La-Net

STA211

PROBABILITY II

Series 1: Advanced Permutations, Combinations and Probability Laws

Series outline

- **Part 1.1: Permutations**
  - 1.1.1 Definition and properties of permutations
  - 1.1.2 Examples and applications of permutations
  - 1.1.3 Advanced counting with restrictions
- **Part 1.2: Combinations**
  - 1.2.1 Definition and properties of combinations
  - 1.2.2 Examples and applications of combinations
  - 1.2.3 Distinction between permutations and combinations
- **Part 1.3: Probability Laws**
  - 1.3.1 Definition of probability laws
  - 1.3.2 Addition and multiplication laws of probability
  - 1.3.3 Applications of probability laws in problem solving
- **Part 1.4: Applications of Permutations, Combinations and Probability Laws**
  - 1.4.1 Real-life applications in probability modelling
  - 1.4.2 Problem-solving strategies using counting and probability laws
  - 1.4.3 Combined applications in discrete probability

Introduction

Probability is a powerful tool that helps us quantify uncertainty and make informed decisions in everyday life and scientific investigations. From arranging objects in specific orders to selecting subsets from larger collections, the concepts of **permutations** and **combinations** provide the foundation for counting methods in probability. These counting techniques, when combined with the fundamental **laws**

of probability, enable us to solve a wide range of problems in mathematics, statistics, and applied fields. This Series builds on your earlier knowledge of basic probability and introduces more advanced methods that will be essential for tackling complex problems later in the course.

The aim of this Series is to equip you with the ability to define, explain, and apply advanced permutations, combinations, and probability laws in solving practical and theoretical problems.

We shall commence this Series by exploring permutations, beginning with their definition, properties, and applications, including cases with restrictions. Next, we will move on to combinations, where we will examine their definition, properties, and applications, as well as the distinction between permutations and combinations. Following that, we will study probability laws, focusing on the addition and multiplication laws and their applications in problem solving. Finally, we will conclude this Series with applications of permutations, combinations, and probability laws, showing how these concepts are used together in real-life probability modelling and discrete probability problems.

Series Learning Outcomes

Upon completing this Series, you should be able to:

- **SLO 1.1** define permutations and explain their properties,
- **SLO 1.2** apply permutations to solve counting problems with and without restrictions,
- **SLO 1.3** define combinations and explain their properties,
- **SLO 1.4** distinguish between permutations and combinations with examples,
- **SLO 1.5** define probability laws and explain the addition and multiplication laws of probability,
- **SLO 1.6** apply probability laws to solve practical problems, and
- **SLO 1.7 evaluate combined applications of permutations, combinations, and probability laws in real-life probability modelling.**

1.1 Permutations

Permutations are a fundamental concept in probability and combinatorics. A **permutation** is an arrangement of objects in a specific order. The order of arrangement is important, which distinguishes permutations from combinations.

For example, arranging the letters A, B, and C as ABC, BAC, or CAB are all different permutations.

Fill in the blank: A permutation is an arrangement of objects in a specific __________.

1.1.1 Definition and properties of permutations

A permutation of ndistinct objects taken rat a time is defined as the number of ways of arranging robjects out of n. It is denoted by P(n,r)and given by the formula:

$$P(n,r)=n!(n-r)!$$

Here, **factorial** (denoted n!) means the product of all positive integers up to n. For example, $5!=5\times4\times3\times2\times1=120$.

When you arrange three distinct items, like our three labeled and colored balls, the different possible orders are called **permutations**.

Since there are three choices for the first ball, two for the second, and only one choice left for the third, the total number of unique permutations is calculated as:

$$3 \times 2 \times 1 = 3! = 6$$

Here is a tree diagram illustrating all 6 possible arrangements (permutations) of the three balls:

Here is a tree diagram illustrating all 6 possible arrangements (permutations) of the three balls:

Listing All Permutations

This systematic approach ensures we don't miss any of the six unique orders:

1. **A, B, C**
2. **A, C, B**
3. **B, A, C**
4. **B, C, A**
5. **C, A, B**
6. **C, B, A**

Each row in the list represents a complete, unique permutation of the set {A, B, C}.

Illustration of permutations of three objects.

1.1.2 Examples and applications of permutations

Let us consider a simple example. Suppose we want to know how many ways we can arrange the letters A, B, and C. Since there are three letters, the number of permutations is $3! = 6$. The arrangements are ABC, ACB, BAC, BCA, CAB, and CBA.

Another example is arranging five students in a line. The number of possible arrangements is $5! = 120$.

Permutations are widely applied in real life, such as in arranging seats, scheduling tasks, and designing codes.

Fill in the blank: The number of ways to arrange three letters A, B, and C is equal to __________.

Examples of permutations

Scenario	**Number of permutations**
Arranging 3 letters	$3! = 6$
Arranging 5 students	$5! = 120$

Selecting 2 out of 4 objects	P(4,2)=12

1.1.3 Advanced counting with restrictions

Sometimes permutations involve restrictions. For example, arranging digits to form numbers where repetition is not allowed, or arranging people in seats where certain positions are fixed.

Consider arranging the digits 1, 2, 3, and 4 to form a three-digit number without repetition. The first digit can be chosen in 4 ways, the second in 3 ways, and the third in 2 ways. Thus, the total number of permutations is $4 \times 3 \times 2 = 24$.

Fill in the blank: If four digits are used to form a three-digit number without repetition, the total number of permutations is __________.

Key points on restricted permutations

- Restrictions may include fixed positions or no repetition.
- Multiply choices step by step to find total arrangements.
- Useful in forming codes, passwords, and seating plans.

Restriction Type	Logic / Method	Formula / Approach
Objects Must Be Together	Treat the group of objects as a **single "super-object."**	$(n-k+1)! \times k!$ (where k is the group size)
Objects Must NOT Be Together	Use the **Complement Method**: Total permutations minus "together" permutations.	$n! - [(n-k+1)! \times k!]$
Fixed Positions	If an object *must* be at the start or end, its position is "locked."	Permute the remaining $(n-1)$ objects: $(n-1)!$
Gap Method	To ensure no two items from a specific group are adjacent.	Arrange the "non-restricted" items first, then place restricted items in the "gaps."

Pilot questions 1.1

1. The formula for permutations of nobjects taken rat a time is: A. n!r!B. n!(n-r)!C. nrD. None of the above

2. The number of ways to arrange three letters A, B, and C is: A. 3 B. 6 C. 9 D. 12
3. The factorial of 5, denoted 5!, is equal to: A. 24 B. 60 C. 120 D. 720
4. If four digits are used to form a three-digit number without repetition, the total number of permutations is: A. 12 B. 24 C. 36 D. 48
5. Which of the following scenarios involves restricted permutations? A. Arranging 5 students in a line B. Arranging digits to form a number without repetition C. Arranging letters in any order D. None of the above

1.2 Combinations

Combinations are another key concept in probability and combinatorics. A **combination** is a selection of objects from a set where the order of selection does not matter. This is the main difference between combinations and permutations. For example, choosing the letters A, B, and C together is the same as choosing B, A, and C, since the arrangement is irrelevant.

Fill in the blank: A combination is a selection of objects where __________ does not matter.

1.2.1 Definition and properties of combinations

A combination of n distinct objects taken rat a time is denoted by C(n,r)or nr. The formula is:

$$C(n,r)=n!r!(n-r)!$$

This formula shows that combinations are derived from permutations but adjusted to remove the effect of order.

For our set of three distinct colored balls (A, B, and C), if we are choosing all three, there is only **one** possible combination.

$$\binom{3}{3} = \frac{3!}{3!(3-3)!} = \frac{6}{6 \times 1} = 1$$

Here is a diagram illustrating that all six permutations result in the same single combination:

The Unique Combination

Whether you pick them as A-B-C, C-B-A, or B-A-C, you still end up with the same set of three balls. The order of selection is irrelevant.

1. **{A, B, C}**

This single set is the only unique combination when selecting three objects from a set of three.

Illustration of combinations of three objects.

1.2.2 Examples and applications of combinations

Suppose we want to know how many ways we can select two letters from A, B, and C. Using the formula:

$$C(3,2)=3!2!(3-2)!=62=3$$

The possible selections are AB, AC, and BC.

Another example is choosing a committee of 3 people from a group of 10. The number of possible committees is:

$$C(10,3)=10!3!\cdot 7!=120$$

Combinations are widely applied in probability problems, such as lottery draws, team selections, and sampling.

Fill in the blank: The number of ways to select two letters from A, B, and C is equal to __________.

Examples of combinations

Scenario	Number of combinations
Selecting 2 letters from 3	$C(3,2)=3$

Choosing 3 people from 10	C(10,3)=120
Selecting 5 cards from 52	C(52,5)

1.2.3 Distinction between permutations and combinations

The key distinction is that **permutations** consider order, while **combinations** do not. For example, arranging A, B, and C in different orders gives six permutations, but selecting any two letters gives only three combinations.

This distinction is crucial in probability problems, as using the wrong method can lead to incorrect results.

Fill in the blank: The main difference between permutations and combinations is that permutations consider __________, while combinations do not.

Key differences between permutations and combinations

- Permutations: order matters.
- Combinations: order does not matter.
- Permutations are used in arrangements, combinations in selections.

Pilot questions 1.2

1. The formula for combinations of nobjects taken rat a time is: A. n!(n-r)!B. n!r!(n-r)!C. nrD. None of the above
2. The number of ways to select two letters from A, B, and C is: A. 2 B. 3 C. 6 D. 9
3. The number of committees of 3 people that can be chosen from 10 is: A. 30 B. 60 C. 120 D. 720
4. Which of the following scenarios involves combinations rather than permutations? A. Arranging 5 students in a line B. Selecting 3 players from a team of 10 C. Arranging digits to form a number D. None of the above
5. The main difference between permutations and combinations is that: A. Permutations ignore order, combinations consider order B. Permutations consider order, combinations ignore order C. Both consider order D. Neither considers order

1.3 Probability Laws

Probability laws provide the foundation for calculating the likelihood of events. A **probability law** is a rule that governs how probabilities are assigned and combined. These laws ensure consistency and allow us to solve problems involving uncertainty. The two most important laws are the **addition law** and the **multiplication law**.

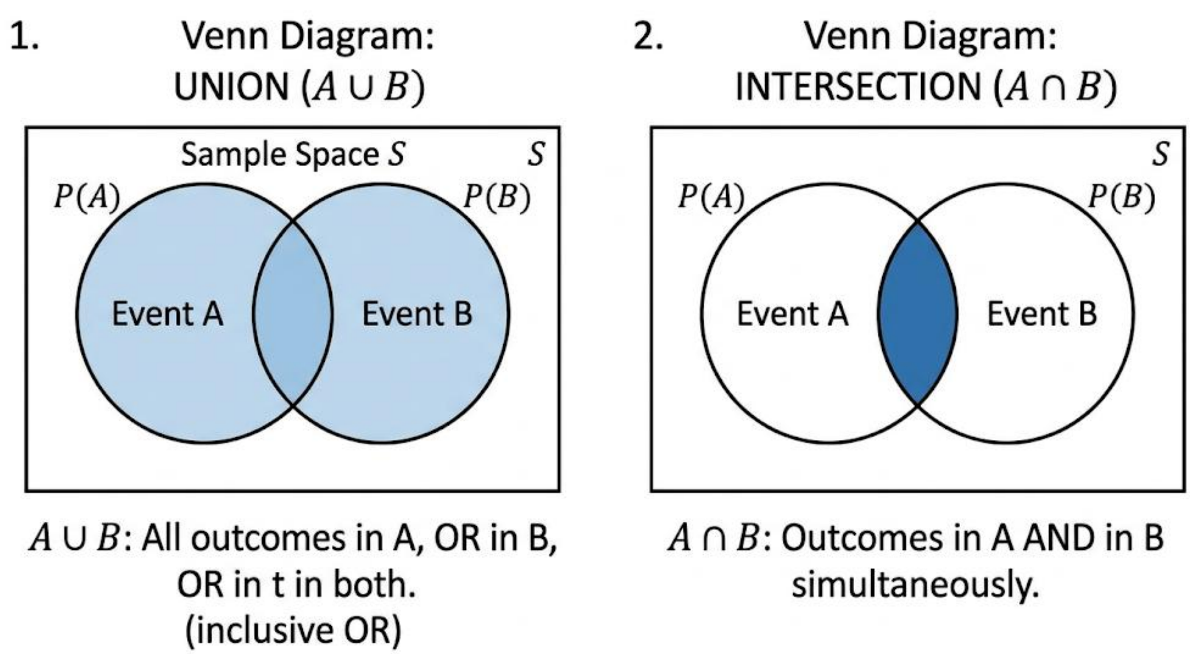
Fill in the blank: The two most important probability laws are the __________ law and the multiplication law.

