runners are competing in a race, and you want to know how many different ways the first 3 places can be awarded. The number of permutations is:

A number and a number

Description automatically generated with medium confidence

This indicates there are 210 possible ways to award the top 3 positions.

**Example 9: Forming Committees** If you need to select a president, vice-president, and treasurer from a group of 8 members, the number of permutations is:

A number and a line

Description automatically generated with medium confidence

There are 336 different ways to assign these positions.

**Example 10: Deck of Cards** If you want to arrange 5 cards from a standard deck of 52, without repetition, the number of permutations is:

A number and a line

Description automatically generated with medium confidence

There are 311,875,200 different ways to arrange 5 cards.

**Example 11: Permutations with Repetition** If you want to create a 2-letter code from the letters A, B, C, D, allowing repetition, the number of permutations is:

A number and a symbol

Description automatically generated with medium confidence

There are 16 possible codes.

Top of Form

Bottom of Form

**Addition Theorem in Probability**

The **Addition Theorem** is used to find the probability of the occurrence of at least one of two events. There are two cases to consider:

1. **Mutually Exclusive Events (Disjoint Events)**: Events that cannot happen simultaneously. For example, rolling a die and getting either a 3 or a 5.
2. **Non-Mutually Exclusive Events**: Events that can occur together. For example, drawing a card from a deck that is both a heart and an ace.

**1. Addition Theorem for Mutually Exclusive Events**

For mutually exclusive events, the probability of either event A or event B occurring is the sum of their individual probabilities. Formula

A black text on a white background

Description automatically generated

A math problem with numbers and equations

Description automatically generated

**2. Addition Theorem for Non-Mutually Exclusive Events**

For non-mutually exclusive events, the probability of either event A or event B occurring is the sum of their individual probabilities, minus the probability of both events occurring together (to avoid double-counting).

A math problem with numbers and lines

Description automatically generated

**Multiplication Theorem in Probability**

The **Multiplication Theorem** is used to find the probability of the occurrence of two or more events together. Like the Addition Theorem, the Multiplication Theorem also has two cases:

1. **Independent Events**: Events where the occurrence of one does not affect the occurrence of the other. For example, tossing two coins.
2. **Dependent Events**: Events where the occurrence of one affects the occurrence of the other. For example, drawing cards from a deck without replacement.

**1. Multiplication Theorem for Independent Events**

For independent events, the probability of both events A and B occurring is the product of their individual probabilities.

A math problem with numbers and equations

Description automatically generated

**2. Multiplication Theorem for Dependent Events**

For dependent events, the probability of both events A and B occurring is the product of the probability of A and the probability of B given that A has occurred.

A white paper with black text

Description automatically generated

**Here are examples for both the Addition and Multiplication Theorems in probability:**

**Example 1: Addition Theorem**

Scenario: You are attending a party where a game involves drawing a card from a standard deck of 52 cards. You win a prize if you draw either a spade or a face card (jack, queen, or king).

Problem: What is the probability of winning the prize?

**\*\*Solution:\*\***

- There are 13 spades in a deck: ( P({Spade}) = 13/52

- There are 12 face cards (3 per suit): (P({Face Card}) = 12/52

- There are 3 face cards that are also spades: ( P({Spade and Face Card}) = 3/52

Using the Addition Theorem for Non-Mutually Exclusive Events:

A math equations with numbers

Description automatically generated

**Example 2: Multiplication Theorem**

Scenario: You are playing a board game where you need to roll two six-sided dice. To win the game, you need to roll a 4 on the first die and a 5 on the second die.

Problem: What is the probability of rolling a 4 on the first die and a 5 on the second die?

**\*\*Solution:\*\***

- Probability of rolling a 4 on the first die: (P({4 on first die}) = 1/6

- Probability of rolling a 5 on the second die: (P({5 on second die}) = 1/6

Since these are independent events:

A white background with black text

Description automatically generated

P(4 on first die and 5 on second die) =P(4 on first die)×P(5 on second die)=

A black and white math equation

Description automatically generated

So, the probability of rolling a 4 on the first die and a 5 on the second die is 1/36.

