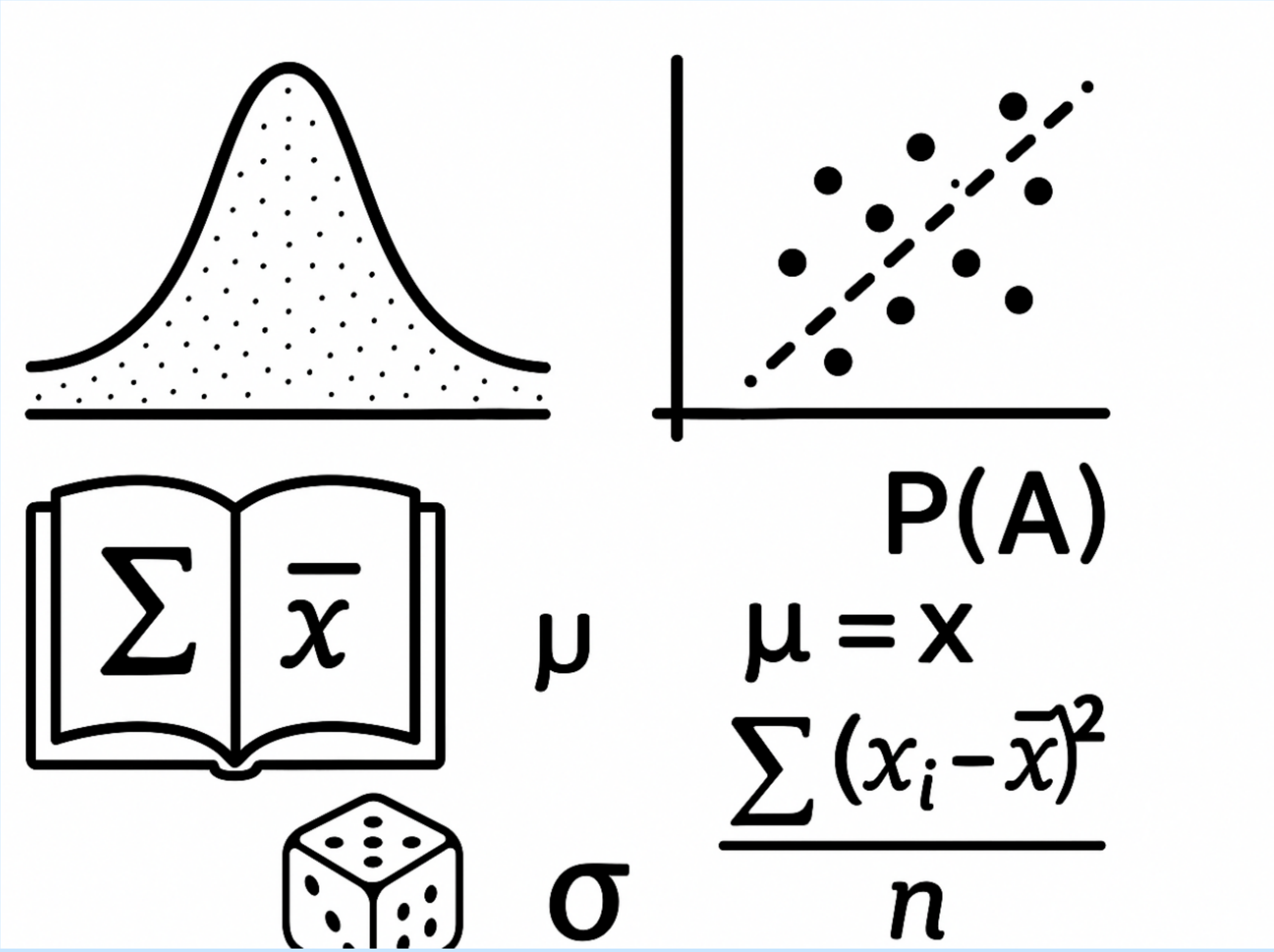
GATE – Computer Science and Engineering (CSE)

Probability and Statistics



GateXAIML

2025

Contents

About the Book

The **GATE Computer Science and Engineering (CSE) exam** serves as a national-level gateway to higher studies, research, and employment in top institutions and organizations. It evaluates a candidate's understanding of core Computer Science subjects, including Calculus, Matrices and Linear Algebra, Probability and Statistics, Discrete Mathematics, Algorithms, Data Structures, Theory of Computation, Databases, Operating Systems, and Computer Networks.

This book is **designed for aspirants of the GATE CSE exam**, focusing on **Probability and Statistics**. It systematically covers theory, solved examples, and practice problems **aligned with the official syllabus**, helping learners build strong conceptual foundations and problem-solving skills.

Selected solutions and topic-wise lectures will be explained on my YouTube channel (@GATEXAiml), providing a **complete resource for GATE CSE preparation**.

Dedicated to all my Gurus and Students.
"Knowledge grows only when shared — and it must remain free, for that is how it thrives."

Probability and Statisitcs - Syllabus

Random variables, Uniform, normal, exponential, Poisson and binomialdistributions. Mean, median, mode and standard deviation. Conditional probability and Bayes theorem.

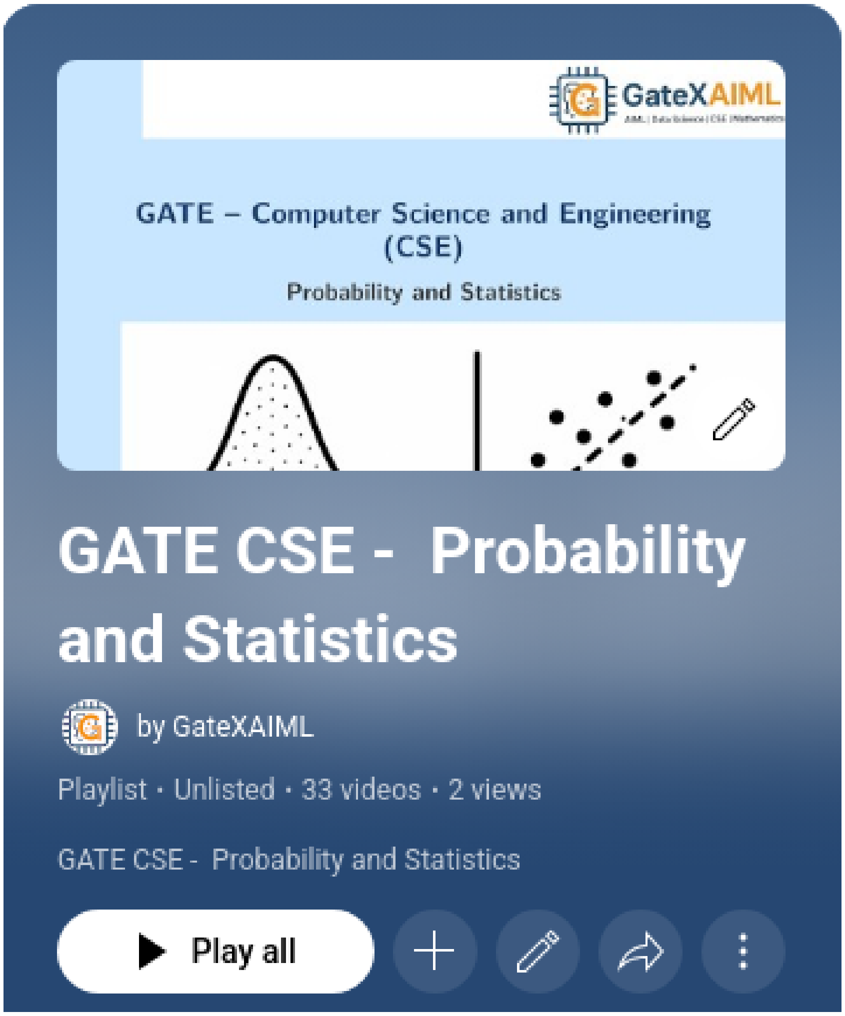
STOP!

✦ Attention! ✦

Some examples solved in video lectures are different from those given in this book.

The procedure to solve problems and examples is well explained in the video lectures, and it is highly recommended to go through the video lectures for complete understanding.

Official Video Playlist



Watch on YouTube

Chapter 1

Counting

NOTE

The counting (Permutations and Combinations) is included in the GATE CSE syllabus under **Discrete Mathematics**; however, in this book, we cover this topic under **Probability and Statistics**.

What is Combinatorics?

Combinatorics is a branch of mathematics focused on *counting, arranging, and analyzing discrete structures*. It has wide applications in computer science, probability theory, and logic.