1.3.1 Definition of probability laws

The addition and multiplication laws are derived from the basic definition of probability. Probability measures the chance of an event occurring, expressed as a number between 0 and 1.

- The **addition law of probability** deals with the probability of the union of two events.
- The **multiplication law of probability** deals with the probability of the intersection of two events.

Fill in the blank: The addition law of probability deals with the probability of the __________ of two events.



Union and intersection of two events in probability.

1.3.2 Addition and multiplication laws of probability

The **addition law** states:

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

This formula ensures that overlapping probabilities are not counted twice.

The **multiplication law** states:

$$P(A\cap B)=P(A)\times P(B|A)$$

where $P(B|A)$is the conditional probability of Bgiven A. If $A$and $B$are independent, then:

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

Probability laws

Law	**Formula**	**Description**	
Addition law	$P(A\cup B)=P(A)+P(B)-P(A\cap B)$	Probability of union of two events	
Multiplication law	\(P(A \cap B) = P(A) X P(B	A	Probability of intersection of two events
Independence case	$P(A\cap B)=P(A)\times P(B)$	Events do not influence each other	

1.3.3 Applications of probability laws in problem solving

Probability laws are applied in diverse areas such as risk assessment, quality control, and decision making. For example, in card games, the addition law helps calculate the probability of drawing a card that is either a heart or a face card. The multiplication law helps calculate the probability of drawing two specific cards in succession.

Fill in the blank: The addition law helps calculate the probability of the ___________ of two events.

Applications of probability laws

- Calculating probabilities in card games

- Assessing risks in insurance
- Analysing outcomes in experiments
- Supporting decision making in uncertain situations

Pilot questions 1.3

1. The addition law of probability is expressed as: A. $P(A\cup B)=P(A)+P(B)$B. $P(A\cup B)=P(A)+P(B)-P(A\cap B)$C. $P(A\cup B)=P(A)\times P(B)$D. None of the above
2. The multiplication law of probability is expressed as: A. $P(A\cap B)=P(A)+P(B)$B. $P(A\cap B)=P(A)\times P(B|A)$C. $P(A\cap B)=P(A)-P(B)$D. None of the above
3. If two events are independent, then $P(A\cap B)$equals: A. $P(A)+P(B)$B. $P(A)\times P(B)$C. $P(A)-P(B)$D. None of the above
4. Which probability law is used to calculate the probability of the union of two events? A. Multiplication law B. Addition law C. Independence law D. None of the above
5. Which of the following is an application of probability laws? A. Calculating probabilities in card games B. Designing seating arrangements C. Counting letters in a word D. None of the above

1.4 Applications of Permutations, Combinations and Probability Laws

Mathematics becomes most meaningful when abstract concepts are applied to real-life situations. The principles of **permutations**, **combinations**, and **probability laws** are not only theoretical but also practical tools used in everyday decision making, scientific research, and professional practice. This Part will highlight how these concepts are applied in diverse contexts, showing their relevance beyond the classroom.

Fill in the blank: Permutations, combinations, and probability laws are practical tools used in everyday __________ making.

1.4.1 Real-life applications in probability modelling

Permutations are applied in situations where order matters, such as arranging seating plans, designing passwords, or scheduling tasks. Combinations are used in contexts where order does not matter, such as lottery draws, team selections, or sampling in surveys. Probability laws underpin risk assessment in insurance, forecasting in business, and decision making in uncertain environments.

For example, in a lottery game where six numbers are chosen from a pool of 49, the number of possible outcomes is calculated using combinations. Similarly, in arranging exam timetables, permutations help determine possible sequences of subjects.

Fill in the blank: In a lottery game where six numbers are chosen from 49, the number of possible outcomes is calculated using __________.

Application of combinations in lottery draws.

How Combinations Work in a Lottery

We use the standard combination formula (often stated as "n choose k") to calculate the number of possible winning tickets:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Where:

- n = Total number of balls in the machine (e.g., 50)
- k = Number of balls drawn in a single selection (e.g., 5)

Why Order Does NOT Matter (Permutations vs. Combinations)

To understand the difference, let's contrast combinations with permutations:

- **Permutations (Order Matters):** If a lottery were run as a permutation, there would be many more possible outcomes. If you had to guess not only the 5 correct numbers but also the *exact order* they were drawn in (first, second, third, etc.), your chances of winning would plummet.
- **Combinations (Order Does Not Matter):** Since order is irrelevant, all the different permutations of the same 5 numbers are compressed into a **single unique combination**. As you can see in the diagram above, multiple distinct sequential drawings all result in the same winning set.

1.4.2 Problem-solving strategies using counting and probability laws

In problem solving, permutations and combinations provide systematic ways to count possibilities, while probability laws help calculate the likelihood of events. For instance, in card games, combinations determine the number of possible hands, while probability laws calculate the chance of drawing a specific card.

In business, probability laws are used to estimate risks, such as the likelihood of product failure or market fluctuations. In science, they are applied in genetics to predict inheritance patterns.

Fill in the blank: In card games, combinations determine the number of possible __________.

Applications of counting and probability laws

Scenario	Concept applied
Lottery draws	Combinations
Exam timetables	Permutations
Card games	Combinations and probability laws
Insurance risk	Probability laws

1.4.3 Combined applications in discrete probability

Often, permutations, combinations, and probability laws are used together. For example, in calculating the probability of winning a card game, combinations are used to count possible hands, permutations to arrange sequences of plays, and probability laws to determine the overall likelihood of success.

Another example is in computer science, where permutations and combinations are used in algorithm design, and probability laws are applied in simulations and modelling.

Fill in the blank: In computer science, permutations and combinations are used in algorithm design, while probability laws are applied in __________.

[Insert box here] Caption: Box 1.4: Combined applications of permutations, combinations, and probability laws

- Card games: counting hands, arranging plays, calculating probabilities
- Genetics: predicting inheritance patterns
- Computer science: algorithm design and simulations

- Business: risk modelling and forecasting

Summary: Combining Counting and Probability

To find the probability $P(A)$ of an event, we typically use the ratio of the number of successful outcomes to the total number of possible outcomes in the sample space S:

$$P(A) = \frac{\text{Number of favorable outcomes}}{\text{Total number of outcomes in } S}$$

Concept	Application in Probability	Formula/Rule
Permutations	Used when the **order** of outcomes matters (e.g., exact sequence of winning horses).	$P(A) = \frac{1}{nPr}$
Combinations	Used when **order does not matter** (e.g., drawing a specific hand of cards).	$P(A) = \frac{nCr\ (\text{favorable})}{nCr\ (\text{total})}$
Addition Law	Calculating the probability of Event A **OR** Event B occurring.	$P(A \cup B) = P(A) + P(B) - P(A \cap B)$
Multiplication Law	Calculating the probability of Event A **AND** Event B occurring.	$P(A \cap B) = P(A) \times P(B

Pilot questions 1.4

1. Which concept is used to calculate the number of possible lottery outcomes? A. Permutations B. Combinations C. Probability laws D. None of the above
2. Which concept is applied in arranging exam timetables? A. Permutations B. Combinations C. Probability laws D. None of the above
3. In card games, combinations are used to determine: A. The number of possible hands B. The order of seating arrangements C. The probability of market fluctuations D. None of the above

4. Which of the following is an application of probability laws? A. Designing seating arrangements B. Forecasting market risks C. Selecting lottery numbers D. None of the above
5. In computer science, permutations and combinations are used in algorithm design, while probability laws are applied in: A. Sampling B. Simulations C. Scheduling D. None of the above

Series Summary

This Series has introduced you to advanced permutations, combinations, and probability laws, which form the foundation of many probability problems. We began by examining permutations, focusing on their definition, properties, and applications, including cases with restrictions where order is important. We then moved on to combinations, highlighting how they differ from permutations by ignoring order, and explored their applications in selections such as committees and lottery draws. Following that, we studied probability laws, particularly the addition and multiplication laws, and saw how they are applied to calculate unions and intersections of events. Finally, we concluded with real-life applications, showing how permutations, combinations, and probability laws are used together in modelling, risk assessment, and problem solving.

By completing this Series, you should now be able to define and apply permutations and combinations, distinguish between them, explain the addition and multiplication laws of probability, and evaluate how these concepts are applied in real-life contexts. These skills directly align with the Series Learning Outcomes, equipping you with both theoretical knowledge and practical problem-solving strategies in probability.

Glossary of Terms

- **Combination**: A selection of objects from a set where order does not matter.
- **Factorial**: The product of all positive integers up to a given number, denoted by n!.
- **Permutation**: An arrangement of objects in a specific order.
- **Probability law**: A rule that governs how probabilities are assigned and combined, such as the addition and multiplication laws.
- **Union of events**: The set of outcomes that belong to either of two events or both.

- **Intersection of events**: The set of outcomes that belong to both events simultaneously.

Concept Check Questions 1

Objective questions

1. The formula for permutations of nobjects taken rat a time is: A. n!r!B. n!(n-r)!C. nrD. None of the above (Linked to SLO 1.1)
2. The formula for combinations of nobjects taken rat a time is: A. n!(n-r)!B. n!r!(n-r)!C. nrD. None of the above (Linked to SLO 1.3)
3. The addition law of probability is expressed as: A. P(A∪B)=P(A)+P(B)B. P(A∪B)=P(A)+P(B)-P(A∩B)C. P(A∪B)=P(A)×P(B)D. None of the above (Linked to SLO 1.5)
4. Which concept is used to calculate the number of possible lottery outcomes? A. Permutations B. Combinations C. Probability laws D. None of the above (Linked to SLO 1.7)
5. Which of the following scenarios involves permutations rather than combinations? A. Selecting 3 players from a team of 10 B. Arranging 5 students in a line C. Choosing 5 cards from a deck D. None of the above (Linked to SLO 1.2)

Short-answer questions

6. Define a permutation. (Linked to SLO 1.1)
7. Give one example of a combination in real life. (Linked to SLO 1.3)
8. Explain the difference between permutations and combinations. (Linked to SLO 1.4)
9. State the multiplication law of probability. (Linked to SLO 1.5)
10. Provide one application of probability laws in business. (Linked to SLO 1.6)

Long-answer questions

11. Analyse the properties of permutations and explain their importance in counting problems. (Linked to SLO 1.1)
12. Evaluate the role of combinations in probability modelling, with examples. (Linked to SLO 1.3)

13. Discuss the addition and multiplication laws of probability, highlighting their applications in real-life scenarios. (Linked to SLO 1.5)

14. Examine combined applications of permutations, combinations, and probability laws in discrete probability problems. (Linked to SLO 1.7)

Answers to Concept Check Questions 1

Objective questions

1. B
2. B
3. B
4. B
5. B

Short-answer questions

6. A permutation is an arrangement of objects in a specific order.

7. Choosing a committee of 3 people from a group of 10.

8. Permutations consider order, while combinations do not.

9. $P(A \cap B) = P(A) \times P(B|A)$. If independent, $P(A \cap B) = P(A) \times P(B)$.

10. Probability laws are used in risk assessment, such as estimating the likelihood of product failure.

Long-answer questions

11. *Basic response*: Permutations involve arrangements where order matters, using factorials to calculate possibilities. *Value-added response*: They are important in scheduling, seating arrangements, and code design, ensuring systematic counting in ordered contexts.