**Bayes’s Theorem: Detailed Concept**

**Definition**

**Bayes’s Theorem** is a fundamental concept in probability theory that describes how to update the probability of a hypothesis, HHH, based on new evidence, EEE. It is named after the Reverend Thomas Bayes, who first provided an equation that allows new evidence to update beliefs about the likelihood of a given event.

Bayes’s Theorem links the conditional probability of the hypothesis given the evidence, P(H∣E)P(H|E)P(H∣E), with the conditional probability of the evidence given the hypothesis, P(E∣H)P(E|H)P(E∣H), along with the prior probability of the hypothesis, P(H)P(H)P(H), and the marginal likelihood of the evidence, P(E)P(E)P(E).

**Concept**

Bayes’s Theorem is built upon the concept of **conditional probability**, which is the probability of an event occurring given that another event has already occurred. In real life, we often encounter situations where we have some prior knowledge about an event, and as we gather more information, we refine our predictions or beliefs about that event.

Bayes’s Theorem is particularly useful in situations where we need to make decisions based on incomplete or evolving information. It allows us to revise our predictions or hypotheses by incorporating new data. This concept is widely used in various fields like medical diagnosis, machine learning, finance, and more.

**The Formula**

The mathematical formula for Bayes’s Theorem is:

A mathematical equation with black text

Description automatically generated

Where:

* P(H∣E) is the **posterior probability**, the probability of the hypothesis HHH given the evidence EEE.
* P(E∣H) is the **likelihood**, the probability of the evidence EEE given that the hypothesis HHH is true.
* P(H) is the **prior probability**, the initial probability of the hypothesis HHH before considering the evidence EEE.
* P(E) is the **marginal likelihood** or **evidence**, the total probability of the evidence EEE under all possible hypotheses.

**Example**

Let’s consider a classic example related to medical diagnosis.

**Scenario:** A patient is being tested for a rare disease. The test for the disease is 99% accurate, meaning it correctly identifies the disease 99% of the time if the patient has it, and it correctly identifies 99% of healthy patients as not having the disease. However, the disease is quite rare, affecting only 1 in 10,000 people.

**Problem:** If the test result comes back positive, what is the probability that the patient actually has the disease?

**Solution:**

* Let H be the event that the patient has the disease.
* Let E be the event that the test result is positive.

We know the following:

A test results with black text

Description automatically generated

First, we need to calculate the marginal likelihood P(E), which is the total probability of getting a positive test result under both scenarios (having the disease or not having the disease):

P(E)=P(E∣H)⋅P(H)+P(E∣¬H)⋅P(¬H)

Substituting the values:

P(E)=(0.99×0.0001)+(0.01×0.9999)

P(E)=0.000099+0.009999=0.010098

Now, using Bayes’s Theorem to calculate the posterior probability P(H∣E):

A mathematical equation with numbers and symbols

Description automatically generated

So, even after a positive test result, the probability that the patient actually has the disease is only about 0.98%, which is surprisingly low. This result is due to the rarity of the disease combined with the fact that the test, while accurate, still has a small false positive rate.

**Intuition and Insights**

* Prior and Posterior Probability: The prior probability reflects what we know before considering new evidence, while the posterior probability updates this belief in light of the new evidence.
* Impact of Rare Events: When dealing with rare events, even highly accurate tests can lead to counterintuitive results. This is known as the base rate fallacy, where the base rate (prior probability) of the event significantly influences the outcome.
* Relevance in Decision-Making: Bayes’s Theorem is crucial in decision-making processes that involve uncertainty. By systematically updating probabilities, it helps make more informed decisions.

**Applications of Bayes’s Theorem**

Bayes’s Theorem is applied in various fields:

* Medical Diagnosis: Estimating the likelihood of a disease given a test result, as illustrated in the example above.
* Machine Learning: Algorithms like Naive Bayes classifiers rely on Bayes’s Theorem for classifying data.
* Finance: Estimating the likelihood of market movements based on new financial data or news.
* Law: Assessing the likelihood of a suspect’s guilt given new evidence in a case.

Bayes’s Theorem provides a robust framework for updating beliefs and making decisions under uncertainty, making it a powerful tool in both theoretical and applied probability.