Key Areas in Combinatorics

1. **Counting (Enumerative Combinatorics)**
   Concerned with determining *how many* ways certain configurations can occur.
   - Example: How many 4-letter passwords can be formed using letters A–Z?
2. **Permutations and Combinations**
   - **Permutation**: The order *matters*.
     - Example: How many ways can 3 books be arranged on a shelf?
       Solution: $P(n, r) = \frac{n!}{(n-r)!}$
   - **Combination**: The order *does not matter*.
     - Example: How many ways can you choose 3 students from a group of 10?
       Solution: $C(n, r) = \binom{n}{r} = \frac{n!}{r!(n-r)!}$

3. **Pigeonhole Principle**
   If more than n items are put into n containers, at least one container must contain more than one item.

4. **Inclusion-Exclusion Principle**
   Used to count the total number of elements in the union of overlapping sets:

$$|A \cup B| = |A| + |B| - |A \cap B|$$

   For three sets:

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |B \cap C| - |C \cap A| + |A \cap B \cap C|$$

5. **Graph Theory**
   The study of graphs, which are mathematical structures used to model pairwise relations between objects.

6. **Partition Theory**
   The number of ways of writing a positive integer as a sum of other positive integers.

   - Example: The number 5 can be written as $5 = 2 + 2 + 1$, among other ways.

7. **Generating Functions**
   Algebraic tools that encode sequences of numbers. Useful for solving recurrence relations and counting problems.

Applications of Combinatorics

- Computer Algorithms
- Cryptography
- Game Theory
- Probability Theory
- Data Structures
- Bioinformatics (e.g., DNA sequencing)

Rule Identification Guide in Combinatorics

1. Sum Rule (Addition)

Use when: You have multiple mutually exclusive options, and you need to pick one.

Keywords to watch for:

- "or"
- "either this or that"

- "one of the following"

Formula: If task A can be done in m ways and task B in n ways (and both can't happen together), then:

$$\text{Total ways} = m + n$$

2. Product Rule (Multiplication)

Use when: You need to do multiple independent tasks in sequence.

Keywords to watch for:

- "and"
- "followed by"
- "for each choice of . . . "

Formula: If task A can be done in m ways and task B in n ways, then:

$$\text{Total ways} = m \times n$$

3. Subtraction Rule (Complement)

Use when: It's easier to count the total number of cases and subtract unwanted/restricted ones.

Keywords to watch for:

- "not containing"
- "at least one"
- "excluding"

Formula:

$$\text{Favorable cases} = \text{Total cases} - \text{Unfavorable cases}$$

4. Division Rule (Correcting Overcounting)

Use when: Your method overcounts due to symmetry or repetition of elements.

Keywords to watch for:

- "arrangements of identical items"
- "rotations/symmetry"
- "repeated elements"

Formula:

$$\text{Correct count} = \frac{\text{Total permutations}}{\text{Repetition factor}}$$

Example:

$$\frac{n!}{p_1! \cdot p_2! \cdot \dots \cdot p_k!} \quad \text{(for repeated elements)}$$

Quick Decision Table

When to Use Which Rule?

Scenario	Rule to Use
You must choose between two options only one of which can happen	Sum Rule
You are doing two or more tasks in sequence (both must happen)	Product Rule
You are asked "at least", "at most", or "not containing something"	Subtraction Rule
You are arranging items with some repeated/symmetric cases	Division Rule

Sum Rule

If a task can be done in m ways or n ways (but not both), then the total number of ways to do the task is $m + n$.

Example 1: Sum Rule

You can take a bus (5 options) or a train (3 options) to reach a city.
Solution: Since these are mutually exclusive choices:

$$\text{Total ways} = 5 + 3 = \boxed{8}$$

Product Rule

If a task consists of two steps, where the first step can be done in m ways and the second in n ways, then the total number of ways of doing the task is $m \times n$.

Example 2: Product Rule

You want to choose a shirt (4 choices) and a pant (5 choices).
Solution:

$$\text{Total outfits} = 4 \times 5 = \boxed{20}$$

Subtraction Rule (Complement Rule)

To count the number of outcomes that satisfy a condition, subtract from the total the number of outcomes that do not satisfy it.

Example 3: Subtraction Rule

How many 2-digit numbers (10–99) *do not* contain the digit 7?
Solution: Total 2-digit numbers $= 90$.
Numbers that contain digit 7 = Numbers with 7 in tens place + Numbers with 7 in ones place − Double counted (77).
$= 10 + 9 - 1 = 18$.
Therefore, numbers without 7 $= 90 - 18 = 72$.

Division Rule

If a task can be done in n ways and there are k indistinguishable repetitions for each outcome, then the actual number of distinct outcomes is $\frac{n}{k}$.

Example 4: Division Rule

In a class of 4 students A, B, C, D, how many distinct 2-person teams can be formed? (Here, the team $\{A, B\}$ is the same as $\{B, A\}$.)
Solution: First count *ordered* selections: $4 \times 3 = 12$.
Each unordered team is counted twice (A, B and B, A). By the Division Rule, divide by 2:

$$\frac{12}{2} = \boxed{6}$$

So there are $\boxed{6}$ distinct 2-person teams.

1.1 Permutations

Permutations refer to the *arrangement of objects in a specific order*. It is used when the **order matters**.

If we are given n distinct objects and we want to arrange r of them in a sequence, the number of such arrangements (permutations) is given by:

$$P(n, r) = \frac{n!}{(n - r)!}$$

How to Identify Permutation Problems

Tip

A problem involves permutations if:

- The order or arrangement matters.
- You are asked to "arrange," "order," "line up," or form "different sequences."
- There are restrictions about positioning (like certain people must be together or not together).

Examples of Permutation Problems

Example 1: Simple Permutation

Example 1

How many different ways can 4 students be seated in 4 chairs?
Solution: Since all 4 students are seated, the number of ways $= P(4,4) = 4! = 24$

Example 2: Permutation of a Subset

Example 2

In how many ways can 3 students be selected and arranged from a group of 6?
Solution: $P(6,3) = \frac{6!}{(6-3)!} = \frac{720}{6} = 120$

Example 3: Digits in a Number

Example 3

How many 3-digit numbers can be formed using the digits 1, 2, 3, 4, 5 with no repetition?
Solution: Choose and arrange 3 digits from 5: $P(5,3) = \frac{5!}{2!} = 60$

Example 4: Arrangement with Restrictions

Example 4

How many ways can 4 books be arranged on a shelf if one specific book must always be at the end?
Solution: Fix the specific book at the end. Now arrange the remaining 3 books: $3! = 6$

Example 5: Circular Permutations

Example 5

In how many ways can 5 people be seated around a circular table?
Solution: Circular permutations: $(n-1)! = 4! = 24$

Example 6: Permutations with Repeated Items

Example 6

How many permutations can be formed from the letters of the word `LEVEL`?
Solution: Total letters = 5; Repeated letters: L (2), E (2)

$$\text{Permutations} = \frac{5!}{2! \cdot 2!} = \frac{120}{4} = 30$$

Example 7: At Least Condition

Example 7

From 7 people, how many ways can a president, vice president, and secretary be chosen?
Solution: These are distinct positions; order matters: $P(7,3) = \frac{7!}{4!} = 210$

Example 8: Arrangement with Items Together

Example 8

In how many ways can the letters of the word `SCHOOL` be arranged such that `S` and `C` are together?
Solution: Treat SC as one unit. So we have: [SC], H, O, O, L $\rightarrow$ 5 letters
But O is repeated.

$$\text{Arrangements} = \frac{5!}{2!} = \frac{120}{2} = 60$$

Also SC and CS both allowed $\rightarrow$ multiply by 2: $60 \times 2 = 120$

Example 9: Number with Constraints

Example 9

How many 4-digit numbers can be formed using the digits 1 to 6, such that the number starts with an odd digit?
Solution: Odd digits: 1, 3, 5 $\rightarrow$ 3 choices for first digit
Remaining 3 digits from 5 unused digits: $P(5,3) = 60$
Total $= 3 \times 60 = 180$