12. *Basic response*: Combinations involve selections where order does not matter. *Value-added response*: They are crucial in lottery draws, team selections, and sampling, providing accurate counts in unordered contexts.

13. *Basic response*: The addition law calculates unions, while the multiplication law calculates intersections. *Value-added response*: These laws are applied in card games, insurance, and scientific experiments, supporting decision making under uncertainty.

14. *Basic response*: Permutations, combinations, and probability laws are often used together in discrete probability problems. *Value-added response*: Their combined use is seen in genetics, computer science algorithms, and business forecasting, where counting and probability laws interact to model complex scenarios.

Answers to Pilot Questions 1

Part 1.1: 1-B, 2-B, 3-C, 4-B, 5-B **Part 1.2**: 1-B, 2-B, 3-C, 4-B, 5-B **Part 1.3**: 1-B, 2-B, 3-B, 4-B, 5-A **Part 1.4**: 1-B, 2-A, 3-A, 4-B, 5-B

References and Attributions 1

- **Figures, Tables, and Boxes**:
    - Figure 1.1: Illustration of permutations of three objects (AI-generated, prompt: "Generate a diagram showing all possible arrangements of three coloured balls labelled A, B, and C").
    - Table 1.1: Examples of permutations (AI-generated, prompt: "Generate a table showing examples of permutations with letters and students").
    - Figure 1.2: Illustration of combinations of three objects (AI-generated, prompt: "Generate a diagram showing combinations of three coloured balls where order does not matter").
    - Table 1.2: Examples of combinations (AI-generated, prompt: "Generate a table showing examples of combinations with letters, committees, and cards").
    - Figure 1.3: Union and intersection of two events in probability (AI-generated, prompt: "Generate a Venn diagram showing union and intersection of two events A and B").
    - Table 1.3: Probability laws (AI-generated, prompt: "Generate a table showing addition and multiplication laws of probability with formulas and descriptions").
    - Figure 1.4: Application of combinations in lottery draws (AI-generated, prompt: "Generate a diagram showing lottery balls being selected to illustrate combinations").
    - Table 1.4: Applications of counting and probability laws (AI-generated, prompt: "Generate a table showing applications of permutations, combinations, and probability laws in real life").

- Box 1.1, 1.2, 1.3, 1.4: AI-generated text boxes summarising restricted permutations, differences between permutations and combinations, applications of probability laws, and combined applications.

Series 2: Conditional Probability, Independence and Bayes' Theorem.

- **Series title**: Conditional Probability, Independence and Bayes' Theorem
- **Series outline**:
    - **2.1 Conditional probability**
        - 2.1.1 Definition of conditional probability
        - 2.1.2 Properties and examples
        - 2.1.3 Applications in real-life problems
    - **2.2 Independence of events**
        - 2.2.1 Definition of independence
        - 2.2.2 Properties and examples
        - 2.2.3 Distinction between independence and conditional probability
    - **2.3 Bayes' theorem**
        - 2.3.1 Statement of Bayes' theorem
        - 2.3.2 Applications in probability problems
        - 2.3.3 Real-life applications of Bayes' theorem

Series Learning Outcomes

Upon completing **Series 2: Conditional Probability, Independence and Bayes' Theorem**, you should be able to:

- **SLO 2.1** define conditional probability and explain its properties,
- **SLO 2.2** apply conditional probability to solve practical problems,
- **SLO 2.3** define independence of events and explain its properties,
- **SLO 2.4** distinguish between independence and conditional probability with examples,
- **SLO 2.5** state Bayes' theorem clearly,

- **SLO 2.6** apply Bayes' theorem to solve probability problems, and
- **SLO 2.7** evaluate real-life applications of Bayes' theorem in contexts such as medical testing and forecasting.

2.1 Conditional Probability

Conditional probability is one of the most important concepts in probability theory. **Conditional probability** is the probability of an event occurring given that another event has already occurred. It allows us to refine our predictions when we have additional information. For example, the probability of a student passing a course may change if we know that the student has attended most of the lectures.

Fill in the blank: Conditional probability is the probability of an event occurring given that another __________ has already occurred.

2.1.1 Definition of conditional probability

The conditional probability of event Agiven event Bis denoted by $P(A|B)$. It is defined as:

$$P(A|B)=P(A\cap B)P(B),\text{provided } P(B)>0$$

Here, $P(A\cap B)$is the probability that both events Aand Boccur, while $P(B)$is the probability that event Boccurs.

Fill in the blank: The formula for conditional probability is P(A|B) = \frac{P(A \cap B)}{P(____)}.

How to Interpret the Diagram

- **The Original Sample Space (S):** The entire rectangle represents all possible outcomes. The probability of any outcome in S is 1.
- **The Conditional "New" Sample Space (B):** Because we are given that event B *has* occurred, we only focus on the region covered by circle B. This is now our entire universe for calculating the probability.
- **The Successful Outcomes (A ∩ B):** Within this new sample space B, the only way for event A to also occur is if the outcome falls into the shaded intersection where circle A and circle B overlap.
- **Calculating $P(A|B)$:** The conditional probability is the size of the "successful" overlap ($A \cap B$) relative to the size of the entire conditional space (B).

Illustration of conditional probability using overlapping events.

2.1.2 Properties and examples

Conditional probability has several important properties:

- If Aand Bare independent, then $P(A|B)=P(A)$.
- $P(A|B)$can be used to update probabilities when new information is available.

Example 1: Suppose the probability of a student passing mathematics is 0.6, and the probability of passing both mathematics and statistics is 0.4. The probability of passing mathematics given that the student passed statistics is:

$$P(\text{Math}|\text{Stats})=P(\text{MathStats})P(\text{Stats})$$

Fill in the blank: If two events are independent, then P(A|B) = P(____).

[Insert table here] Caption: Table 2.1: Examples of conditional probability

Real-World Examples of Conditional Probability

1. Medical Diagnostics (The False Positive Paradox)

This is a classic "exam-style" application. Suppose a rare disease affects **1%** of the population. A test for this disease is **99% accurate**.

- **The Condition:** You just tested positive.
- **The Probability:** Even with a "99% accurate" test, the probability that you *actually* have the disease might only be **50%**. This is because the small number of true positives is balanced against the small percentage of false positives from the much larger healthy population.

2. Weather Forecasting

Meteorologists constantly use updated data to shift probabilities.

- **Event A:** It will rain today.
- **Event B (The Condition):** Dark clouds have formed, and the barometric pressure has dropped.
- **The Shift:** $P(\text{Rain})$ might generally be 10% for a month, but $P(\text{Rain} \mid \text{Dark Clouds})$ might jump to 80%.

3. Card Games (Blackjack/Poker)

In games played without replacement, every card dealt changes the probability of the next card.

- **The Condition:** You are dealt a King.
- **The Probability:** Before the first card, $P(\text{King}) = 4/52$. Once you hold one King, the probability that the next card is also a King drops to $P(\text{King} \mid 1 \text{ King out}) = 3/51$.

Formal Logic and Formula

The notation for conditional probability is $P(A \mid B)$, read as "the probability of A given B."

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

Component	Meaning
$P(A \cap B)$	The probability that **both** events A and B happen.
$P(B)$	The probability of the condition (the "given" event) happening.
The Logic	You are measuring how much of "Event B" is actually made up of "Event A."

2.1.3 Applications in real-life problems

Conditional probability is widely applied in real life. In medicine, it is used to calculate the probability of a patient having a disease given a positive test result. In business, it helps in forecasting sales given certain market conditions. In sports, it can be used to estimate the probability of winning given past performance.

Fill in the blank: In medicine, conditional probability is used to calculate the probability of a patient having a __________ given a positive test result.

Applications of conditional probability

- Medical testing and diagnosis
- Business forecasting under market conditions
- Sports performance analysis
- Risk assessment in insurance

Pilot questions 2.1

1. The formula for conditional probability is: A. $P(A|B)=P(A)P(B)$B. $P(A|B)=P(A\cap B)P(B)$C. $P(A|B)=P(A)\times P(B)$D. None of the above
2. If two events are independent, then $P(A|B)$equals: A. $P(A)$B. $P(B)$C. $P(A\cap B)$D. None of the above

3. Which of the following is an application of conditional probability? A. Designing seating arrangements B. Medical testing and diagnosis C. Counting letters in a word D. None of the above

4. The notation P(A|B)represents: A. Probability of A and B occurring together B. Probability of A occurring given B has occurred C. Probability of B occurring given A has occurred D. None of the above

5. In card games, the probability of drawing a red card given that it is a heart is equal to: A. 0.5 B. 1 C. 0.25 D. None of the above

2.2 Independence of Events

Independence is another central concept in probability. Two events are said to be **independent** if the occurrence of one does not affect the probability of the other. For example, tossing a coin and rolling a die are independent events because the outcome of the coin toss does not influence the outcome of the die roll.

Fill in the blank: Two events are independent if the occurrence of one does not affect the __________ of the other.

2.2.1 Definition of independence

Formally, events Aand Bare independent if:

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

This means that the probability of both events occurring together equals the product of their individual probabilities.

Fill in the blank: Events Aand Bare independent if P(A \cap B) = P(A) \times P(____).

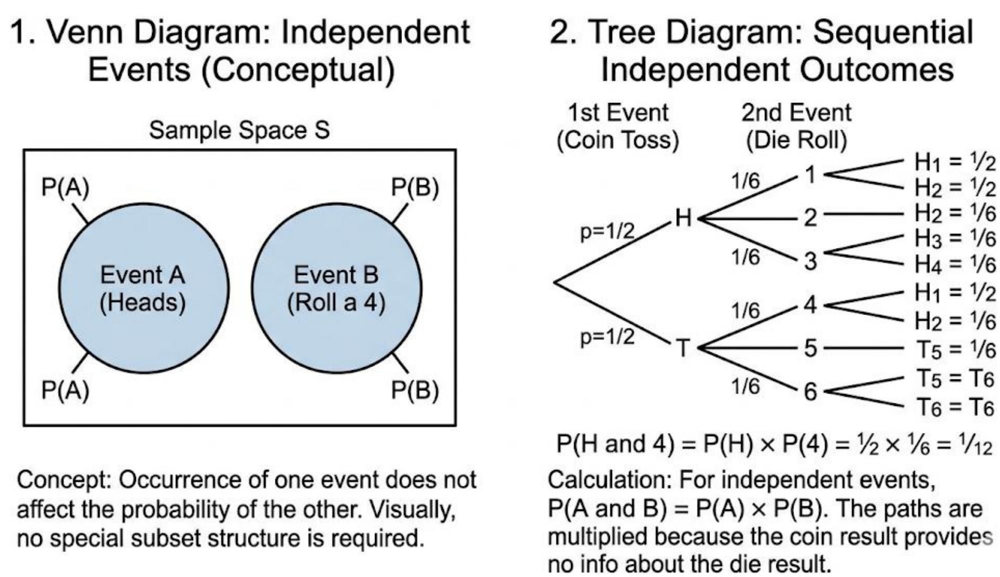


Illustration of independent events.

2.2.2 Properties and examples

Key properties of independence include:

- If Aand Bare independent, then $P(A|B)=P(A)$.
- Independence is symmetric: if Ais independent of B, then Bis independent of A.
- Independence is different from mutual exclusivity.

Example 1: Tossing a coin and rolling a die. The probability of getting heads is 0.5, and the probability of rolling a 4 is 16. Since these events are independent:

$$P(Heads4)=P(Heads)\times P(4)=0.5\times 16=112$$

Fill in the blank: If two events are independent, then $P(A|B)$equals __________.

Examples of independent events

Scenario	Probability
Tossing a coin and rolling a die	P(Heads4)=112
Drawing a card and tossing a coin	Independent outcomes

2.2.3 Distinction between independence and conditional probability

It is important to distinguish independence from conditional probability. Independence means that the occurrence of one event does not affect the other, while conditional probability measures the probability of one event given that another has occurred.

For example, tossing a coin and rolling a die are independent, but the probability of drawing a red card given that it is a heart is a conditional probability.