**Permutation vs Combination**

The key differences between [permutation and combination](https://www.geeksforgeeks.org/permutations-and-combinations/), some of those differences are listed as follows:

| **Aspect** | **Permutations** | **Combinations** |
| --- | --- | --- |
| **Definition** | Arrangements of elements in a specific order. | Selections of elements without considering the order. |
| **Formula** | nPr = n!​/(n−r)! | nPr = n!​/[(n−r)! × r!] |
| **Notation** | nPr  OR P(n, r) | nCr  OR C(n, r) |
| **Order Matters** | Yes, order matters. | No, order doesn’t matter. |
| **Example** | Arranging books on a shelf. | Selecting members for a committee. |
| **Sample Problems** | How many ways to arrange 3 books out of 5? | How many ways to choose 2 fruits from a basket of 7? |
| **Application** | Permutations are used when order matters, such as arranging items in a sequence or forming a code. | Combinations are used when order doesn’t matter, like selecting a group of people or choosing items without caring about their order. |

**Unit 2:**

**Topics**

Random Variable Concept

Discrete and Continuous Random Variable

Probability density function

Mathematical Expectation and their Theorem

**Random Variable Concept**

**Definition**

A \*\*Random Variable\*\* is a fundamental concept in probability and statistics, representing a variable whose possible values are outcomes of a random phenomenon. It is a function that assigns a numerical value to each outcome in a sample space, which is the set of all possible outcomes of a random experiment.

In simpler terms, a random variable is a way to quantify the outcomes of a random process. For example, when you roll a die, the outcome can be any number between 1 and 6. If we define a random variable X to represent the outcome, then X can take any of these six values.

**Types of Random Variables**

There are two main types of random variables: \*\*Discrete\*\* and \*\*Continuous\*\*.

**a. Discrete Random Variable**

A **\*\*Discrete Random Variable\*\*** takes on a countable number of distinct values. These values are often integers and can be listed out. Common examples include the number of heads when flipping a coin multiple times, the number of students in a classroom, or the number rolled on a die.

\*\*Example:\*\*

Let X represent the number of heads in three flips of a fair coin. The possible values of X are 0, 1, 2, or 3, because you can get anywhere from 0 to 3 heads in three flips.

**b. Continuous Random Variable**

A **\*\*Continuous Random Variable\*\*** can take on an infinite number of possible values within a given range. These values are often real numbers, and they are typically measured rather than counted. Examples include the height of students in a class, the time it takes to run a race, or the temperature at a particular location.

\*\*Example:\*\*

Let Y represent the time it takes for a runner to complete a marathon. Y can take any value from, say, 2 hours to 6 hours, including any fractional value within this range (e.g., 3.5 hours, 4.1 hours).

**Probability Distribution**

A **\*\*Probability Distribution\*\*** describes how the probabilities are distributed over the values of a random variable. The probability distribution depends on whether the random variable is discrete or continuous.

**a. Probability Mass Function (PMF) for Discrete Random Variables**

For a discrete random variable, the probability distribution is described by a \*\*Probability Mass Function (PMF)\*\*. The PMF gives the probability that a random variable is exactly equal to some value.

**\*\*Example:\*\***

For a fair six-sided die, let X be the outcome when the die is rolled. The PMF is:

A black text on a white background

Description automatically generated

This means each outcome (1 through 6) has an equal probability of 1/6.

**b. Probability Density Function (PDF) for Continuous Random Variables**

For a continuous random variable, the probability distribution is described by a **\*\*Probability Density Function (PDF)\*\*.** Unlike the PMF, the PDF does not give probabilities directly but rather describes the density of the probability at each point. The probability of the variable falling within a specific range is given by the area under the curve of the PDF over that range.

**\*\*Example:\*\***

Let Y represent the height of adult men in a population, and assume Y ) follows a normal distribution with a mean of 70 inches and a standard deviation of 3 inches. The PDF of Y describes how heights are distributed around the mean. The probability that a randomly chosen man is between 68 and 72 inches tall is given by the area under the PDF curve between these two values.