Example 10: All Letters Used with Restrictions

Example 10

How many ways can the letters of the word `APPLE` be arranged if P's must be together?
Solution: Treat PP as a block → Remaining: [PP], A, L, E → 4 elements

$$\text{Arrangements} = 4! = 24$$

Example 11: Repeated Digits Constraint

Example 11

How many 5-digit numbers can be formed using digits 1–5 such that no digit repeats and the number is even?
Solution: For a number to be even, it must end in 2 or 4.
Case 1: Last digit is 2 → Remaining 4 digits from {1,3,4,5}: $4!$
Case 2: Last digit is 4 → Remaining digits {1,2,3,5}: $4!$

$$\text{Total} = 2 \times 4! = 2 \times 24 = \boxed{48}$$

Example 12: Vowels Together

Example 12

How many ways can the letters of the word `ENGINEER` be arranged so that all vowels are together?
Solution: Vowels = E, I, E, E → 4 vowels (E repeated 3 times) Consonants = N, G, N, R → 4 consonants (N repeated 2 times)
Treat vowels as one block → total blocks = 5

$$\text{Ways to arrange blocks} = \frac{5!}{2!} = 60$$

$$\text{Vowel arrangements} = \frac{4!}{3!} = 4$$

$$\text{Total} = 60 \times 4 = \boxed{240}$$

Example 13: At Least One Condition

Example 13

How many 4-digit numbers can be formed using digits 0-9 such that no digit repeats and at least one digit is 7?
Solution: Total 4-digit numbers with distinct digits:
- Thousands digit: 1–9 → 9 options - Remaining 3 digits from remaining 9 digits → $P(9, 3)$

$$\text{Total} = 9 \times P(9,3) = 9 \times 504 = 4536$$

Now, subtract those without 7:
- Choose 4 digits from $\{0–9\}$ excluding 7 $\to$ 9 digits - Thousands place $\neq 0 \implies$ valid starting digits $=$ 8 options (excluding 0 and 7) - Remaining 3 digits: choose from remaining 8

$$\text{Without 7} = 8 \times P(8,3) = 8 \times 336 = 2688$$

$$\text{With at least one 7} = 4536 - 2688 = \boxed{1848}$$

Example 14: Positions Fixed

Example 14

In how many ways can 6 people be arranged in a row if 2 particular people (A and B) must always occupy the first and last positions (in any order)?
Solution: A and B can be in (1st,6th) or (6th,1st) $\to$ 2 ways Remaining 4 people can be arranged in $4!$

$$\text{Total} = 2 \times 4! = 2 \times 24 = \boxed{48}$$

Example 15: Word Arrangement with Restrictions

Example 15

How many arrangements of the word `ASSASSIN` are there such that no two S's are adjacent?
Solution: Total letters $=$ 8 S appears 4 times.
First, arrange the non-S letters: A, A, I, N $\to$ 4 letters Arrangements $= \frac{4!}{2!} = 12$
Place 4 S's in the 5 available gaps between/around non-S's: _ A _ A _ I _ N _ $\to$ 5 positions
Choose 4 positions out of 5 for S's: $\binom{5}{4} = 5$

$$\text{Total} = 12 \times 5 = \boxed{60}$$

Example 16: Conditional Digit Use

Example 16

How many 3-digit numbers can be formed using digits 2, 3, 4, 5, 6, 7 such that 5 is always used and digits are not repeated?
Solution: Total 3-digit numbers using digits from set of 6.
We fix 5 in one position. Then choose 2 more from remaining 5 digits.
Case 1: 5 at hundreds place $\to$ choose 2 digits from 5 $\to P(5,2) = 20$
Case 2: 5 at tens place $\to$ hundreds digit $\neq$ 0, 5 $\to$ choose from 5 digits (excluding 5), then one more from remaining 4 $\to 5 \times 4 = 20$

Case 3: 5 at units place $\rightarrow$ similar $\rightarrow 5 \times 4 = 20$

$$\text{Total} = 20 + 20 + 20 = \boxed{60}$$

Example 17: Circular Permutation with a Gap Constraint

Example 17

In how many distinct ways can 8 different people be seated around a circular table if persons A and B have *exactly two people* sitting between them?
Solution: Fix A's position to remove rotational symmetry. Around the circle, B must be placed so that exactly two seats lie between A and B. There are 2 such positions for B (clockwise or counterclockwise).
Arrange the remaining 6 people in the remaining seats: $6!$ ways.

$$\text{Total} = 2 \times 6! = 2 \times 720 = \boxed{1440}.$$

Example 18: Permutations with Restricted Positions

Example 18

In how many ways can 6 people be arranged in a row if person A must be in one of the two end positions?
Solution: First, choose the position for A: There are 2 choices (left end or right end).
The remaining 5 people can be arranged in any order in the remaining 5 seats: $5! = 120$ ways.

$$\text{Total arrangements} = 2 \times 120 = \boxed{240}.$$

Example 19: Permutation with Minimum Gap

Example 19

How many ways can the letters of `EXAMPLE` be arranged such that no two vowels are adjacent?
Solution: Vowels = E, A, E $\rightarrow$ total = 3 (E repeated twice)
Consonants = X, M, P, L $\rightarrow$ 4 letters
Arrange consonants: $4! = 24$
Now place 3 vowels in gaps between consonants: _ C _ C _ C _ C _ $\rightarrow$ 5 gaps
Choose 3 of 5 gaps: $\binom{5}{3} = 10$
Vowel arrangements $= \frac{3!}{2!} = 3$

$$\text{Total} = 24 \times 10 \times 3 = \boxed{720}$$

Example 20: Arrangements with Fixed Distance

Example 20

How many ways can 5 boys and 3 girls be seated in a row such that no two girls sit together?
Solution: First, place 5 boys: $5! = 120$
Gaps between boys = 6 positions: _ B _ B _ B _ B _ B _
Choose 3 gaps from 6 to place girls: $\binom{6}{3} = 20$
Arrange girls in those positions: $3! = 6$

$$\text{Total} = 120 \times 20 \times 6 = \boxed{14400}$$

Example 21: Permutations with Banned Positions

Example 21

In how many ways can the letters of the word HUNTER be arranged such that the word does not start with H?
Solution: Total arrangements $= 6! = 720$
Fix H at first position: Remaining 5 letters $\rightarrow 5! = 120$

$$\text{Valid arrangements} = 720 - 120 = \boxed{600}$$

Example 22: Permutations with Digit Reuse Prohibited

Example 22

How many 4-digit numbers greater than 3000 can be formed using digits 2, 3, 4, 5, 6 without repetition?
Solution: Thousands digit must be 3, 4, 5, or 6 (4 options) Then choose 3 more digits from remaining 4 $\rightarrow P(4, 3) = 24$

$$\text{Total} = 4 \times 24 = \boxed{96}$$

Example 23: Permutations with Specific Characters Together

Example 23

In how many ways can the letters of the word COMMITTEE be arranged such that all T's are together?
Solution: Letters = 9; Repeats: M(2), T(2), E(2)
Treat TT as a block → 8 items: C, O, M, M, I, T**T**, E, E

$$\text{Total} = \frac{8!}{2! \cdot 2! \cdot 1!} = \frac{40320}{4} = \boxed{10080}$$

Example 24: At Least One Fixed Digit

Example 24

How many 4-digit numbers can be formed from digits 1 to 9 such that digit 5 appears at least once?
Solution: Total 4-digit numbers using 1–9 (no 0): $9 \times 8 \times 7 \times 6 = 3024$
Exclude numbers without 5:
- Choose 4 digits from 8 (excluding 5): $P(8,4) = 1680$

$$\text{With at least one 5} = 3024 - 1680 = \boxed{1344}$$

Example 25: Group Arrangements with Subgroup Together

Example 25

In how many ways can 5 boys and 3 girls be seated in a row such that all the girls sit together?
Solution: Treat the 3 girls as a block $\rightarrow$ now 6 "people" to arrange $\rightarrow$ $6! = 720$
Girls within their block can be arranged in $3! = 6$

$$\text{Total} = 720 \times 6 = \boxed{4320}$$

Example 26: Letter Arrangement with Repetition and Restrictions

Example 26

How many arrangements of the word `BALLOON` can be made such that both L's are not together?
Solution: Total arrangements without restriction: Letters = 7; L(2), O(2)

$$\text{Total} = \frac{7!}{2! \cdot 2!} = \frac{5040}{4} = 1260$$

Now count cases where L's are together:
Treat LL as a block $\rightarrow$ total = 6 items: B, A, LL, O, O, N O(2):

$$\text{With L's together} = \frac{6!}{2!} = 360$$

$$\text{Not together} = 1260 - 360 = \boxed{900}$$

Example 27: Permutation with Parity Constraints

Example 27

How many 3-digit even numbers can be formed using digits 2, 3, 4, 5, 6 with no digit repeated?
Solution: Even numbers end in 2, 4, or 6 $\rightarrow$ 3 cases
Case 1: End with 2 $\rightarrow$ choose 2 digits from {3,4,5,6} $\rightarrow$ $P(4,2) = 12$

Case 2: End with 4 $\to$ {2,3,5,6} $\to P(4,2) = 12$
Case 3: End with 6 $\to$ {2,3,4,5} $\to P(4,2) = 12$

$$\text{Total} = 3 \times 12 = \boxed{36}$$

Example 28: Multiset with At Least Condition

Example 28

From the word STATISTICS, how many distinct arrangements are there where all S's are together?
Solution: Letters: 10; S(3), T(3), I(2), A(1), C(1)
Group all S's $\to$ treat as a block: 8 items: SSS, T(3), I(2), A, C

$$\text{Total} = \frac{8!}{3! \cdot 2!} = \frac{40320}{12} = \boxed{3360}$$

Example 29: Permutations with Min Distance Constraint

Example 29

How many ways can 6 people be seated in a row such that A and B have exactly one person between them?
Easy Solution (Position Method): Valid seat pairs for A and B with exactly one seat between them are: $(1,3),(2,4),(3,5),(4,6)$ — 4 pairs.
For each pair: - A and B can swap places: 2 ways. - Choose who sits between them: from the remaining 4 people: 4 ways. - Arrange the remaining 3 people in the remaining 3 seats: $3! = 6$ ways.

$$\text{Total} = 4 \times 2 \times 4 \times 3! = 4 \times 2 \times 4 \times 6 = \boxed{192}.$$

(Quick Check) Alternatively, fix a valid pair (e.g., $1,3$) and order A/B in 2 ways; then the other 4 people fill the remaining seats in $4! = 24$ ways. So $4 \times 2 \times 24 = \boxed{192}$.