Fill in the blank: Independence means events do not affect each other, while conditional probability measures the probability of

Key differences between independence and conditional probability

- Independence: events do not influence each other.
- Conditional probability: probability of one event given another has occurred.
- Independence implies $P(A|B)=P(A)$.
- Conditional probability adjusts probabilities based on information.

Pilot questions 2.2

1. Events Aand Bare independent if: A. $P(A \cap B)=P(A)+P(B)$B. $P(A \cap B)=P(A) \times P(B)$C. $P(A \cap B)=P(A)-P(B)$D. None of the above
2. If two events are independent, then $P(A|B)$equals: A. $P(A)$B. $P(B)$C. $P(A \cap B)$D. None of the above
3. Which of the following scenarios involves independent events? A. Tossing a coin and rolling a die B. Drawing two cards without replacement C. Selecting students from the same class D. None of the above
4. Independence is different from: A. Conditional probability B. Mutual exclusivity C. Both A and B D. None of the above
5. Which statement is true about independence? A. Independence means events always occur together B. Independence means events do not influence each other C. Independence means probabilities are always equal D. None of the above

2.3 Bayes' Theorem

Bayes' theorem is a powerful result in probability theory that allows us to update probabilities when new information becomes available. It connects conditional probability with prior knowledge, making it especially useful in situations where we want to revise our beliefs after observing evidence. For example, in medical testing, Bayes' theorem helps calculate the probability that a patient truly has a disease given a positive test result.

Fill in the blank: Bayes' theorem allows us to update probabilities when new __________ becomes available.

2.3.1 Statement of Bayes' theorem

Bayes' theorem states that:

$$P(A|B)=P(B|A).P(A)P(B)$$

Here:

- $P(A|B)$is the probability of event Agiven event B.
- $P(B|A)$is the probability of event Bgiven event A.
- $P(A)$is the prior probability of event A.
- $P(B)$is the probability of event B.

2.3.2 Applications in probability problems

Bayes' theorem is applied in solving problems where conditional probabilities are involved. For example, suppose a factory produces 60% of its products from Machine A and 40% from Machine B. Machine A produces 2% defective items, while Machine B produces 5% defective items. If a defective item is found, Bayes' theorem can be used to calculate the probability that it came from Machine A.

Fill in the blank: Bayes' theorem is especially useful in problems involving __________ probabilities.

Example of Bayes' theorem in manufacturing

Machine	Proportion of output	Defective rate	Contribution to defectives
A	60%	2%	0.012
B	40%	5%	0.020

2.3.3 Real-life applications of Bayes' theorem

Bayes' theorem has wide applications in real life. In medicine, it is used to interpret diagnostic tests. In law, it helps assess the probability of guilt given evidence. In artificial intelligence, it underpins Bayesian networks used for decision making. In everyday life, it can be applied to revise expectations based on new information.

Fill in the blank: In artificial intelligence, Bayes' theorem underpins __________ networks used for decision making.

[Insert box here] Caption: Box 2.3: Real-life applications of Bayes' theorem

- Medical testing and diagnosis
- Legal reasoning with evidence
- Artificial intelligence and Bayesian networks
- Everyday decision making with updated information

Pilot questions 2.3

1. Bayes' theorem is expressed as: A. P(A∣B)=P(A)P(B)B. P(A∣B)=P(B∣A)⋅P(A)P(B)C. P(A∣B)=P(A)×P(B)D. None of the above
2. In Bayes' theorem, P(A)represents: A. Posterior probability B. Prior probability C. Likelihood D. None of the above
3. Which of the following is a real-life application of Bayes' theorem? A. Designing seating arrangements B. Medical testing and diagnosis C. Counting letters in a word D. None of the above
4. Bayes' theorem is especially useful in problems involving: A. Independent events B. Conditional probabilities C. Mutually exclusive events D. None of the above
5. In artificial intelligence, Bayes' theorem underpins: A. Neural networks B. Bayesian networks C. Decision trees D. None of the above

Series Summary

This Series has focused on conditional probability, independence of events, and Bayes' theorem, which are essential tools for analysing uncertainty. We began by exploring conditional probability, learning how to calculate the probability of one event given that another has occurred, and applying it to practical contexts such as medical testing and card games. We then examined independence of events, distinguishing it from conditional probability and showing how independent events do not influence each other. Finally, we studied Bayes' theorem, understanding its statement, solving probability problems with it, and appreciating its wide applications in medicine, law, and artificial intelligence.

By completing this Series, you should now be able to define and apply conditional probability, explain independence of events, distinguish independence from conditional probability, state Bayes' theorem, and evaluate its applications in real-life scenarios. These achievements align with the Series Learning Outcomes, equipping you with both theoretical knowledge and practical problem-solving skills in probability.

Glossary of Terms

- **Bayes' theorem**: A probability rule that allows updating of probabilities when new information is available.
- **Conditional probability**: The probability of an event occurring given that another event has already occurred.
- **Independence of events**: A situation where the occurrence of one event does not affect the probability of another.
- **Intersection of events**: The set of outcomes that belong to both events simultaneously.
- **Likelihood**: In Bayes' theorem, the probability of evidence occurring given a particular hypothesis.
- **Posterior probability**: The updated probability of an event after considering new information.
- **Prior probability**: The initial probability of an event before new information is considered.

Concept Check Questions 2

Objective questions

1. The formula for conditional probability is: A. $P(A|B)=P(A)P(B)$B. $P(A|B)=P(A\cap B)P(B)$C. $P(A|B)=P(A)\times P(B)$D. None of the above (Linked to SLO 2.1)
2. Events Aand Bare independent if: A. $P(A\cap B)=P(A)+P(B)$B. $P(A\cap B)=P(A)\times P(B)$C. $P(A\cap B)=P(A)-P(B)$D. None of the above (Linked to SLO 2.3)
3. Bayes' theorem is expressed as: A. $P(A|B)=P(A)P(B)$B. $P(A|B)=P(B|A)\cdot P(A)P(B)$C. $P(A|B)=P(A)\times P(B)$D. None of the above (Linked to SLO 2.5)
4. Which of the following is a real-life application of Bayes' theorem? A. Medical testing and diagnosis B. Designing seating arrangements C. Counting letters in a word D. None of the above (Linked to SLO 2.7)
5. If two events are independent, then $P(A|B)$equals: A. $P(A)$B. $P(B)$C. $P(A\cap B)$D. None of the above (Linked to SLO 2.4)

Short-answer questions

6. Define conditional probability. (Linked to SLO 2.1)
7. Give one example of independence of events. (Linked to SLO 2.3)
8. State Bayes' theorem. (Linked to SLO 2.5)

9. Explain the difference between independence and conditional probability. (Linked to SLO 2.4)

10. Provide one application of Bayes' theorem in artificial intelligence. (Linked to SLO 2.7)

Long-answer questions

11. Analyse the properties of conditional probability and explain its importance in problem solving. (Linked to SLO 2.2)

12. Evaluate the role of independence in probability modelling, with examples. (Linked to SLO 2.3)

13. Discuss Bayes' theorem, highlighting its applications in medicine and law. (Linked to SLO 2.6, 2.7)

14. Examine how conditional probability, independence, and Bayes' theorem work together in real-life decision making. (Linked to SLO 2.7)

Answers to Concept Check Questions 2

Objective questions

1. B
2. B
3. B
4. A
5. A

Short-answer questions

6. Conditional probability is the probability of an event occurring given that another event has already occurred.

7. Tossing a coin and rolling a die.

8. $P(A|B)=P(B|A).P(A)P(B)$.

9. Independence means events do not affect each other, while conditional probability measures the probability of one event given another has occurred.

10. Bayesian networks used for decision making.

Long-answer questions

11. *Basic response*: Conditional probability refines predictions when new information is available. *Value-added response*: It is crucial in medical testing, business forecasting, and sports analysis, where probabilities depend on prior conditions.

12. *Basic response*: Independence means events do not influence each other. *Value-added response*: It is important in modelling random processes, such as coin tosses and dice rolls, ensuring accurate probability calculations.

13. *Basic response*: Bayes' theorem updates probabilities using prior and conditional information. *Value-added response*: In medicine, it helps interpret diagnostic tests; in law, it supports reasoning with evidence.

14. *Basic response*: Conditional probability, independence, and Bayes' theorem are interconnected concepts in probability. *Value-added response*: Together, they underpin decision making in fields such as artificial intelligence, genetics, and risk assessment, where updated and accurate probabilities are essential.

Answers to Pilot Questions 2

Part 2.1: 1-B, 2-A, 3-B, 4-B, 5-B **Part 2.2**: 1-B, 2-A, 3-A, 4-B, 5-B **Part 2.3**: 1-B, 2-B, 3-B, 4-B, 5-B

References and Attributions 2

- **Figures, Tables, and Boxes**:
    - Figure 2.1: Illustration of conditional probability using overlapping events (AI-generated, prompt: "Generate a Venn diagram showing conditional probability with event A overlapping event B").
    - Table 2.1: Examples of conditional probability (AI-generated, prompt: "Generate a table showing examples of conditional probability in exams and card games").
    - Figure 2.2: Illustration of independent events (AI-generated, prompt: "Generate a diagram showing independence of coin toss and die roll outcomes").
    - Table 2.2: Examples of independent events (AI-generated, prompt: "Generate a table showing examples of independent events with coin toss and die roll").
    - Figure 2.3: Illustration of Bayes' theorem components (AI-generated, prompt: "Generate a diagram showing Bayes' theorem with prior, likelihood, and posterior probabilities").
    - Table 2.3: Example of Bayes' theorem in manufacturing (AI-generated, prompt: "Generate a table showing Bayes' theorem example with two machines and defective rates").
    - Box 2.1, 2.2, 2.3: AI-generated text boxes summarising applications of conditional probability, differences between independence and conditional probability, and real-life applications of Bayes' theorem.
- **Search strings used**:
    - "conditional probability Venn diagram OER"
    - "conditional probability examples table OER"
    - "independent events coin toss die roll OER"
    - "examples of independent events probability OER"
    - "Bayes theorem diagram prior likelihood posterior OER"
    - "Bayes theorem manufacturing example OER"

- "Bayes theorem real life applications OER"

- **Attribution**: Explanations and instructional content are based on standard open educational resources in probability theory. AI-generated illustrations were created using prompts as noted above.

Series 4 Outline

4.1 Expectations

- 4.1.1 Definition of expectation
- 4.1.2 Properties of expectation
- 4.1.3 Applications of expectation

4.2 Moments of random variables

- 4.2.1 Definition of moments
- 4.2.2 Variance and standard deviation
- 4.2.3 Higher-order moments and their applications

4.3 Inequalities in probability

- 4.3.1 Markov's inequality
- 4.3.2 Chebyshev's inequality
- 4.3.3 Applications of probability inequalities

Introduction

Expectations, moments, and inequalities are fundamental concepts in probability theory that help us understand the behaviour of random variables beyond just their distributions. The **expectation** (or mean) provides a measure of the central tendency of a random variable. **Moments** (such as variance and higher-order moments) describe the spread and shape of distributions. **Inequalities in probability** (such as Markov's and Chebyshev's inequalities) provide bounds on probabilities, offering powerful tools when exact calculations are difficult.

The aim of this Series is to enable you to compute expectations and moments of random variables, understand their significance, and apply probability inequalities to solve problems and interpret real-life scenarios.

We shall commence this Series by examining expectations, beginning with their definition, properties, and applications. Next, we will move on to moments of random variables, including variance, standard deviation, and higher-order moments. Finally, we will conclude with inequalities in probability, studying Markov's inequality, Chebyshev's inequality, and their applications in bounding probabilities

Series Learning Outcomes

Upon completing **Series 4: Expectations, Moments and Inequalities in Probability**, you should be able to:

- **SLO 4.1** define expectation of a random variable and explain its properties,
- **SLO 4.2** compute expectations for discrete and continuous random variables,
- **SLO 4.3** define and explain moments of random variables, including variance and standard deviation,
- **SLO 4.4** calculate variance, standard deviation, and higher-order moments,
- **SLO 4.5** state and apply Markov's inequality,
- **SLO 4.6** state and apply Chebyshev's inequality, and
- **SLO 4.7** evaluate real-life applications of expectations, moments, and inequalities in probability.