**Expectation (Mean) and Variance**

**a. Expectation (Mean)**

The **\*\*Expectation\*\* or \*\*Mean\*\*** of a random variable is the long-run average value of repetitions of the experiment it represents. It gives a measure of the central tendency of the distribution.

- For a **\*\*Discrete Random Variable\*\*** X with possible values x1, x2….., xn ) and corresponding probabilities P(X = x1), P(X = x2), ….., P(X = xn) , the expectation E(X) is:

A black and white math equation

Description automatically generated

- For a **\*\*Continuous Random Variable\*\*** Y with a probability density function f(y) , the expectation E(Y) is:

A black and white math symbols

Description automatically generated with medium confidence

**\*\*Example:\*\***

If you roll a fair die, the expected value (mean) of the outcome is:

A number and plus symbol

Description automatically generated with medium confidence

**b. Variance and Standard Deviation**

The **\*\*Variance\*\*** of a random variable measures the spread or dispersion of the values around the mean. It is the expected value of the squared deviation of the random variable from its mean.

- For a \*\*Discrete Random Variable\*\* X with mean μ :

A mathematical equation with numbers and symbols

Description automatically generated

- For a **\*\*Continuous Random Variable\*\*** Y with mean μ :

A black and white math equation

Description automatically generated with medium confidence

The \*\*Standard Deviation\*\* is the square root of the variance, providing a measure of spread in the same units as the original variable.

\*\*Example:\*\*

If you roll a fair die, the variance of the outcome can be calculated as:

A number with numbers on it

Description automatically generated with medium confidence

**Applications of Random Variables**

Random variables are extensively used in statistical analysis and modeling across various fields:

- Finance: Modeling stock prices, where the price at any given time can be considered a random variable.

- Insurance: Estimating risks, where claims and losses are modeled as random variables.

- Engineering: Reliability analysis, where the failure time of a component is treated as a random variable.

- Science: Experiment outcomes, such as particle counts in physics experiments, are modeled using random variables.

**Real-Life Example**

**Scenario:** Imagine a company produces light bulbs, and historically, 5% of the bulbs are defective. Let X be the random variable representing the number of defective bulbs in a sample of 100 bulbs.

**\*\*Solution:\*\***

- X is a discrete random variable because it counts the number of defective bulbs.

- The probability of a bulb being defective is P(\text{defective}) = 0.05 .

- The expected number of defective bulbs in a sample of 100 is:

E(X) = n x p = 100 \* 0.05 = 5

- The variance can be calculated using the formula for the variance of a binomial distribution

Var(X) = n x p x (1 - p) :

Var(X) = 100 x 0.05 x 0.95 = 4.75

- The standard deviation is sqrt(4.75) = approx 2.18 .

This analysis helps the company understand the expected number of defective bulbs and the variability around this expectation, which is crucial for quality control and decision-making.

Random variables are the building blocks of statistical analysis, allowing us to model and understand the uncertainty inherent in various processes. By defining, analyzing, and interpreting random variables, we can make informed decisions in fields as diverse as finance, engineering, science, and beyond.

**Discrete and Continuous Random Variable**

**Introduction to Random Variables**

A Random Variable is a variable that takes on different values based on the outcomes of a random experiment. It quantifies the outcomes of random phenomena and is a key concept in probability and statistics. Random variables can be categorized into two main types: Discrete and Continuous.

**Discrete Random Variables**

**Definition**

A Discrete Random Variable is a random variable that can take on a countable number of distinct values. These values are typically integers and can be listed individually. The term "discrete" indicates that there are gaps between the possible values of the variable, meaning the variable can only assume specific points on the number line.

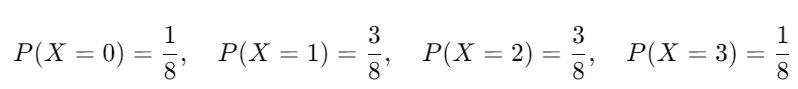
**Characteristics**

* Countable Outcomes: The values a discrete random variable can take are finite or countably infinite. For example, the number of students in a classroom or the number of heads when flipping a coin multiple times.
* Probability Mass Function (PMF): The probability distribution of a discrete random variable is described by a Probability Mass Function (PMF). The PMF gives the probability that the random variable is exactly equal to each possible value. The sum of all probabilities in the PMF equals 1.