Example 30: Distribution of Items in Permutations

Example 30

In how many ways can 5 identical red balls and 4 identical green balls be arranged in a row?
Solution: Total objects = 9 Red balls identical, Green balls identical

$$\text{Total arrangements} = \frac{9!}{5! \cdot 4!} = \frac{362880}{120 \cdot 24} = \boxed{126}$$

1.2 Combinations

In combinatorics, a **combination** is a selection of items where the **order does not matter**.

The number of combinations of r items that can be selected from n distinct items is given by:

$$^nC_r = \binom{n}{r} = \frac{n!}{r!(n-r)!}$$

When to Identify a Problem as Combination

Tip: When to Use Combinations

Use **combinations** when:

- The selection is made without regard to order.
- Groupings or committees are formed.
- No specific arrangement or position is mentioned.

Avoid using permutations unless the sequence/arrangement matters.

Examples of Combinations

Example 1: Committee Selection

Example 1

From a group of 10 people, how many ways can we select 4 to form a committee?
Solution:

$$\binom{10}{4} = \frac{10!}{4! \cdot 6!} = \frac{5040}{24} = \boxed{210}$$

Example 2: Handshake Problem

Example 2

In a party of 12 people, if each person shakes hands with every other person once, how many handshakes occur?
Solution: Each handshake is a combination of 2 people:

$$\binom{12}{2} = \frac{12 \cdot 11}{2} = \boxed{66}$$

Example 3: Choosing Balls

Example 3

From a bag of 15 balls, how many ways can 6 balls be drawn simultaneously?
Solution: Order doesn't matter:

$$\binom{15}{6} = \boxed{5005}$$

Example 4: Forming Teams

Example 4

A class has 8 boys and 7 girls. In how many ways can a team of 5 students be chosen such that it includes at least 2 girls?
Solution: We sum combinations for cases:

- 2 girls, 3 boys: $\binom{7}{2} \cdot \binom{8}{3} = 21 \cdot 56 = 1176$
- 3 girls, 2 boys: $\binom{7}{3} \cdot \binom{8}{2} = 35 \cdot 28 = 980$
- 4 girls, 1 boy: $\binom{7}{4} \cdot \binom{8}{1} = 35 \cdot 8 = 280$
- 5 girls: $\binom{7}{5} = 21$

Total $= 1176 + 980 + 280 + 21 = \boxed{2457}$

Example 5: Card Selection

Example 5

How many 5-card hands can be drawn from a deck of 52 cards?
Solution:

$$\binom{52}{5} = \boxed{2,598,960}$$

Example 6: Dividing Prizes

Example 6

In how many ways can 3 identical prizes be given to 10 students?
Solution: Choose 3 students from 10:

$$\binom{10}{3} = \boxed{120}$$

Example 7: Selecting Questions

Example 7

Out of 20 questions in an exam, a student must attempt 15. In how many ways can this be done?
Solution:

$$\binom{20}{15} = \binom{20}{5} = \boxed{15504}$$

Example 8: Subsets

Example 8

How many subsets of 4 elements can be formed from a set of 10 elements?
Solution:

$$\binom{10}{4} = \boxed{210}$$

Example 9: Lottery Selection

Example 9

A lottery draws 6 numbers out of 49. How many combinations are possible?
Solution:

$$\binom{49}{6} = \boxed{13,983,816}$$

Example 10: Choosing Books

Example 10

From 12 different books, in how many ways can 5 books be selected for reading?
Solution:

$$\binom{12}{5} = \boxed{792}$$

Example 11

In how many ways can a committee of 4 people be selected from 7 men and 5 women such that the committee contains at least 2 women?
Solution:

We consider all valid combinations:

$$\text{2 women, 2 men: } \binom{5}{2} \cdot \binom{7}{2} = 10 \cdot 21 = 210$$
$$\text{3 women, 1 man: } \binom{5}{3} \cdot \binom{7}{1} = 10 \cdot 7 = 70$$
$$\text{4 women: } \binom{5}{4} = 5$$

$$\text{Total} = 210 + 70 + 5 = \boxed{285}$$

Example 12

From a group of 10 people, in how many ways can a team of 5 be formed such that two particular persons are never together?

Solution:

Let the two particular persons be A and B.

Total combinations without any restriction: $\binom{10}{5} = 252$

Number of teams where A and B are together: Choose A and B (fixed), then choose 3 more from remaining 8:

$$\binom{8}{3} = 56$$

Required: $252 - 56 = \boxed{196}$

Example 13

Question: For a game in which two partners oppose two other partners, a total of 6 men are available. If every possible pair must play with every other pair, how many games are played?

Solution:

We interpret a "game" as selecting two disjoint unordered pairs (pair A vs pair B). Each distinct unordered pair-of-pairs corresponds to one game.

Method 1 (choose pairs stepwise):

1. Choose the first pair of players: there are
$$\frac{6 \times 5}{2 \times 1} = 15$$
ways to pick 2 out of 6.
2. From the remaining 4 men choose the opposing pair:
$$\frac{4 \times 3}{2 \times 1} = 6$$
ways.

3. The ordered choice (first pair, then second pair) counts each matchup twice (pair P vs Q and pair Q vs P are the same game), so divide by 2:

$$\frac{15 \times 6}{2} = 45.$$

Method 2 (choose 4 players then partition):

1. Choose which 4 of the 6 men will play in a particular game:

$$\frac{6 \times 5 \times 4 \times 3}{4 \times 3 \times 2 \times 1} = 15$$

ways.

2. For those 4 chosen players, count the number of ways to split them into two unordered pairs. The pairings of 4 distinct people into two unordered pairs equals 3 (you can check: fix one person, they have 3 choices for partner, then remaining 2 form the other pair; divide by 2 to remove order of the two pairs: $\frac{3}{1} = 3$).

3. Multiply:

$$15 \times 3 = 45.$$

$$\boxed{45}$$

Answer: $\boxed{45}$ games.

Example 14

A question paper has 10 questions. A student has to answer exactly 7. However, questions 1 and 2 are mutually exclusive (i.e., both cannot be selected together). In how many ways can the student choose 7 questions?

Solution:

Case 1: Choose Q1 but not Q2 → choose 6 from remaining 8: $\binom{8}{6} = 28$

Case 2: Choose Q2 but not Q1 → again, $\binom{8}{6} = 28$

Case 3: Choose neither Q1 nor Q2 → choose 7 from remaining 8: $\binom{8}{7} = 8$

Total $= 28 + 28 + 8 = \boxed{64}$

Example 15

How many ways can you choose a group of 6 people from 12 such that two particular people are either both in or both out of the group?

Solution:

Let A and B be the two particular people.

Case 1: Both A and B are in: choose 4 more from 10 → $\binom{10}{4} = 210$

Case 2: Both A and B are out: choose 6 from 10 → $\binom{10}{6} = 210$

Total $= 210 + 210 = \boxed{420}$

Example 16

From 8 different books, how many ways can 5 be chosen such that at least one specific book is included?
Solution:
Total ways without restriction: $\binom{8}{5} = 56$
Ways that exclude the specific book: $\binom{7}{5} = 21$
Required: $56 - 21 = \boxed{35}$

Example 17

From 9 items, including 3 identical ones, how many distinct sets of 4 items can be chosen?
Solution:
We count based on how many of the identical items are included.
Case 1: 0 identical → choose 4 from 6 distinct: $\binom{6}{4} = 15$
Case 2: 1 identical → $\binom{6}{3} = 20$
Case 3: 2 identical → $\binom{6}{2} = 15$
Case 4: 3 identical → $\binom{6}{1} = 6$
Total $= 15 + 20 + 15 + 6 = \boxed{56}$

Example 18

How many ways can a team of 4 be formed from 6 girls and 4 boys such that the team includes more girls than boys?
Solution:
Valid compositions: - 3 girls, 1 boy $\to \binom{6}{3} \cdot \binom{4}{1} = 20 \cdot 4 = 80$ - 4 girls $\to \binom{6}{4} = 15$
Total $= 80 + 15 = \boxed{95}$