4.1 Expectations

The **expectation** (or mean) of a random variable is the long-run average value it takes when an experiment is repeated many times. It provides a measure of the central tendency of the distribution. For example, if you roll a fair die many times, the average outcome tends toward 3.5, which is the expectation of the die roll.

Fill in the blank: The expectation of a random variable is the long-run __________ value it takes when an experiment is repeated many times.

4.1.1 Definition of expectation

For a **discrete random variable** Xwith possible values xiand probabilities pi:

E[X]=ixipi

For a **continuous random variable** Xwith probability density function f(x):

$$E[X]=-xf(x)\ dx$$

Fill in the blank: For a discrete random variable, the expectation is calculated as E[X]=ixipi.

4.1.2 Properties of expectation

Key properties include:

- **Linearity**: $E[aX+bY]=aE[X]+bE[Y]$.
- **Constant rule**: $E[c]=c$, where cis a constant.
- **Non-negativity**: If $X\geq 0$, then $E[X]\geq 0$.

Fill in the blank: The expectation operator is __________, meaning $E[aX+bY]=aE[X]+bE[Y]$.

Properties of expectation

Property	Formula	Example
Linearity	$E[aX+bY]=aE[X]+bE[Y]$	Weighted averages
Constant rule	$E[c]=c$	Expectation of fixed value
Non-negativity	If $X\geq 0$, then $E[X]\geq 0$	Positive random variables

4.1.3 Applications of expectation

Expectation is widely applied in real life:

- In economics, expected profit guides business decisions.
- In insurance, expected claims determine premium rates.
- In gambling, expectation helps evaluate fairness of games.
- In science, expectation is used in statistical estimation.

Fill in the blank: In insurance, expected __________ determine premium rates.

Applications of expectation

- Economics: expected profit
- Insurance: expected claims
- Gambling: fairness of games
- Science: statistical estimation

Pilot questions 4.1

1. The expectation of a random variable is: A. The maximum value it can take B. The long-run average value C. The minimum possible outcome D. None of the above
2. For a discrete random variable, expectation is calculated as: A. E[X]=ipiB. E[X]=ixipiC. E[X]=ixiD. None of the above
3. The expectation operator is: A. Non-linear B. Linear C. Constant D. None of the above
4. In insurance, expectation is used to determine: A. Premium rates B. Customer satisfaction C. Marketing strategies D. None of the above
5. Which of the following is a real-life application of expectation? A. Designing seating arrangements B. Evaluating fairness of gambling games C. Counting letters in a word D. None of the above

4.2 Moments of Random Variables

Moments are numerical measures that describe the shape and spread of a probability distribution. While expectation (mean) captures the central tendency, **moments** provide deeper insights into variability, skewness, and kurtosis.

Fill in the blank: Moments are numerical measures that describe the __________ and spread of a probability distribution.

4.2.1 Definition of moments

The r-th moment of a random variable Xabout the origin is defined as:

$$r'=E[Xr]$$

The r-th moment about the mean (central moment) is defined as:

$$r=E[(X-\mu)r]$$

Where $\mu=E[X]$.

Fill in the blank: The second central moment of a random variable is its __________.

4.2.2 Variance and standard deviation

- **Variance** measures the spread of a distribution:

$$Var(X)=E[(X-\mu)2]$$

- **Standard deviation** is the square root of variance:

$$\sigma = \text{Var}(X)$$

Variance and standard deviation quantify how much values deviate from the mean.

Fill in the blank: The square root of variance is called the __________ deviation.

Curve Color	Standard Deviation (σ)	Variance (σ2)	Visual Characteristic
Purple	= 0.5	2 = 0.25	**Tall and Narrow:** Most data points cluster very closely to the mean.
Light Blue	=1.0	2= 1.0	**Standard Normal:** This is the baseline shape of the bell curve.
Orange	σ= 2.0	2= = 4.0	**Short and Wide:** The data points are highly dispersed and spread far from the mean.

Variance and standard deviation as measures of spread.

4.2.3 Higher-order moments and their applications

- **Third central moment (skewness)**: Measures asymmetry of the distribution.
- **Fourth central moment (kurtosis)**: Measures the "peakedness" or heaviness of tails.

Applications:

- Skewness helps detect bias in data.
- Kurtosis helps identify whether data has heavy tails (risk assessment in finance).

Fill in the blank: The third central moment of a distribution is called __________.

Moments of random variables

Moment	Formula	Interpretation
First	$E[X]$	Mean
Second	$E[(X-\mu)2]$	Variance
Third	$E[(X-\mu)3]$	Skewness
Fourth	$E[(X-\mu)4]$	Kurtosis

Pilot questions 4.2

1. The second central moment of a random variable is: A. Mean B. Variance C. Skewness D. None of the above
2. The square root of variance is called: A. Mean B. Standard deviation C. Kurtosis D. None of the above
3. The third central moment of a distribution is: A. Variance B. Skewness C. Kurtosis D. None of the above
4. Variance measures: A. Central tendency B. Spread of distribution C. Symmetry of distribution D. None of the above
5. Kurtosis measures: A. Peakedness or heaviness of tails B. Average value C. Number of trials D. None of the above

4.3 Inequalities in Probability

Probability inequalities provide bounds on probabilities when exact calculations are difficult or impossible. They are powerful tools for estimating risks and making decisions under uncertainty. Two of the most important inequalities are **Markov's inequality** and **Chebyshev's inequality**.

Fill in the blank: Probability inequalities provide __________ on probabilities when exact calculations are difficult.

4.3.1 Markov's inequality

Markov's inequality states that for a non-negative random variable Xand any a>0:

$$P(X \geq a) \leq E[X]a$$

This inequality provides an upper bound on the probability that a random variable exceeds a certain value.

Example: If the expected number of defective items in a batch is 2, then the probability of having at least 10 defective items is at most 210=0.2.

Fill in the blank: Markov's inequality applies to __________ random variables.

4.3.2 Chebyshev's inequality

Chebyshev's inequality states that for any random variable Xwith mean and variance 2:

$$P(|X-\mu| \geq k\sigma) \leq 1k2$$

This inequality provides a bound on the probability that a random variable deviates from its mean by more than kstandard deviations.

Example: For any distribution, at least 75% of values lie within 2 standard deviations of the mean (since 1-122=0.75).

Fill in the blank: Chebyshev's inequality provides a bound on the probability that a random variable deviates from its __________ by more than kstandard deviations.

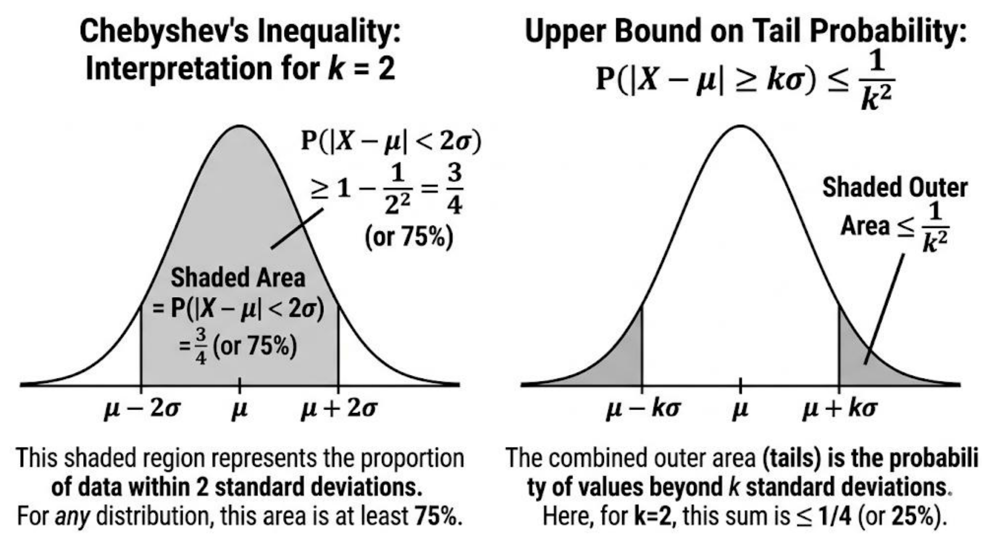


Illustration of Chebyshev's inequality.

4.3.3 Applications of probability inequalities

- **Risk management**: Bounding probabilities of extreme losses in finance.
- **Quality control**: Estimating probability of large deviations in manufacturing.
- **Statistics**: Providing guarantees when exact distributions are unknown.

- **Computer science**: Analysing algorithms with probabilistic performance.

Fill in the blank: In finance, probability inequalities are used to bound probabilities of extreme __________.

Applications of probability inequalities

- Finance: bounding extreme losses
- Manufacturing: quality control
- Statistics: guarantees without exact distributions
- Computer science: algorithm analysis

Pilot questions 4.3

1. Markov's inequality applies to: A. Negative random variables B. Non-negative random variables C. Continuous random variables only D. None of the above
2. Markov's inequality states: A. P(X≥a)≤E[X]aB. P(X≥a)=E[X]aC. P(X≥a)≥E[X]aD. None of the above
3. Chebyshev's inequality provides a bound on: A. Probability of exact outcomes B. Probability of deviations from the mean C. Probability of independent events D. None of the above
4. According to Chebyshev's inequality, at least 75% of values lie within: A. 1 standard deviation of the mean B. 2 standard deviations of the mean C. 3 standard deviations of the mean D. None of the above
5. In finance, probability inequalities are used to bound probabilities of: A. Customer satisfaction B. Extreme losses C. Marketing strategies D. None of the above

Series Summary

This Series explored **expectations, moments, and inequalities in probability**, which are essential tools for understanding the behaviour of random variables. We began with expectations, defining them as the long-run average value of a random variable, examining their properties, and highlighting applications in economics, insurance, gambling, and science. Next, we studied moments of random variables, including variance, standard deviation, skewness, and kurtosis, which describe spread, asymmetry, and peakedness of distributions. Finally, we examined inequalities in probability, focusing on Markov's and Chebyshev's inequalities, which provide bounds on probabilities when exact calculations are difficult, with applications in risk management, quality control, statistics, and computer science.

By completing this Series, you should now be able to compute expectations and moments, interpret their significance, and apply probability inequalities to solve problems and evaluate real-life scenarios.

Glossary of Terms

- **Expectation (mean)**: The long-run average value of a random variable.
- **Linearity of expectation**: Property that $E[aX+bY]=aE[X]+bE[Y]$.
- **Moment**: A numerical measure describing the shape of a distribution.
- **Variance**: The second central moment, measuring spread around the mean.
- **Standard deviation**: The square root of variance, quantifying deviation from the mean.
- **Skewness**: The third central moment, measuring asymmetry of a distribution.
- **Kurtosis**: The fourth central moment, measuring peakedness or heaviness of tails.
- **Markov's inequality**: Bound on the probability that a non-negative random variable exceeds a given value.
- **Chebyshev's inequality**: Bound on the probability that a random variable deviates from its mean by more than a given number of standard deviations.