**Example 1: Number of Heads in Coin Flips**

Let’s say you flip a fair coin three times. Define the discrete random variable XXX as the number of heads obtained.

* Possible values of X: 0, 1, 2, 3.
* The probabilities can be calculated using the binomial distribution:



The PMF for X is:

A black and white math equation

Description automatically generated

**Example 2: Rolling a Die**

Consider rolling a fair six-sided die. Define the discrete random variable YYY as the outcome of the roll.

* Possible values of Y: 1, 2, 3, 4, 5, 6.
* Each outcome has an equal probability of 1/6​.

The PMF for YYY is:

A black and white math equation

Description automatically generated

**Continuous Random Variables**

**Definition**

A Continuous Random Variable is a random variable that can take on an infinite number of possible values within a given range. These values are uncountable and typically include real numbers, meaning the variable can assume any value within a certain interval.

Characteristics

* Uncountable Outcomes: The values of a continuous random variable are uncountable and can include any real number within a certain range. Examples include the height of people, the time taken to complete a task, or the temperature in a room.
* Probability Density Function (PDF): The probability distribution of a continuous random variable is described by a Probability Density Function (PDF). Unlike the PMF, the PDF does not give the probability that the variable takes a specific value; instead, it gives the density of probability at each point. The probability that the variable falls within a specific range is found by calculating the area under the curve of the PDF over that range.

**Example 1: Height of Students**

Let’s say the height of students in a class is normally distributed with a mean of 170 cm and a standard deviation of 10 cm. Define the continuous random variable ZZZ as the height of a randomly selected student.

* The PDF of Z is described by the normal distribution:

A math equation with a number of equations

Description automatically generated with medium confidence

where μ=170 and σ=10

* The probability that a student's height is between 160 cm and 180 cm is given by the area under the PDF curve from 160 to 180.

**Example 2: Time to Complete a Task**

Consider the time TTT it takes to complete a task, which could be any value between, say, 0 and 10 hours. If TTT is uniformly distributed, the PDF is constant over the interval.

* The PDF for T is:

A math equations on a white background

Description automatically generated

* The probability that the task takes between 4 and 6 hours is the area under the PDF from 4 to 6, calculated as:

A black and white math equation

Description automatically generated with medium confidence

**Key Differences Between Discrete and Continuous Random Variables**

| **Feature** | **Discrete Random Variable** | **Continuous Random Variable** |
| --- | --- | --- |
| Possible Values | Countable, distinct values (e.g., 0, 1, 2, 3) | Uncountable, infinite values (e.g., 1.5, 2.75, 3.14) |
| Probability Distribution | Probability Mass Function (PMF) | Probability Density Function (PDF) |
| Probability Calculation | Direct probabilities for exact values | Probability for intervals (area under the curve) |
| Example | Number of students in a class, number of heads | Height of students, time to complete a task |
| Sum of Probabilities | Sum of all probabilities equals 1 | Total area under the PDF curve equals 1 |

**Applications in Statistical Analysis**

**a. Discrete Random Variables**

Discrete random variables are often used in situations where the data is inherently countable. Some applications include:

* **Quality Control:** Number of defective items in a batch.
* **Epidemiology:** Number of new cases of a disease.
* **Insurance:** Number of claims filed in a year.

**b. Continuous Random Variables**

Continuous random variables are used when the data can take any value within a range. Some applications include:

* **Finance:** Modeling stock prices, where prices can take any real number within a range.
* **Environmental Science:** Measuring pollutants in the air, which can take any value within a given concentration range.
* **Engineering:** Analyzing the time until failure of a machine component.

**Probability density function**

**Probability Density Function**

[**Probability**](https://www.geeksforgeeks.org/card-probability/)**Density Function** is the function of probability defined for various distributions of variables and is the less common topic in the study of probability throughout the academic journey of students. However, this function is very useful in many areas of real life such as predicting rainfall, financial modelling such as the stock market, income disparity in social sciences, etc.

This article explores the topic of the Probability Density Function in detail including its definition, condition for existence of this function, as well as various examples.