Example 19

From 12 students, how many ways can 5 be chosen such that two specific students are not both included?
Solution:
Total combinations: $\binom{12}{5} = 792$
Combinations with both A and B: Choose 3 more from 10 $\to \binom{10}{3} = 120$
Required: $792 - 120 = \boxed{672}$

Example 20

From 6 red, 5 green, and 4 blue balls, how many ways can 5 balls be selected with at least one of each color?
Solution:
We count cases with at least one red, one green, one blue.
Approach: Use generating functions or inclusion-exclusion, but simpler to manually count valid (r, g, b) combinations that sum to 5 with each ≥ 1:

Possible (r, g, b) combinations:
- (3,1,1): $\binom{6}{3} \cdot \binom{5}{1} \cdot \binom{4}{1} = 20 \cdot 5 \cdot 4 = 400$ - (2,2,1): $\binom{6}{2} \cdot \binom{5}{2} \cdot \binom{4}{1} = 15 \cdot 10 \cdot 4 = 600$ - (2,1,2): $15 \cdot 5 \cdot 6 = 450$ - (1,3,1): $6 \cdot 10 \cdot 4 = 240$ - (1,2,2): $6 \cdot 10 \cdot 6 = 360$ - (1,1,3): $6 \cdot 5 \cdot 4 = 120$
Total $= 400 + 600 + 450 + 240 + 360 + 120 = \boxed{2170}$

1.3 Problems

Problem 1 *How many ways can 6 men and 4 women be seated around a round table such that no two women sit next to each other?*

Problem 2 *How many 5-digit numbers can be formed using digits from 1 to 9 such that digits are in strictly increasing order?*

Problem 3 *How many ways can 8 people be seated in a row such that persons A and B are never together and C and D are always together?*

Problem 4 *How many 7-letter words (not necessarily meaningful) can be formed using all the letters of the word "MATHLET" such that no vowel occupies an even position?*

Problem 5 *In how many ways can the word "PROGRAMMER" be rearranged such that the vowels appear in alphabetical order?*

Problem 6 *How many ways can the digits 1 to 9 be arranged in a row such that all odd digits come before even digits?*

Problem 7 *How many ways can 10 people be divided into 2 groups of 5 each such that person A and person B are in different groups?*

Problem 8 *In how many ways can the digits 1 through 6 be arranged such that 1 is never at either end and 6 is never in the middle?*

Problem 9 *In how many ways can a committee of 5 be chosen from 7 Engineers and 5 Doctors such that there are at least 2 Engineers and at least 1 Doctor?*

Problem 10 *How many ways are there to choose 6 non-consecutive integers from the first 20 positive integers?*

Problem 11 *From 12 distinct books, how many ways can you choose 5 books such that no two selected books are adjacent when placed on a shelf?*

Problem 12 *How many ways are there to select 5 letters from the word "MATHEMATICS" such that no letter repeats?*

Problem 13 *How many ways can 5 students be selected from 8 such that two particular students are either both selected or both excluded?*

Problem 14 *How many subsets of size 5 can be chosen from the set* $\{1, 2, \ldots, 15\}$ *such that no two elements are consecutive?*

Problem 15 *In how many ways can 4 teams be selected from 8 teams such that two rival teams are not selected together?*

Problem 16 *A committee of 4 is to be formed from 6 men and 5 women. Find the number of ways to form the committee such that the number of women is even.*

Problem 17 *From a group of 10 people, including A and B, 4 people are to be selected. Find how many selections include A but exclude B.*

Problem 18 *Find the number of ways of choosing 6 balls from a bag of 10 red and 8 green balls such that at least 2 red balls are included.*

Problem 19 *From a set of 12 different flowers, how many bouquets of 5 flowers can be made if two specific flowers must not appear together?*

Problem 20 *How many 4-digit even numbers have all 4 digits distinct?*

1.4 Try it Yourself

Exercise 1 *How many 10-digit numbers can be formed using the digits 0–9, such that each digit is used exactly once and the number does not start with 0?*

Exercise 2 *How many permutations of the letters in the word "STATISTICS" are there such that all the S's are together?*

Exercise 3 *From the word "ARRANGEMENT", how many distinct permutations can be formed where no two N's are adjacent?*

Exercise 4 *In how many ways can the letters of the word "ENGINEERING" be arranged such that the three E's are not together?*

Exercise 5 *How many ways can 5 boys and 5 girls be arranged in a row so that no two boys sit together?*

Exercise 6 *How many permutations of the numbers 1 to 7 are there in which exactly 3 elements are in their original positions?*

Exercise 7 *From the word "ASSESSMENT", how many permutations are possible such that no two S's are adjacent?*

Exercise 8 *How many 6-digit even numbers can be formed using digits 1 to 9 without repetition?*

Exercise 9 *How many permutations of the letters in the word "EXCELLENCE" exist such that no two E's are together?*

Exercise 10 *How many 8-letter permutations of the letters in "BASEBALL" can be made such that all vowels come together?*

Exercise 11 *From a group of 10 men and 8 women, how many ways can a committee of 6 people be formed such that it contains at least 3 women?*

Exercise 12 *In how many ways can a team of 5 people be selected from 8 men and 6 women such that the team has more women than men?*

Exercise 13 *In how many ways can a group of 4 people be selected from 6 boys and 4 girls such that the group has at least one girl?*

Exercise 14 *In how many ways can 5 persons be chosen from 10 persons such that two particular persons are never selected together?*

Exercise 15 *From 8 Mathematics books and 5 Computer books, how many ways can 4 books be selected including at least 2 Mathematics books?*

Exercise 16 *From 7 white and 6 black balls, how many ways can you choose 5 balls such that the number of white balls is at least 2?*

Exercise 17 *How many committees of 5 people can be formed from 9 people if two of them refuse to be on the same committee?*

1.5 YouTube Links and QR codes

Lecture	Details	YouTube Link	QR Code
1	Chapter 1: Introduction to counting, Applications, Permutations and Combinations Explanation	https://youtu.be/h63tqKw4e58	
2	Chapter 1: Rules, Quick Decision Table, Examples. Rules for Identifying Permutations and Combinations	https://youtu.be/-V5dn57PsGw	
3	Chapter 1.1: Permutations, Examples 1-10	https://youtu.be/Pvl8WlhbRhA	
4	Chapter 1.1: Permutations Examples 11-20	https://youtu.be/W9AyUy8268o	

5	Chapter 1.1: Permutations Examples 21-30	https://youtu.be/4TZ0tHUp1is	
6	Chapter 1.2: Combinations, Examples 1-10	https://youtu.be/PweafckcKjo	
7	Chapter 1.2: Combinations, Examples 11-20	https://youtu.be/hIc_y8EeLVM	
8	Chapter 1.3: Solutions to Problems 1-5	https://youtu.be/Zcl7dzRsde8	
9	Chapter 1.3: Solutions to Problems 6-10	https://youtu.be/gDD8iECtZRA	

10	Chapter 1.3: Solutions to Problems 11-20	https://youtu.be/6P9hPGiYjbk	

Chapter 2

Introduction to Statistics

What is Probability?

Probability is a branch of mathematics that deals with measuring the likelihood of events happening.

- Starts with a **known model or process**, like tossing a coin or rolling a die.
- Answers: *"Given a known situation, what outcomes are likely, and how likely?"*

Example: If a coin is fair, the probability of heads is 0.5.

What is Statistics?

Statistics is the science of collecting, analyzing, interpreting, and making decisions based on data.

- Starts with **sample data** and tries to learn about the underlying population.
- Answers: *"Based on what I observe, what can I say about the larger group?"*

Example: A survey of 500 voters shows that 60% favor a policy — what does this say about all voters?

How Do They Work Together?

Statistics uses **probability theory** to make **inferences** about the population.

Aspect	Probability	Statistical Inference
Starts with	Known population	Sample data
Aims to find	Likelihood of outcomes	Population parameters
Direction	From population to data	From data to population

Inference uses probability to estimate how likely an observed result is.

2.1 Motivation

Key Relationship

- **Probability** starts from the known and predicts data outcomes.
- **Statistics** starts from observed data and infers about the unknown.
- Together, they enable evidence-based decisions under uncertainty.

Applications in Real Life

1. Business Forecasting

- *Goal:* Predict future sales.
- *Method:* Use historical data to build a predictive model.
- *Inference:* Make data-driven decisions on inventory and marketing.