Concept Check Questions 4

Objective questions

1. The expectation of a random variable is: A. Maximum value B. Long-run average value C. Minimum value D. None of the above (Linked to SLO 4.1)
2. The second central moment of a random variable is: A. Mean B. Variance C. Skewness D. None of the above (Linked to SLO 4.3)
3. The square root of variance is called: A. Mean B. Standard deviation C. Kurtosis D. None of the above (Linked to SLO 4.4)

4. Markov's inequality applies to: A. Negative random variables B. Non-negative random variables C. Continuous random variables only D. None of the above (Linked to SLO 4.5)
5. Chebyshev's inequality provides a bound on: A. Exact outcomes B. Deviations from the mean C. Independent events D. None of the above (Linked to SLO 4.6)

Short-answer questions

6. Define expectation of a random variable. (Linked to SLO 4.1)

7. State the formula for variance. (Linked to SLO 4.4)

8. What does skewness measure? (Linked to SLO 4.3)

9. State Markov's inequality. (Linked to SLO 4.5)

10. Provide one application of Chebyshev's inequality in real life. (Linked to SLO 4.7)

Long-answer questions

11. Analyse the properties of expectation and explain their importance in probability modelling. (Linked to SLO 4.2)

12. Evaluate the role of variance and standard deviation in describing distributions, with examples. (Linked to SLO 4.4)

13. Discuss skewness and kurtosis, highlighting their applications in finance and data analysis. (Linked to SLO 4.3, 4.7)

14. Examine Markov's and Chebyshev's inequalities, explaining how they provide bounds and their applications in risk management. (Linked to SLO 4.5, 4.6, 4.7)

Answers to Concept Check Questions 4

Objective questions

1. B
2. B
3. B
4. B
5. B

Short-answer questions

6. Expectation is the long-run average value of a random variable.

7. Var(X)=E[(X-μ)2].

8. Skewness measures asymmetry of a distribution.

9. P(X≥a)≤E[X]a, for non-negative X.

10. In quality control, Chebyshev's inequality helps estimate probability of large deviations in product measurements.

Long-answer questions

11. *Basic response*: Expectation captures the average outcome of a random variable. *Value-added response*: Its linearity makes it essential in economics, insurance, and statistical modelling, simplifying complex calculations.

12. *Basic response*: Variance and standard deviation measure spread around the mean. *Value-added response*: They are crucial in finance (risk assessment), education (exam score variability), and manufacturing (quality control).

13. *Basic response*: Skewness measures asymmetry, kurtosis measures peakedness. *Value-added response*: In finance, skewness indicates bias in returns, while kurtosis identifies heavy tails, important for risk management.

14. *Basic response*: Markov's and Chebyshev's inequalities provide bounds on probabilities. *Value-added response*: They are applied in risk management, algorithm analysis, and statistics when exact distributions are unknown, offering guarantees and safety margins.

Answers to Pilot Questions 4

Part 4.1: 1-B, 2-B, 3-B, 4-A, 5-B **Part 4.2**: 1-B, 2-B, 3-B, 4-B, 5-A **Part 4.3**: 1-B, 2-A, 3-B, 4-B, 5-B

References and Attributions 4

- **Figures, Tables, and Boxes**:
  - Figure 4.1: Expectation as the centre of mass of a distribution (AI-generated, prompt: "Generate a diagram showing expectation as the balance point of a probability distribution").
  - Table 4.1: Properties of expectation (AI-generated, prompt: "Generate a table showing properties of expectation with formulas and examples").

- Figure 4.2: Variance and standard deviation as measures of spread (AI-generated, prompt: “Generate a diagram showing a bell curve with variance represented by spread around the mean”).
- Table 4.2: Moments of random variables (AI-generated, prompt: “Generate a table showing first four moments of random variables with formulas and interpretations”).
- Figure 4.3: Illustration of Chebyshev’s inequality (AI-generated, prompt: “Generate a diagram showing Chebyshev’s inequality with shaded region within 2 standard deviations of the mean”).
- Box 4.1, 4.3: AI-generated text boxes summarising applications of expectation and probability inequalities.

- **Search strings used**:
  - “expectation mean probability distribution diagram OER”
  - “properties of expectation probability OER”
  - “variance standard deviation bell curve diagram OER”
  - “moments of random variables table OER”
  - “Chebyshev inequality diagram probability OER”
  - “applications of Markov Chebyshev inequality OER”
- **Attribution**: Explanations and instructional content are based on standard open educational resources in probability theory. AI-generated illustrations were created using prompts as noted above.

Series 5 Outline

5.1 Joint distributions

- 5.1.1 Definition of joint distributions
- 5.1.2 Properties of joint distributions
- 5.1.3 Examples of joint distributions

5.2 Marginal distributions

- 5.2.1 Definition of marginal distributions

- 5.2.2 Relationship between joint and marginal distributions
- 5.2.3 Applications of marginal distributions

5.3 Limiting distributions

- 5.3.1 Definition of limiting distributions
- 5.3.2 Convergence in distribution
- 5.3.3 Applications of limiting distributions

5.4 Law of Large Numbers (LLN)

- 5.4.1 Statement of the LLN
- 5.4.2 Weak and strong forms of LLN
- 5.4.3 Applications of LLN

5.5 Central Limit Theorem (CLT)

- 5.5.1 Statement of the CLT
- 5.5.2 Importance of the CLT
- 5.5.3 Applications of the CLT

Introduction

In this Series, we move into **joint, marginal, and limiting distributions**, alongside two cornerstone results in probability: the **Law of Large Numbers (LLN)** and the **Central Limit Theorem (CLT)**. These concepts extend our understanding of random variables from individual behaviour to collective and asymptotic behaviour.

- **Joint distributions** describe the probability behaviour of two or more random variables together.
- **Marginal distributions** are derived from joint distributions, focusing on one variable at a time.
- **Limiting distributions** capture the behaviour of sequences of random variables as the number of observations grows.
- The **Law of Large Numbers** formalises the idea that averages converge to expected values as sample size increases.
- The **Central Limit Theorem** explains why normal distributions appear so frequently in practice, showing that sums (or averages) of many independent random variables tend toward a normal distribution.

Series Learning Outcomes

Upon completing **Series 5: Joint, Marginal and Limiting Distributions with Laws of Large Numbers and the Central Limit Theorem**, you should be able to:

- **SLO 5.1** define and explain joint distributions of random variables,
- **SLO 5.2** derive marginal distributions from joint distributions,
- **SLO 5.3** explain limiting distributions and convergence in distribution,
- **SLO 5.4** state and apply the Law of Large Numbers (LLN), distinguishing between weak and strong forms,
- **SLO 5.5** state and apply the Central Limit Theorem (CLT),
- **SLO 5.6** evaluate the importance of LLN and CLT in probability and statistics, and
- **SLO 5.7** apply joint, marginal, limiting distributions, LLN, and CLT to real-life problems in science, economics, and data analysis.

5.1 Joint Distributions

A **joint distribution** describes the probability behaviour of two or more random variables considered together. It captures how variables interact and whether they are dependent or independent. For example, the joint distribution of rainfall and crop yield shows how these two random variables are related.

Fill in the blank: A joint distribution describes the probability behaviour of __________ random variables considered together.

5.1.1 Definition of joint distributions

For **discrete random variables** Xand Y, the joint probability mass function (pmf) is:

$$P(X=x, Y=y)$$

For **continuous random variables**, the joint probability density function (pdf) is:

$$f(x,y), \text{with} \int\int f(x,y)\,dx\,dy=1$$

Fill in the blank: For continuous random variables, the joint distribution is described by a joint __________ density function.

Joint distribution of two continuous random variables.

A **joint probability density function (PDF)** describes the probability distribution of two or more continuous random variables.

For two variables, X and Y, the joint PDF is a surface defined by the function $f(x, y)$. This function must satisfy two key properties:

1. **Non-negativity:** $f(x, y) \geq 0$ for all (x, y).
2. **Total Volume = 1:** The volume under the entire surface must be exactly 1, representing the total probability of all possible outcomes. This is expressed mathematically by a double integral:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1$$

To find the probability that (X, Y) falls in a specific rectangular region defined by $[a, b] \times [c, d]$, you would calculate the volume under the surface over that specific area:

$$P(a \leq X \leq b, c \leq Y \leq d) = \int_{c}^{d} \int_{a}^{b} f(x, y)\, dx\, dy$$

For a visualization of such a surface, imagine a 3D plot where the x and y axes represent the values of the random variables, and the z axis represents the value of the joint PDF, $f(x, y)$, creating a shape like a mountain or a 3D bell curve.

Marginal PDFs

If you have the joint PDF $f(x, y)$, you can find the individual PDF for each variable (called the **marginal PDF**) by "integrating out" the other variable.

- The **marginal PDF of X**, $f_X(x)$, is found by integrating the joint PDF with respect to Y across all possible Y values:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy$$

- Similarly, the **marginal PDF of Y**, $f_Y(y)$, is found by integrating with respect to X:

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx$$

The diagram above illustrates this concept. The 3D surface is the joint PDF, $f(x,y)$. The two curves projected on the back and side planes represent the marginal PDFs, $f_X(x)$ and $f_Y(y)$, respectively. Visually, these marginal distributions can be seen as the "shadows" cast by the joint distribution surface.

Calculating Marginal Probabilities from Joint PDF

Let's look at an example to see how to calculate probabilities using a joint PDF. Suppose we have the joint PDF:
$f(x,y) = 4xy$ for $0 \le x \le 1, 0 \le y \le 1$, and $f(x,y) = 0$ otherwise.

Problem: Find $P(X \le 0.5 \text{ and } Y \le 0.75)$.

Solution: This region is defined by $0 \le x \le 0.5$ and $0 \le y \le 0.75$. We set up the double integral over these limits.

$$P(X \le 0.5 \text{ and } Y \le 0.75) = \int_0^{0.75} \int_0^{0.5} 4xy\,dx\,dy$$

1. **Integrate with respect to x:**

$$= \int_0^{0.75} \left[\frac{4x^2y}{2}\right]_{x=0}^{x=0.5} dy$$

$$= \int_0^{0.75} \left[2x^2y\right]_{x=0}^{x=0.5} dy$$

$$= \int_0^{0.75} \left(2(0.5)^2y - 2(0)^2y\right)\,dy$$

$$= \int_0^{0.75} \left(2(0.25)y\right)\,dy$$

$$= \int_0^{0.75} 0.5y \, dy$$

2. **Integrate with respect to y:**

$$= \left[\frac{0.5y^2}{2}\right]_{y=0}^{y=0.75}$$

$$= \left[0.25y^2\right]_{y=0}^{y=0.75}$$

$$= 0.25(0.75)^2 - 0.25(0)^2$$

$$= 0.25(0.5625)$$

$$= 0.140625$$

So, the probability that X 0.5 and Y 0.75 simultaneously is approximately 0.1406.

5.1.2 Properties of joint distributions

- **Non-negativity**: P(X=x,Y=y)≥0.
- **Normalization**: The sum (discrete) or integral (continuous) of all joint probabilities equals 1.
- **Marginalization**: Marginal distributions can be derived by summing or integrating over one variable.
- **Independence**: If P(X=x,Y=y)=P(X=x)P(Y=y), then Xand Yare independent.

Fill in the blank: If P(X=x,Y=y)=P(X=x)P(Y=y), then Xand Yare __________.

[Insert table here] Caption: Table 5.1: Properties of joint distributions

Property	**Formula**	**Explanation**
Non-negativity	P(X=x,Y=y)≥0	Probabilities are non-negative
Normalization	xyP(X=x,Y=y)=1	Total probability equals 1
Marginalization	P(X=x)=yP(X=x,Y=y)	Deriving marginal distributions
Independence	P(X=x,Y=y)=P(X=x)P(Y=y)	Variables are independent

5.1.3 Examples of joint distributions

- **Weather and crop yield**: Joint distribution of rainfall and crop yield.
- **Economics**: Joint distribution of income and expenditure.
- **Health**: Joint distribution of age and blood pressure.
- **Sports**: Joint distribution of goals scored by two teams.

Fill in the blank: The joint distribution of rainfall and crop yield is an example from __________.

Real-life examples of joint distributions

- Weather and agriculture
- Economics and business
- Health and medicine
- Sports and competition

Pilot questions 5.1

1. A joint distribution describes the probability behaviour of: A. One random variable B. Two or more random variables C. Only continuous random variables D. None of the above
2. For continuous random variables, the joint distribution is described by: A. Probability mass function B. Probability density function C. Cumulative distribution function D. None of the above
3. If $P(X=x,Y=y)=P(X=x)P(Y=y)$, then Xand Yare: A. Dependent B. Independent C. Identical D. None of the above
4. Marginal distributions can be derived from joint distributions by: A. Multiplication B. Summation or integration C. Subtraction D. None of the above
5. The joint distribution of rainfall and crop yield is an example from: A. Economics B. Weather and agriculture C. Sports D. None of the above

5.2 Marginal Distributions

A **marginal distribution** describes the probability behaviour of a single random variable within a joint distribution. It is obtained by summing (for discrete variables) or integrating (for continuous variables) over the other variables. Marginal distributions allow us to focus on one variable at a time while still acknowledging its relationship with others.