**What is Probability Density Function(PDF)?**

Probability Density Function is used for calculating the probabilities for continuous random variables. When the cumulative distribution function (CDF) is differentiated we get the probability density function (PDF). Both functions are used to represent the probability distribution of a continuous random variable.

The probability density function is defined over a specific range. By differentiating CDF we get PDF and by integrating the probability density function we can get the cumulative density function.

**Probability Density Function Definition**

*Probability density function is the function that represents the density of probability for a continuous random variable over the specified ranges.*

Probability Density Function is abbreviated as PDF and for a continuous [random variable](https://www.geeksforgeeks.org/random-variable/) X, Probability Density Function is denoted by f(x).

PDF of the random variable is obtained by differentiating CDF (Cumulative Distribution Function) of X. The probability density function should be a positive for all possible values of the variable. The total area between the density curve and the x-axis should be equal to 1.

**Necessary Conditions for PDF**

Let X be the continuous random variable with probability density function f(x). For a function to be valid probability function should satisfy below conditions.

* *f(x) ≥ 0, ∀ x ∈ R*
* *f(x) should be piecewise continuous.*

A black math equation

Description automatically generated with medium confidence

So, the PDF should be the non-negative and piecewise continuous function whose total value evaluates to 1.

**Example of a Probability Density Function**

Let X be a continuous random variable and the probability density function pdf is given by f(x) = x – 1 , 0 < x ≤ 5. We have to find P (1 < x ≤ 2).

To find the probability P (1 < x ≤ 2) we integrate the pdf f(x) = x – 1 with the limits 1 and 2. This results in the probability P (1 < x ≤ 2) = 0.5

**Probability Density Function Formula**

Let Y be a continuous random variable and F(y) be the cumulative distribution function (CDF) of Y. Then, the probability density function (PDF) f(y) of Y is obtained by differentiating the CDF of Y.

***A mathematical equation with a few square objects

Description automatically generated with medium confidence***

If we want to calculate the probability for X lying between the interval a and b, then we can use the following formula:

***A black text on a white background

Description automatically generated***

**Key Points about PDF Formula**

*A math equations on a white background

Description automatically generated* **What Does a Probability Density Function (PDF) Tell Us?**

A Probability Density Function (PDF) is a function that describes the likelihood of a continuous random variable taking on a particular value. Unlike discrete random variables, where probabilities are assigned to specific outcomes, continuous random variables can take on any value within a range. Probability Density Function (PDF) tells us

* Relative Likelihood
* Distribution Shape
* Expected Value and Variance, etc.

**How to Find Probability from Probability Density Function**

To find the probability from the probability density function we have to follow some steps.

***Step 1:****First check the PDF is valid or not using the necessary conditions.*

***Step 2:****If the PDF is valid, use the formula and write the required probability and limits.*

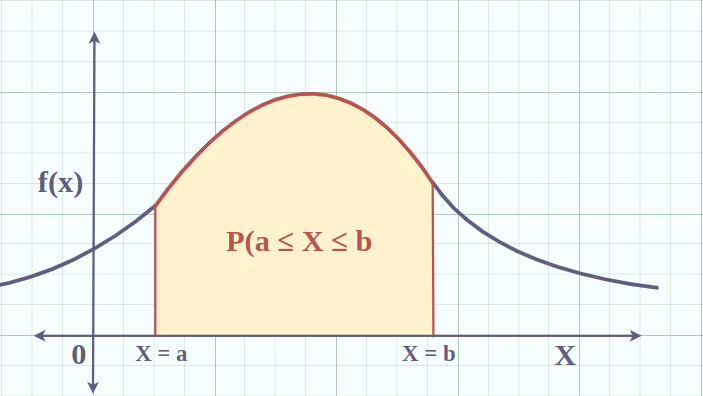
***Step 3:****Divide the integration according to the given PDF.*

***Step 4:****Solve all integrations.*

***Step 5:****The resultant value gives the required probability.*

**Graph for Probability Density Function**

If X is continuous random variable and f(x) be the probability density function. The probability for the random variable is given by area under the pdf curve. The graph of PDF looks like bell curve, with the probability of X given by area below the curve. The following graph gives the probability for X lying between interval a and b.