2. Legal Evidence

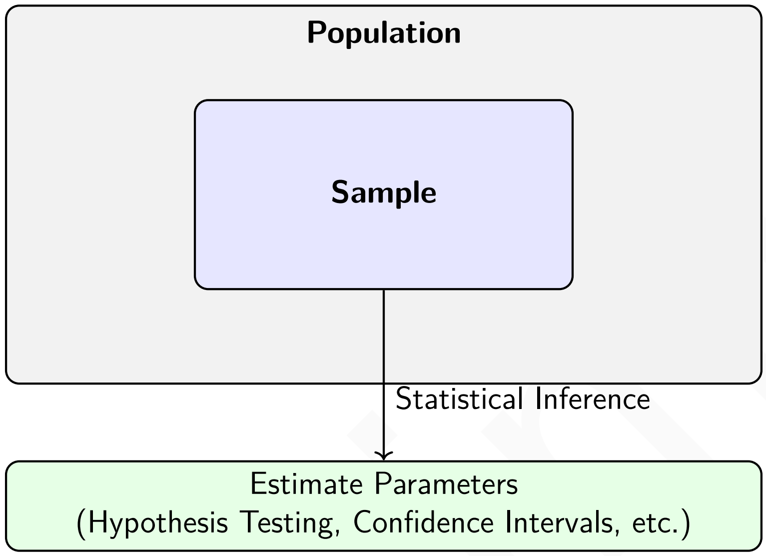
- *Goal:* Determine if a DNA match is reliable.
- *Method:* Use probability to evaluate random match likelihood.
- *Inference:* Decide strength of evidence in court.

3. Quality Control

- *Goal:* Detect faulty products.
- *Method:* Sample and analyze a few items from a batch.
- *Inference:* Accept/reject batch based on probability of defects.

Core Concept

In statistics, we are often interested in studying a large group of individuals or items called the **population**. However, it's usually impractical or impossible to collect data on every member of the population. Instead, we select a **sample** — a smaller, manageable subset — and use the information gathered from this sample to make **inferences** about the entire population.



Why Do We Sample?

- **Cost-effective:** It's cheaper to collect data from a sample than the whole population.
- **Time-saving:** Sampling takes less time than a full population study.
- **Practicality:** Often, the entire population isn't accessible.
- **Feasibility:** Data analysis becomes manageable.

How Does Inference Work?

Inference Flow

1. Define the population and a variable of interest (e.g., average income).
2. Select a representative sample from the population.
3. Collect data from the sample.
4. Compute statistics (like sample mean, variance).
5. Use probability theory to estimate population parameters or test hypotheses.

Fundamental Relationship between Probability and Statistics

Key Insight

Probability is used to model uncertainty in the data generation process.
Statistics uses the collected data (from a sample) to estimate the underlying probabilities and infer characteristics of the population.

Real-World Motivation and Example

Example: Medical Study

A pharmaceutical company wants to test the effectiveness of a new drug. Instead of testing on all patients with a disease, they:

- Select a random sample of 500 patients.
- Administer the drug and collect recovery data.
- Use statistical inference to estimate the drug's effectiveness for the entire patient population.

2.2 Data Collection

What is Sampling?

Sampling is the process of selecting a subset (sample) from a larger group (population) to gather insights and make inferences about the entire population. The goal is to collect data efficiently while maintaining accuracy and representativeness.

Steps in the Sampling and Data Collection Process

Step 1. Define the Target Population
Identify the entire group you are interested in studying (e.g., all college students in India).

Step 2. Specify the Sampling Frame
This is a list or mechanism that gives access to the population (e.g., list of students enrolled in universities).

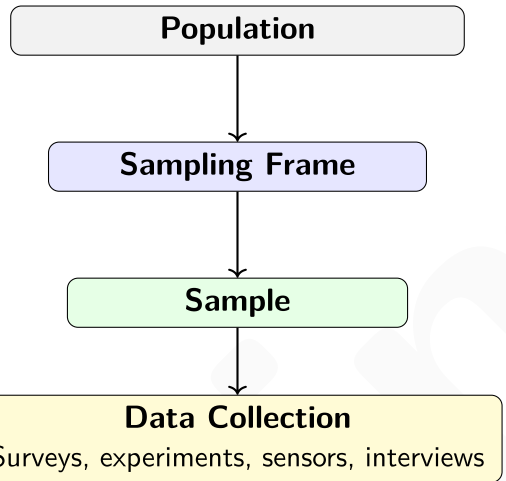
Step 3. Choose the Sampling Method
Select a technique to draw the sample:

- **Simple Random Sampling:** Every individual has an equal chance.
- **Stratified Sampling:** Divide into strata (groups) and sample from each.
- **Systematic Sampling:** Every k^{th} member is selected.
- **Cluster Sampling:** Randomly select entire groups.

Step 4. Determine the Sample Size
Choose how many observations you need based on budget, precision, and confidence level.

Step 5. Collect the Data
Use surveys, experiments, interviews, observations, or other tools to gather the information.

Step 6. Check for Bias and Errors
Ensure the data is accurate, reliable, and free from systematic errors.



Why is Sampling Important?

Advantages of Sampling

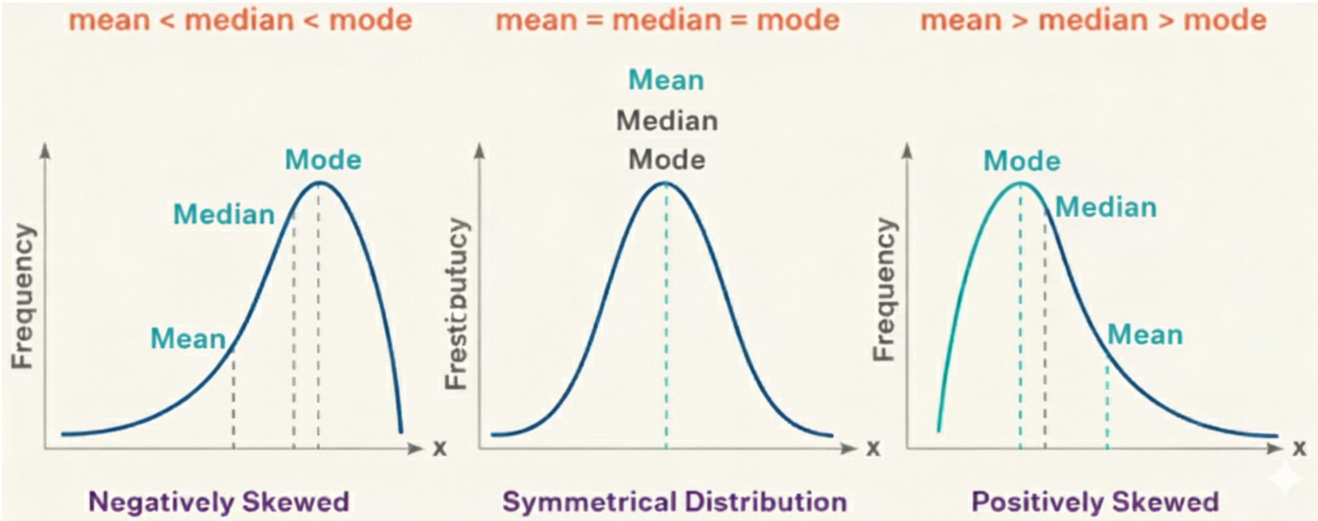
- **Cost-effective:** Saves resources by studying a small group.
- **Time-saving:** Quicker than studying the entire population.
- **Practical:** Ideal when the population is infinite or hard to reach.
- **Accurate:** Well-designed samples can produce highly reliable results.

Common Pitfalls in Data Collection

Beware of These Issues

- **Sampling Bias:** Sample is not representative of the population.
- **Non-response Bias:** Some selected participants do not respond.
- **Measurement Errors:** Faulty instruments or survey design can distort results.
- **Data Entry Errors:** Mistakes during manual or automatic input.

2.3 Measures of Location



Understanding Measures of Location

Measures of location describe the central tendency of a dataset. The two most commonly used measures are the **sample mean** and the **sample median**. These values summarize the dataset with a single representative value.

Sample Mean ($\bar{x}$)

Suppose that the observations in a sample are $x_1, x_2, \ldots, x_n$. The sample mean, denoted by $\bar{x}$, is computed as:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

- $x_1, x_2, \ldots, x_n$ are the sample observations.
- n is the sample size.
- $\bar{x}$ gives the arithmetic average of the values.

Example:

Let the sample be: $5,\ 7,\ 9,\ 4,\ 6$

$$\bar{x} = \frac{5+7+9+4+6}{5} = \frac{31}{5} = 6.2$$

Importance of Sample Mean

The sample mean is sensitive to all data points and is a good measure when data is symmetric and not heavily skewed. It is used in many statistical methods, including confidence intervals and hypothesis testing.

Sample Median

Given a set of n observations in a sample arranged in **increasing order** of magnitude as:

$$x_1 \leq x_2 \leq \cdots \leq x_n,$$

the **sample median**, denoted by $\tilde{x}$, is defined as:

$$\tilde{x} = \begin{cases} x_{(n+1)/2}, & \text{if } n \text{ is odd} \\ \frac{1}{2}\left(x_{n/2} + x_{n/2+1}\right), & \text{if } n \text{ is even} \end{cases}$$

Importance:

- The sample median provides a measure of the **central location** of the data.
- It is particularly useful when the data contains **outliers** or is **skewed**, as it is less affected by extreme values than the mean.