Fill in the blank: A marginal distribution describes the probability behaviour of a __________ random variable within a joint distribution.

5.2.1 Definition of marginal distributions

For **discrete random variables** Xand Y:

P(X=x)=yP(X=x,Y=y)

For **continuous random variables**:

fX(x)=-f(x,y) dy

Fill in the blank: For continuous random variables, the marginal distribution is obtained by __________ the joint density function over the other variable.

5.2.2 Relationship between joint and marginal distributions

- The **joint distribution** contains complete information about two or more variables.
- The **marginal distribution** is derived from the joint distribution by focusing on one variable.
- Marginals are useful when we want to study one variable independently, but they do not capture dependence between variables.

Fill in the blank: Marginal distributions are derived from __________ distributions by focusing on one variable.

Marginal distribution derived from a joint distribution.

The Marginal PDF: A Projection

If you have the joint PDF $f(x, y)$, you can find the individual PDF for one of the variables (called the **marginal PDF**) by "integrating out" the other variable. This process is visually equivalent to projecting the volume of the 3D surface onto one of the 2D planes.

- **Projecting onto the X-Axis:** To find the marginal PDF of X, $f_X(x)$, you integrate the joint PDF with respect to Y across all possible Y values. Visually, you can imagine "squashing" the 3D surface onto the X-Z plane.

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy$$

In the diagram above, the 3D surface represents the joint PDF, $f(x, y)$. The colored curve on the back wall (the X-Z plane) is the **projection**, which is the **marginal PDF** $f_X(x)$. Every point on this marginal curve represents the total probability of X taking a specific value x, summed across all possible values of Y.

5.2.3 Applications of marginal distributions

- **Economics**: Studying income distribution without considering expenditure.
- **Health**: Analysing blood pressure distribution without considering age.

- **Sports**: Looking at goals scored by one team without considering the opponent.
- **Weather**: Examining rainfall distribution without considering temperature.

Fill in the blank: In economics, marginal distributions can be used to study __________ distribution without considering expenditure.

Real-life applications of marginal distributions

- Economics: income distribution
- Health: blood pressure distribution
- Sports: team goals
- Weather: rainfall distribution

Pilot questions 5.2

1. A marginal distribution describes the probability behaviour of: A. Two random variables together B. A single random variable within a joint distribution C. Only continuous random variables D. None of the above
2. For discrete random variables, the marginal distribution is obtained by: A. Multiplication B. Summation C. Integration D. None of the above
3. For continuous random variables, the marginal distribution is obtained by: A. Summation B. Integration C. Subtraction D. None of the above
4. Marginal distributions are derived from: A. Joint distributions B. Limiting distributions C. Independent distributions D. None of the above
5. In economics, marginal distributions can be used to study: A. Expenditure distribution B. Income distribution C. Both A and B D. None of the above

5.3 Limiting Distributions

A **limiting distribution** describes the behaviour of a sequence of random variables as the number of observations grows large. It helps us understand the long-term or asymptotic behaviour of random processes. Limiting distributions are central to probability theory because they connect finite samples to infinite behaviour.

Fill in the blank: A limiting distribution describes the behaviour of a sequence of random variables as the number of __________ grows large.

5.3.1 Definition of limiting distributions

If a sequence of random variables Xnconverges in distribution to a random variable X, then Xis said to have the **limiting distribution** of Xn.

Formally:

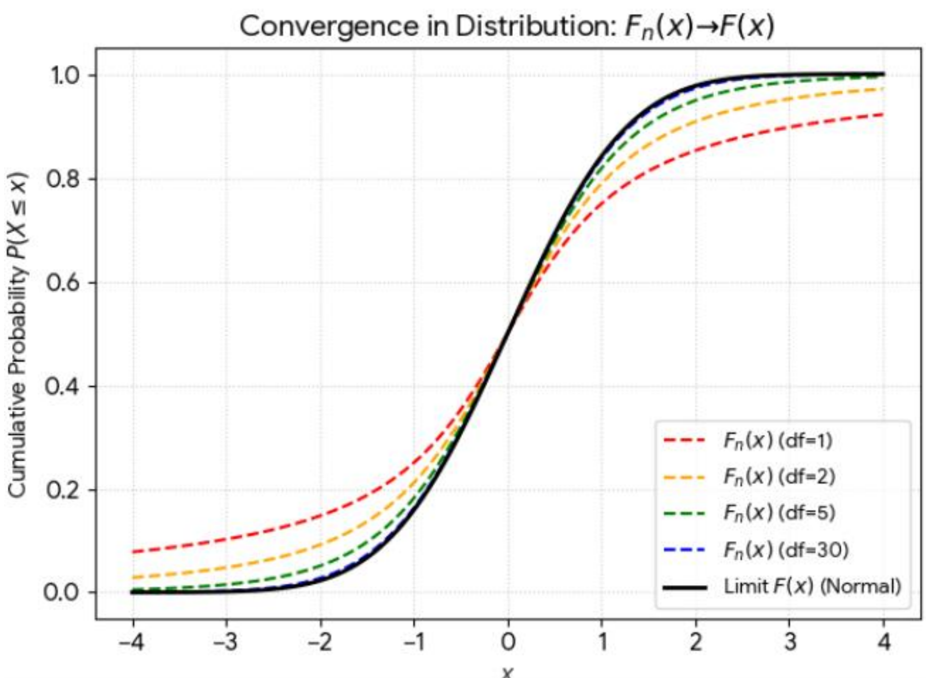
$$X_n \xrightarrow{d} X \text{ if } \lim_{n\to\infty} P(X_n \le x) = P(X \le x), x$$

Fill in the blank: If $X_n \xrightarrow{d} X$, then X is called the __________ distribution of the sequence X_n.

5.3.2 Convergence in distribution

- **Convergence in distribution** means that the cumulative distribution functions (CDFs) of X_n approach the CDF of X.
- It does not require the random variables themselves to converge pointwise.
- This concept is crucial in proving the Central Limit Theorem.

Fill in the blank: Convergence in distribution means that the __________ of X_n approach that of X.



Convergence in distribution illustrated with cumulative distribution functions.

5.3.3 Applications of limiting distributions

- **Statistics**: Sampling distributions converge to normal distributions under the CLT.
- **Economics**: Long-run behaviour of averages in economic data.
- **Physics**: Limiting distributions describe equilibrium states.
- **Computer science**: Algorithm performance analysis in large inputs.

Fill in the blank: In statistics, limiting distributions are central to the __________ Limit Theorem.

Real-life applications of limiting distributions

- Statistics: sampling distributions
- Economics: long-run averages
- Physics: equilibrium states
- Computer science: algorithm analysis

Pilot questions 5.3

1. A limiting distribution describes the behaviour of a sequence of random variables as: A. The number of trials decreases B. The number of observations grows large C. The mean becomes zero D. None of the above
2. If XndX, then Xis called the: A. Marginal distribution B. Limiting distribution C. Joint distribution D. None of the above
3. Convergence in distribution means that: A. Random variables converge pointwise B. CDFs of Xnapproach that of XC. Means of Xnapproach zero D. None of the above
4. Limiting distributions are central to proving: A. Law of Large Numbers B. Central Limit Theorem C. Markov's inequality D. None of the above
5. In physics, limiting distributions describe: A. Equilibrium states B. Random shocks C. Marginal probabilities D. None of the above

5.4 Law of Large Numbers (LLN)

The **Law of Large Numbers (LLN)** formalises the idea that as the number of trials or observations increases, the sample average converges to the expected value. This principle explains why averages become more stable with larger samples.

Fill in the blank: The Law of Large Numbers states that the sample __________ converges to the expected value as the number of observations increases.

5.4.1 Statement of the LLN

- **Weak Law of Large Numbers (WLLN)**: The sample mean converges to the expected value in probability as the sample size grows.
- **Strong Law of Large Numbers (SLLN)**: The sample mean converges to the expected value almost surely (with probability 1).

Formally, if X1,X2,...,Xnare i.i.d. random variables with mean :

X¯n=1ni=1nXin→

Fill in the blank: The strong law of large numbers states that the sample mean converges to the expected value __________ surely.

5.4.2 Weak and strong forms of LLN

- **Weak form**: Convergence in probability.

- **Strong form**: Convergence almost surely.
- Both forms guarantee that averages stabilise as sample size increases, but the strong law is a stronger statement.

Fill in the blank: The weak law of large numbers states that the sample mean converges to the expected value in __________.

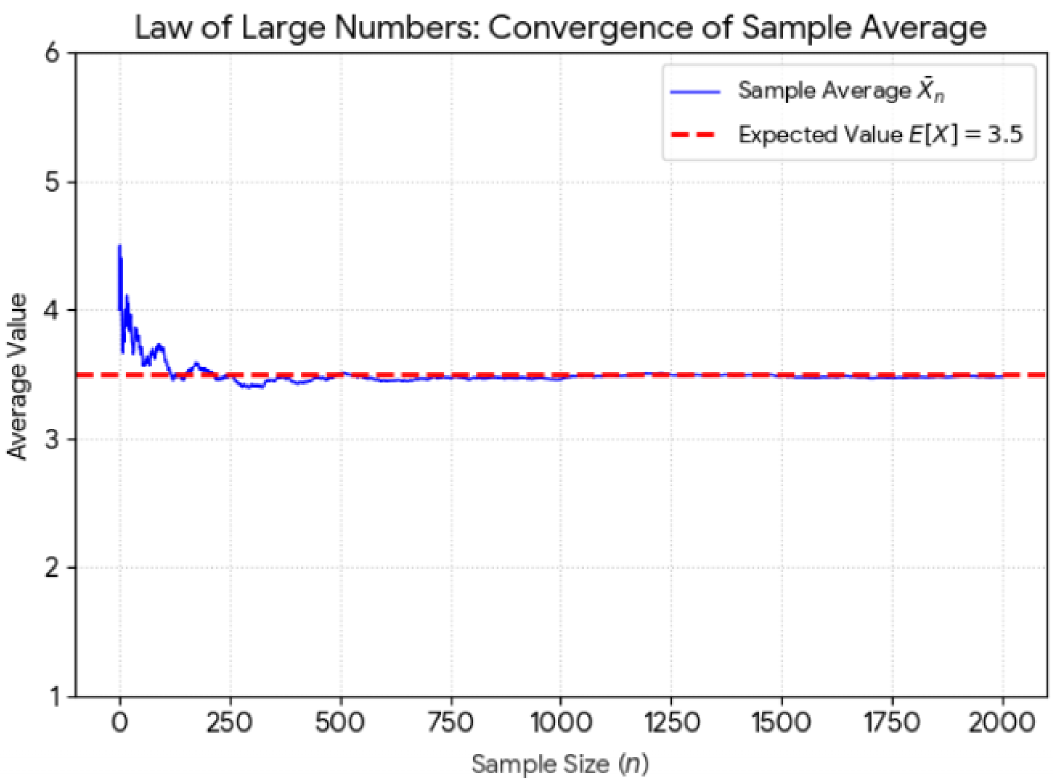


Illustration of the Law of Large Numbers.

5.4.3 Applications of LLN

- **Insurance**: Premiums are based on average claims stabilising over large numbers of policyholders.
- **Gambling**: Explains why casinos profit in the long run.
- **Economics**: Large samples provide reliable estimates of population averages.
- **Science**: Experimental averages converge to true values with repeated trials.

Fill in the blank: In gambling, the Law of Large Numbers explains why __________ profit in the long run.