**Probability Density Function Properties**

Let f(x) be the probability density function for continuous random variable x. Following are some probability density function properties:

* Probability density function is always positive for all the values of x.

**f(x) ≥ 0, ∀ x ∈ R**

* Total area under probability density curve is equal to 1.

A black text with a white background

Description automatically generated

* For continuous random variable X, while calculating the random variable probabilities end values of the interval can be ignored i.e., for X lying between interval a and b

**P (a ≤ X ≤ b) = P (a ≤ X < b) = P (a < X ≤ b) = P (a < X < b)**

* Probability density function of a continuous random variable over a single value is zero.

**A black text with a white background

Description automatically generated with medium confidence**

* Probability density function defines itself over the domain of the variable and over the range of the continuous values of the variable.

**Mean of Probability Density Function**

[Mean](https://www.geeksforgeeks.org/what-is-mean/) of the probability density function refers to the average value of the random variable. The mean is also called as expected value or expectation. It is denoted by μ or E[X] where, X is random variable.

Mean of the probability density function f(x) for the continuous random variable X is given by:

*A mathematical equation with symbols

Description automatically generated with medium confidence*

**Median of Probability Density Function**

[Median](https://www.geeksforgeeks.org/median/) is the value which divides the probability density function graph into two equal halves. If x = M is the median then, area under curve from -∞ to M and area under curve from M to ∞ are equal which gives the median value = 1/2.

Median of the probability density function f(x) is given by:

A mathematical equation with a number of symbols

Description automatically generated with medium confidence

**Variance Probability Density Function**

[Variance](https://www.geeksforgeeks.org/variance/) of probability density function refers to the squared deviation from the mean of a random variable. It is denoted by Var(X) where, X is random variable.

Variance of the probability density function f(x) for continuous random variable X is given by:

***A black text with a white background

Description automatically generated with medium confidence***

**Standard Deviation of Probability Density Function**

[Standard Deviation](https://www.geeksforgeeks.org/how-to-find-standard-deviation-in-r/) is the square root of the variance. It is denoted by σ and is given by:

***A black square and square symbol

Description automatically generated***

**Probability Density Function Vs Cumulative Distribution Function**

The key differences between Probability Density Function (PDF) and Cumulative Distribution Function (CDF) are listed in the following table:

| **Aspect** | **Probability Density Function (PDF)** | **Cumulative Distribution Function (CDF)** |
| --- | --- | --- |
| **Definition** | The PDF gives the probability that a random variable takes on a specific value within a certain range. | The CDF gives the probability that a random variable is less than or equal to a specific value. |
| **Range of Values** | Defined for continuous random variables. | Defined for both continuous and discrete random variables. |
| **Mathematical Expression** | f(x), where f(x)≥0 and ∫−∞∞​f(x)dx=1 | F(x), where 0≤F(x)≤1 for all x, and F(−∞)=0 and F(∞)=1 |
| **Interpretation** | Represents the likelihood of the random variable taking on a specific value. | Represents the probability that the random variable is less than or equal to a specific value. |
| **Area Under the Curve** | The area under the PDF curve over a certain interval gives the probability that the random variable falls within that interval. | The value of the CDF at a specific point gives the probability that the random variable is less than or equal to that point. |
| **Relationship with CDF** | The PDF can be obtained by differentiating the CDF with respect to the random variable. | The CDF can be obtained by integrating the PDF with respect to the random variable. |
| **Probability Calculation** | The probability of a random variable falling within a specific interval (a,b) is given by ∫ab​f(x)dx. | The probability of a random variable being less than or equal to a specific value x is given by F(x). |
| **Properties** | The PDF is always non-negative: f(x)≥0 for all x. The total area under the PDF curve is equal to 1. | The CDF is a monotonically increasing function: F(x1​) ≤ F(x2​) if x1​ ≤ x2​. 0≤F(x)≤1 for all x. |
| **Examples** | Normal Distribution PDF:  1σ2πe−(x−μ)22σ2*σ*2*π*​1​*e*−2*σ*2(*x*−*μ*)2​  Exponential distribution PDF: λe−λx | Normal Distribution CDF:  12(1+erf(x−μσ2))21​(1+erf(*σ*2​*x*−*μ*​))  Exponential distribution CDF: 1−e−λx |