Example:

Example

Consider the following sample of $n = 7$ ordered values:

$$2,\ 4,\ 7,\ 10,\ 12,\ 15,\ 18$$

Since $n = 7$ is odd, the median is the $\left(\frac{7+1}{2}\right)^{\text{th}} = 4^{\text{th}}$ observation:

$$\tilde{x} = x_4 = 10$$

Now consider another sample of $n = 8$ ordered values:

$$1,\ 3,\ 6,\ 9,\ 11,\ 14,\ 17,\ 19$$

Since $n = 8$ is even, the median is the average of the 4th and 5th observations:

$$\tilde{x} = \frac{1}{2}(x_4 + x_5) = \frac{1}{2}(9 + 11) = 10$$

Comparison

- The **mean** considers all values but can be affected by outliers.
- The **median** is less sensitive to outliers and skewed distributions.

- In symmetric distributions, mean and median are approximately equal.

NOTE: $\bar{x}$ represents the mean of the sample, whereas μ represents the mean of the population.

2.4 Measures of Variability

Just as there are many measures of central tendency or location, there are many measures of spread or variability. Variability tells you how far apart points lie from each other and from the center of a distribution or a data set. Variability is also referred to as spread, scatter or dispersion. Along with measures of central tendency, measures of variability give you descriptive statistics that summarize your data. While the central tendency, or average, tells you where most of your points lie, variability summarizes how far apart they are. This is important because the amount of variability determines how well you can generalize results from the sample to your population. Low variability is ideal because it means that you can better predict information about the population based on sample data. High variability means that the values are less consistent, so it's harder to make predictions. Variability is most commonly measured with range, standard deviation and variance .

Sample Range

Definition:
The **sample range** is the difference between the largest and smallest values in a sample.

$$\text{Range} = x_{\max} - x_{\min}$$

Why It Is Used:

- Gives a quick sense of the spread.
- Easy to compute.
- Sensitive to outliers.

Example:
Sample: $5,\ 9,\ 12,\ 18,\ 22$

$$\text{Range} = 22 - 5 = 17$$

Sample Variance

Definition:
Sample variance measures the average squared deviation from the mean. It is denoted by s^2 and defined as:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Where:

- n is the number of observations,
- x_i is each observation,
- $\bar{x}$ is the sample mean.

Why It Is Used:

- Reflects the overall spread in the data.
- Fundamental in inferential statistics, including hypothesis testing and confidence intervals.

Example:
Sample: $4,\ 6,\ 8,\ 10$

$$\bar{x} = \frac{4+6+8+10}{4} = 7$$

Squared deviations:

$$(4-7)^2 = 9, \quad (6-7)^2 = 1, \quad (8-7)^2 = 1, \quad (10-7)^2 = 9$$

$$\sum (x_i - \bar{x})^2 = 9 + 1 + 1 + 9 = 20$$

$$s^2 = \frac{20}{4-1} = \frac{20}{3} \approx 6.667$$

Sample Standard Deviation

Definition:
The **sample standard deviation** is the square root of the sample variance:

$$s = \sqrt{s^2} = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2}$$

Why It Is Used:

- Represents the typical distance of data points from the mean.
- Expressed in the same units as the data.
- Widely used in data analysis and reporting.

Continuing the Previous Example:

$$s = \sqrt{6.667} \approx 2.582$$

Thus, the standard deviation gives a more intuitive sense of variability than variance.

Population Variance and Standard Deviation

NOTE: σ^2 represents the *population variance*, and σ represents the *population standard deviation*.

The formulas for population variance and standard deviation are:

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$$

Explanation:

- x_i is each individual data point.
- μ is the population mean.
- N is the total number of data points in the population.
- σ^2 is the variance — it measures the average squared deviation from the mean.
- σ is the standard deviation — it gives a sense of how spread out the data is.

Example: Given a population dataset: 2, 4, 6, 8

$$\mu = \frac{2+4+6+8}{4} = 5$$

$$\sigma^2 = \frac{(2-5)^2 + (4-5)^2 + (6-5)^2 + (8-5)^2}{4} = \frac{9+1+1+9}{4} = \frac{20}{4} = 5$$

$$\sigma = \sqrt{5} \approx 2.236$$

Understanding Sample Variance and Standard Deviation

The **sample variance** and **sample standard deviation** both measure how spread out data is, but:

- Variance uses *squared units* (e.g., cm^2), while
- Standard deviation uses *linear units* (e.g., cm).

Why are both important?

- They help understand the variability in data.
- **Variance** is used more in statistical theory.
- **Standard deviation** is more common in real-world applications.

In statistics, we often want to know about a population:

- **Population Mean** (μ)
- **Population Variance** (σ^2)

Sample variance helps us estimate σ^2 (population variance), while **sample standard deviation** is used with the sample mean $\bar{x}$ to estimate μ (population mean).

2.5 Problems

Problem 21 *Find the sample mean of the data: 5, 7, 3, 6, 9.*

Problem 22 *Calculate the sample variance for the dataset: 4, 8, 6, 5, 3.*

Problem 23 *Compute the sample standard deviation for: 12, 15, 10, 14, 13.*

Problem 24 *Given the population data: 6, 9, 5, 8, 7, find the population mean.*

Problem 25 *For the population: 4, 6, 5, 7, 8, compute the population variance.*

Problem 26 *Given population data: 10, 12, 14, 16, find the population standard deviation.*

2.6 Try it Yourself

Exercise 18 *A sample of 6 students scored: 70, 65, 80, 75, 60, 85. Calculate the sample mean and sample standard deviation.*

Exercise 19 *If a population has the data values: 5, 10, 15, 20, 25, compute the population mean and population variance.*

Exercise 20 *Find the sample variance and standard deviation for: 13, 17, 19, 21, 23.*

Exercise 21 *From a population of salaries: 25k, 30k, 35k, 40k, find the population standard deviation.*

2.7 YouTube Links and QR codes

Lecture	Details	YouTube Link	QR Code
11	Chapter 2.1–2.2: Introduction, Motivation, Data Collection and Sampling	https://youtu.be/S1w4oJRxP0Y	
12	Chapter 2.3–2.4: Measures of Central Tendency and Variability — Solutions to Problems 21-26	https://youtu.be/j0wNY7c2WEQ	

Chapter 3

The Probability Theory

Random Experiment

A **random experiment** is a process or action that leads to one of several possible outcomes, where the outcome cannot be predicted with certainty in advance.
Examples:

- Tossing a coin (possible outcomes: heads or tails)
- Rolling a die (possible outcomes: 1 to 6)
- Drawing a card from a deck
- Measuring the temperature on a given day

Sample Space

The **sample space**, denoted by S, is the set of all possible outcomes of a random experiment.
Examples:

- Tossing one coin: $S = \{H, T\}$
- Rolling a six-sided die: $S = \{1, 2, 3, 4, 5, 6\}$
- Tossing two coins: $S = \{HH, HT, TH, TT\}$

Event

An **event** is a subset of the sample space. It consists of one or more outcomes of the experiment.
Types of Events:

- **Simple Event:** Contains only one outcome.

Example: Getting a 4 when a die is rolled.

- **Compound Event:** Contains more than one outcome.
  Example: Getting an even number when a die is rolled: $\{2, 4, 6\}$
- **Certain Event:** The entire sample space S (always occurs)
- **Impossible Event:** The empty set $\varnothing$ (never occurs)

Example:
If we roll a die:

- Sample space: $S = \{1, 2, 3, 4, 5, 6\}$
- Event $A =$ rolling an even number $= \{2, 4, 6\}$

Why These Concepts Are Important

Understanding random experiments, sample space, and events is fundamental to the study of probability and statistics.
They help us:

- Model uncertainty
- Define probabilities
- Analyze real-world phenomena like games of chance, weather forecasting, and medical studies

Probability – Definition & Explanation

Probability is a measure of the **likelihood or chance** that a particular event will occur when a random experiment is performed.
Probability values range from **0 to 1**:

- **0** means the event *cannot occur*.
- **1** means the event *is certain to occur*.
- Values between 0 and 1 indicate degrees of likelihood.

To every point in the sample space, a probability (or weight) is assigned such that:

$$\sum_{x \in S} P(x) = 1$$

Where:

- S is the sample space.
- $P(x)$ is the probability of the sample point x.

Definition: If A is an event composed of multiple sample points, then the **probability of event** A, denoted $P(A)$, is the sum of the probabilities of the sample points that make up A:

$$P(A) = \sum_{x \in A} P(x)$$

Example 1: Tossing a Fair Die

Let the sample space be:

$$S = \{1, 2, 3, 4, 5, 6\}$$

Each outcome is equally likely, so:

$$P(1) = P(2) = \cdots = P(6) = \frac{1}{6}$$

Let event $A =$ getting an even number $= \{2, 4, 6\}$
Then:

$$P(A) = P(2) + P(4) + P(6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = 0.5$$

Example 2: Impossible Event

If $B = \{7\}$ when tossing a fair die, then $B \notin S$ and:

$$P(B) = 0$$

Explanation: 7 is not a possible outcome of a die roll, so the event cannot occur.