Real-life applications of the Law of Large Numbers

- Insurance: average claims

- Gambling: casino profits
- Economics: population averages
- Science: experimental averages

Pilot questions 5.4

1. The Law of Large Numbers states that the sample mean converges to: A. The maximum value B. The expected value C. The minimum value D. None of the above
2. The weak law of large numbers states that the sample mean converges in: A. Probability B. Almost surely C. Distribution D. None of the above
3. The strong law of large numbers states that the sample mean converges: A. In probability B. Almost surely C. In distribution D. None of the above
4. In gambling, the Law of Large Numbers explains why: A. Players profit in the long run B. Casinos profit in the long run C. Both A and B D. None of the above
5. In insurance, the Law of Large Numbers ensures that: A. Premiums are unpredictable B. Average claims stabilise C. Losses always occur D. None of the above

5.5 Central Limit Theorem (CLT)

The **Central Limit Theorem (CLT)** is one of the most important results in probability and statistics. It explains why the normal distribution appears so frequently in practice. The theorem states that the distribution of the sum (or average) of a large number of independent, identically distributed random variables tends toward a normal distribution, regardless of the original distribution.

Fill in the blank: The Central Limit Theorem states that the distribution of the __________ of many independent random variables tends toward a normal distribution.

5.5.1 Statement of the CLT

Formally, if X1,X2,...,Xnare i.i.d. random variables with mean and variance 2, then the standardised sum:

Zn=i=1nXi-nμn

converges in distribution to a standard normal random variable as n→.

Fill in the blank: The CLT states that the standardised sum of i.i.d. random variables converges to a __________ normal distribution.

5.5.2 Importance of the CLT

- Explains why normal distribution is common in statistics.
- Provides foundation for hypothesis testing and confidence intervals.
- Allows approximation of distributions when exact forms are unknown.

- Connects probability theory with statistical inference.

Fill in the blank: The CLT provides the foundation for hypothesis testing and __________ intervals.

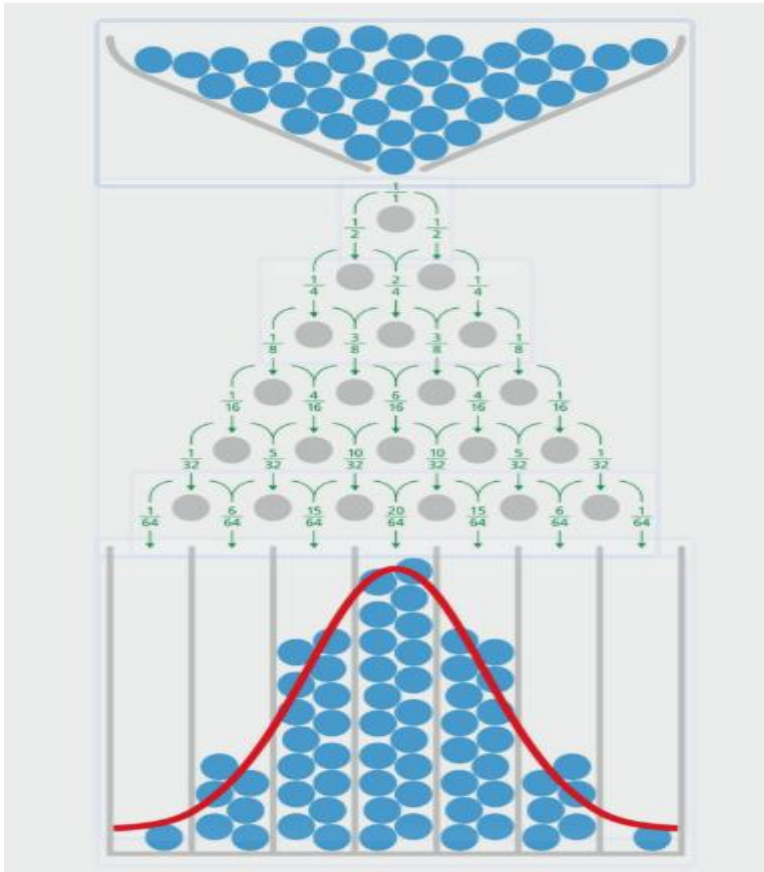


Illustration of the Central Limit Theorem.

5.5.3 Applications of the CLT

- **Economics**: Average income distributions approximate normality.
- **Science**: Measurement errors aggregate to normal distributions.
- **Manufacturing**: Quality control relies on normal approximations.
- **Data science**: Sampling distributions of estimators are approximately normal.

Fill in the blank: In science, measurement __________ aggregate to normal distributions under the CLT.

Real-life applications of the Central Limit Theorem

- Economics: average incomes
- Science: measurement errors
- Manufacturing: quality control

- Data science: sampling distributions

Pilot questions 5.5

1. The Central Limit Theorem states that the distribution of the sum of many independent random variables tends toward: A. Uniform distribution B. Normal distribution C. Binomial distribution D. None of the above
2. The CLT applies to random variables that are: A. Independent and identically distributed B. Dependent and skewed C. Always discrete D. None of the above
3. The CLT states that the standardised sum of i.i.d. random variables converges to: A. Standard normal distribution B. Uniform distribution C. Poisson distribution D. None of the above
4. The CLT provides the foundation for: A. Hypothesis testing and confidence intervals B. Counting outcomes C. Marginal distributions D. None of the above
5. In science, measurement errors aggregate to: A. Uniform distribution B. Normal distribution C. Binomial distribution D. None of the above

Series Summary

This Series introduced **joint, marginal, and limiting distributions**, along with the **Law of Large Numbers (LLN)** and the **Central Limit Theorem (CLT)**. We began with joint distributions, which describe the probability behaviour of two or more random variables together, and marginal distributions, which focus on one variable derived from a joint distribution. We then studied limiting distributions, which capture the long-term behaviour of sequences of random variables. Next, we examined the LLN, which formalises the convergence of sample averages to expected values, distinguishing between weak and strong forms. Finally, we explored the CLT, which explains why sums or averages of many independent random variables tend toward a normal distribution, forming the foundation of modern statistical inference.

By completing this Series, you should now be able to define and apply joint, marginal, and limiting distributions, state and apply the LLN and CLT, and evaluate their importance in probability, statistics, and real-life applications.

Glossary of Terms

- **Joint distribution**: Probability distribution of two or more random variables considered together.
- **Marginal distribution**: Distribution of a single random variable derived from a joint distribution.
- **Limiting distribution**: Distribution describing the asymptotic behaviour of a sequence of random variables.
- **Convergence in distribution**: When the cumulative distribution functions of a sequence of random variables approach that of a limiting random variable.

- **Law of Large Numbers (LLN)**: Principle that sample averages converge to expected values as sample size increases.
- **Weak LLN**: Convergence in probability.
- **Strong LLN**: Convergence almost surely.
- **Central Limit Theorem (CLT)**: Theorem stating that sums or averages of many independent, identically distributed random variables tend toward a normal distribution.

Concept Check Questions 5

Objective questions

1. A joint distribution describes the probability behaviour of: A. One random variable B. Two or more random variables C. Only continuous random variables D. None of the above (Linked to SLO 5.1)
2. A marginal distribution is obtained by: A. Multiplication B. Summation or integration C. Subtraction D. None of the above (Linked to SLO 5.2)
3. Convergence in distribution means that: A. Random variables converge pointwise B. CDFs of Xnapproach that of XC. Means of Xnapproach zero D. None of the above (Linked to SLO 5.3)
4. The Law of Large Numbers states that the sample mean converges to: A. Maximum value B. Expected value C. Minimum value D. None of the above (Linked to SLO 5.4)
5. The Central Limit Theorem states that sums of many independent random variables tend toward: A. Uniform distribution B. Normal distribution C. Binomial distribution D. None of the above (Linked to SLO 5.5)

Short-answer questions

6. Define a joint distribution. (Linked to SLO 5.1)

7. State the formula for a marginal distribution of a continuous random variable. (Linked to SLO 5.2)

8. What does convergence in distribution mean? (Linked to SLO 5.3)

9. Distinguish between weak and strong forms of the LLN. (Linked to SLO 5.4)

10. State the Central Limit Theorem. (Linked to SLO 5.5)

Long-answer questions

11. Analyse the relationship between joint and marginal distributions with examples. (Linked to SLO 5.1, 5.2)

12. Evaluate the role of limiting distributions in probability theory and statistics. (Linked to SLO 5.3)

13. Discuss the Law of Large Numbers, explaining its weak and strong forms and applications in economics and science. (Linked to SLO 5.4, 5.7)

14. Examine the Central Limit Theorem, highlighting its importance in hypothesis testing and real-life applications. (Linked to SLO 5.5, 5.6, 5.7)

Answers to Concept Check Questions 5

Objective questions

1. B
2. B
3. B
4. B
5. B

Short-answer questions

6. A joint distribution describes the probability behaviour of two or more random variables together.

7. fX(x)=-f(x,y) dy.

8. Convergence in distribution means the CDFs of a sequence of random variables approach the CDF of a limiting random variable.

9. Weak LLN: convergence in probability; Strong LLN: convergence almost surely.

10. The CLT states that the standardised sum of i.i.d. random variables converges to a normal distribution as sample size grows.

Long-answer questions

11. *Basic response*: Joint distributions describe multiple variables, while marginal distributions focus on one variable. *Value-added response*: For example, the joint distribution of rainfall and crop yield can be marginalised to study rainfall alone.

12. *Basic response*: Limiting distributions describe asymptotic behaviour of random variables. *Value-added response*: They are central in proving the CLT and in analysing long-run averages in economics and physics.

13. *Basic response*: LLN states that sample averages converge to expected values. *Value-added response*: Weak LLN ensures convergence in probability, strong LLN ensures convergence almost surely. Applications include insurance premiums and experimental averages.

14. *Basic response*: CLT states that sums of many independent random variables tend toward a normal distribution. *Value-added response*: It underpins hypothesis testing, confidence intervals, and explains why normal distributions appear in economics, science, and data science.

Answers to Pilot Questions 5

Part 5.1: 1-B, 2-B, 3-B, 4-B, 5-B **Part 5.2**: 1-B, 2-B, 3-B, 4-A, 5-B **Part 5.3**: 1-B, 2-B, 3-B, 4-B, 5-A **Part 5.4**: 1-B, 2-A, 3-B, 4-B, 5-B **Part 5.5**: 1-B, 2-A, 3-A, 4-A, 5-B

References and Attributions 5

- **Figures, Tables, and Boxes**:
    - Figure 5.1: Joint distribution of two continuous random variables (AI-generated, prompt: “Generate a 3D surface diagram showing a joint probability density function of two continuous random variables”).
    - Table 5.1: Properties of joint distributions (AI-generated, prompt: “Generate a table showing properties of joint distributions with formulas and explanations”).
    - Figure 5.2: Marginal distribution derived from a joint distribution (AI-generated, prompt: “Generate a diagram showing a joint distribution surface with projection onto one axis representing the marginal distribution”).
    - Box 5.1, 5.2, 5.3: AI-generated text boxes summarising real-life examples of joint, marginal, and limiting distributions.
    - Figure 5.3: Convergence in distribution illustrated with cumulative distribution functions (AI-generated, prompt: “Generate a graph showing sequence of cumulative distribution functions converging to a limiting distribution”).
    - Figure 5.4: Illustration of the Law of Large Numbers (AI-generated, prompt: “Generate a graph showing sample averages converging to the expected value as sample size increases”).
    - Box 5.4: AI-generated text box summarising real-life applications of LLN.
    - Figure 5.5: Illustration of the Central Limit Theorem (AI-generated, prompt: “Generate a diagram showing different distributions converging to a normal curve as sample size increases”).
    - Box 5.5: AI-generated text box summarising real-life applications of CLT.
- **Search strings used**:
    - “joint probability distribution diagram OER”
    - “properties of joint probability distribution OER”
    - “marginal distribution diagram probability OER”
    - “marginal distribution real life examples OER”
    - “convergence in distribution CDF diagram OER”
    - “limiting distribution real life examples OER”
    - “law of large numbers illustration OER”
    - “law of large numbers real life applications OER”
    - “central limit theorem illustration OER”
    - “central limit theorem real life applications OER”