**Types of Probability Density Function**

There are different types of probability density functions given below:

* Uniform Distribution
* Binomial Distribution
* Normal Distribution
* Chi-Square Distribution

**Difference Between PDF and Joint PDF**

The PDF is the function defined for single variable whereas joint PDF is the function defined for two or more than two variables, and other key differences between these both concepts are listed in the following table:

| **PDF (Probability Density Function)** | **Joint PDF** |
| --- | --- |
| Probability Density Function is the probability function defined for single variable. | Joint Probability Density Function is the probability function defined for more than one variable. |
| It is denoted as f(x). | It is denoted as f (x, y, …). |
| Probability Density Function is obtained by differentiating the CDF. | Joint Probability Density Function is obtained by differentiating the joint CDF |
| It can be calculated by single integral. | It can be calculated using multiple integrals as there are multiple variables. |

**Applications of Probability Density Function**

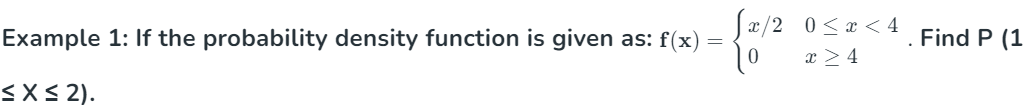
Some of the applications of Probability Density function are:

* Probability density functions are used in statistics for calculating probabilities for random variables.
* It is used in modelling various scientific data.

**Read More,**

* [***Cumulative Frequency Distribution***](https://www.geeksforgeeks.org/cumulative-frequency/)
* [***Probability Distribution Function***](https://www.geeksforgeeks.org/probability-distribution-function/)

**Examples on Probability Density Function**

** Solution:**

*A math equations on a white background

Description automatically generated*

A screenshot of a math problem

Description automatically generated

**A screenshot of a math problem

Description automatically generated**

**Mathematical Expectation and their Theorem**

Because random variables are *random*, knowing the outcome on any one realisation of the random process is not possible. Instead, we can talk about what we might *expect* to happen, or what might happen *on average*.

This is the idea of *mathematical expectation*. In more usual terms, the mathematical expression of the probability distribution of a random variable is the *mean* of the random variable. Mathematical expectation goes far beyond just computing means, but we begin here as the idea of a *mean* is easily understood.

The definition looks different in detail for discrete and continuous random variables, but the intention is the same.

**Definition 3.1 (Expectation)**The *expectation* or *expected value* (or *mean*) of a random variable X is defined as

A screenshot of a computer

Description automatically generated

Effectively E(X) is a weighted average of the points in RX, the weights being the probabilities in the discrete case and probability densities in the continuous case.

**Example - (Expectation for discrete variables) Consider the discrete random variable U with probability function**

A math equations on a white background

Description automatically generated

**Example - (Expectation for continuous variables) Consider a continuous random variable X with pdf**

**A white paper with black text and numbers

Description automatically generated**

**Example - (Expectation for a coin toss)**Consider tossing a coin *once* and counting the number of tails. Let this random variable be T. The probability function is

A math equations and formulas

Description automatically generated

Of course, 0.50.5 tails can never actually be observed in practice on one toss. But it would be silly to round up (or down) and say that the expected number of tails on one toss of a coin is one (or zero). The expected value of 0.50.5 simply means that over a large number of repeats of this random process, we expect a tail to occur in half of those repeats.

**Example (Mean not defined)**Consider the distribution of Z, with the probability density function

A math equations and formulas

Description automatically generated with medium confidence

A graph of a function

Description automatically generated

**Expectation of a function of a random variable**

While the mean can be expressed in terms of mathematical expectation, mathematical expectation is a more general concept.

Let X be a discrete random variable with a probability function pX(x), or a continuous random variable with pdf fX(x). Also assume g(X) is a real-valued function of X. We can then define the expected value of g(X).

**Definition (Expectation for function of a random variable)**The *expected value* of some function g(⋅) of a random variable X is:

A white background with black text

Description automatically generated