Example 3: Certain Event

Let $C = \{1, 2, 3, 4, 5, 6\}$ be the event "any outcome occurs". This includes the entire sample space.
Then:

$$P(C) = P(1) + P(2) + P(3) + P(4) + P(5) + P(6) = \frac{6}{6} = 1$$

Explanation: The event always occurs since it covers all possible outcomes.

We can also define Probability as

Probability is a numerical measure of the likelihood that a specific event will occur as a result of a random experiment. It is assigned as a real number between 0 and 1 (inclusive).

- A probability close to 1 indicates that the event is very likely to occur.
- A probability close to 0 suggests that the event is unlikely to occur.
- If all outcomes in the sample space are equally likely, then:

$$P(A) = \frac{\text{Number of favorable outcomes for event A}}{\text{Total number of possible outcomes}} = \frac{n}{N}$$

If an event cannot possibly occur (i.e., it is outside the sample space), then it is assigned a probability of 0. The sum of probabilities of all outcomes in the sample space is always 1.

Example: Online Course Quiz Attempts

An online learning platform records quiz attempts by 2500 students for a particular module:

- 30 students attempted 1 quiz
- 15 students attempted 2 quizzes
- 5 students attempted 3 quizzes
- Remaining students did not attempt any quizzes

(a) Probability that a student attempted a quiz:
Total students who attempted at least one quiz: $30 + 15 + 5 = 50$ Total students: 2500

$$P(\text{attempt at least one quiz}) = \frac{50}{2500} = 0.02$$

(b) Probability that a student attempted at least two quizzes:
Students who attempted 2 or 3 quizzes: $15 + 5 = 20$

$$P(\text{attempt at least two quizzes}) = \frac{20}{2500} = 0.008$$

(c) Probability that a student did not attempt any quizzes:
Students who did not attempt any quiz: $2500 - 50 = 2450$

$$P(\text{no attempts}) = \frac{2450}{2500} = 0.98$$

Axiom 1: Non-negativity

For any event A,

$$P(A) \geq 0.$$

Explanation: Probabilities cannot be negative — each event must have a probability that is zero or positive.

Axiom 2: Normalization

The probability of the sample space S is 1:

$$P(S) = 1$$

Explanation: This means that the sum of probabilities of all elementary outcomes must be equal to 1.

Axiom 3: Additivity

For any two mutually exclusive (disjoint) events A and B,

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

Explanation: If A and B cannot happen at the same time, the probability of either occurring is the sum of their individual probabilities.

Examples

Example 1: Non-negativity

Let the probability of rain tomorrow be $P(R) = 0.6$.
This satisfies Axiom 1 because $0.6 \geq 0$.

Example 2: Normalization

Suppose a fair die is rolled. The sample space is

$$S = \{1, 2, 3, 4, 5, 6\}$$

Each outcome has probability $\frac{1}{6}$. Then,

$$P(S) = \sum_{i=1}^{6} P(i) = 6 \times \frac{1}{6} = 1$$

Example 3: Additivity

Let event A = getting an even number when a die is rolled, i.e., $A = \{2, 4, 6\}$
Let event B = getting an odd number, i.e., $B = \{1, 3, 5\}$
These are disjoint, so:

$$P(A \cup B) = P(A) + P(B) = \frac{3}{6} + \frac{3}{6} = 1$$

Summary

The three axioms of probability form the foundation of probability theory:

1. Probabilities are always non-negative.
2. The total probability of all possible outcomes is 1.

3. For disjoint events, the probability of their union is the sum of their individual probabilities.

Understanding and applying these axioms is essential for solving any probability problem.

3.1 Probability Rules

Rule 1: Probability lies between 0 and 1

For any event A, the probability is a number between 0 and 1:

$$0 \leq P(A) \leq 1$$

Example 1: Tossing a coin. Let event A be getting heads. Then $P(A) = 0.5 \in [0, 1]$
Example 2: Rolling a die. Let event B be getting number 3. $P(B) = \frac{1}{6} \approx 0.167 \in [0, 1]$

Rule 2: Sum of all probabilities equals 1

The sum of probabilities of all possible elementary events in a sample space is always 1.

$$\sum_{i=1}^{n} P(E_i) = 1$$

Example 1: Tossing a coin.

$$P(H) + P(T) = 0.5 + 0.5 = 1$$

Example 2: Rolling a die.

$$P(1) + P(2) + P(3) + P(4) + P(5) + P(6) = 6 \times \frac{1}{6} = 1$$

Rule 3: Probability of the Impossible Event is 0

The probability of the empty set (an impossible event) is 0.

$$P(\emptyset) = 0$$

Example 1: Rolling a 7 on a standard die: $P(7) = 0$
Example 2: Drawing a purple card from a standard deck: $P(\text{purple}) = 0$

Rule 4: Probability of the Certain Event is 1

The probability of the sample space S, representing a certain event, is 1.

$$P(S) = 1$$

Example 1: Drawing a card from a 52-card deck. Event: "drawing a card from the deck" — always true.
Example 2: Tossing a coin. Event: getting either head or tail.

$$P(H \cup T) = 1$$

Rule 5: Complement Rule

The probability of the complement of an event A is:

$$P(A^c) = 1 - P(A)$$

Example 1: Rolling a die. Let A be the event of getting an even number: $P(A) = \frac{3}{6} = 0.5$
Then $P(A^c) = 1 - 0.5 = 0.5$
Example 2: Tossing a coin. Let A be getting heads, $P(A) = 0.5 \Rightarrow P(A^c) = 0.5$

Rule 6: Addition Rule for Mutually Exclusive Events

If A and B are mutually exclusive events, then:

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

Example 1: Tossing a coin: $A = \text{Head}, B = \text{Tail}$. Then:

$$P(A \cup B) = 0.5 + 0.5 = 1$$

Example 2: Rolling a die: $A = 2, B = 5 \Rightarrow P(A \cup B) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$

Rule 7: Addition Rule for Any Two Events

If A and B are any two events:

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

Example 1: Drawing from a deck, $A = \text{red card}, B = \text{king}$

$$P(A) = \frac{26}{52}, \quad P(B) = \frac{4}{52}, \quad P(A \cap B) = \frac{2}{52}$$
$$P(A \cup B) = \frac{26 + 4 - 2}{52} = \frac{28}{52}$$

Example 2: In a class: - 40 play football (A), 30 play cricket (B), 10 play both. - Total = 100 students.

$$P(A \cup B) = \frac{40 + 30 - 10}{100} = \frac{60}{100} = 0.6$$

Rule 8: Multiplication Rule for Independent Events

If A and B are independent:

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

Example 1: Tossing a coin and rolling a die. - A = Heads, $P(A) = 0.5$ - $B = 3$ on die, $P(B) = \frac{1}{6}$

$$P(A \cap B) = 0.5 \cdot \frac{1}{6} = \frac{1}{12}$$

Example 2: Two separate bulbs with failure probabilities of 0.1 and 0.2:

$$P(\text{both fail}) = 0.1 \cdot 0.2 = 0.02$$

Mutually Exclusive Events

Two events A and B are said to be **mutually exclusive** (or disjoint) if they cannot occur at the same time. That is,

$$P(A \cap B) = 0$$

This means that the occurrence of one event excludes the possibility of the other.
Key Identification Tip: If two events describe outcomes that cannot happen together in a single trial of the experiment, they are mutually exclusive.

Example 1: Tossing a die

- Let event A = getting an even number.
- Let event B = getting an odd number.
- Then $A = \{2, 4, 6\}$, $B = \{1, 3, 5\}$
- $A \cap B = \emptyset$, so $P(A \cap B) = 0$

Example 2: Drawing cards

- Let A = drawing a red card
- Let B = drawing a black card
- Both cannot occur in the same draw, hence they are mutually exclusive

Independent Events

Two events A and B are said to be **independent** if the occurrence of one does not affect the probability of the other occurring. That is,

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

Key Identification Tip: If you can say "the probability of one event is not influenced by the outcome of the other," they are likely independent.

Example 1: Tossing two coins

- Event A = first coin shows heads
- Event B = second coin shows tails
- $P(A) = 0.5, P(B) = 0.5$
- $P(A \cap B) = P(A) \cdot P(B) = 0.5 \cdot 0.5 = 0.25$
- Hence, A and B are independent

Example 2: Rolling a die and tossing a coin

- Event A = die shows a 6
- Event B = coin shows heads
- One doesn't influence the other, so:
- $P(A) = 1/6, P(B) = 1/2, P(A \cap B) = 1/12$
- $P(A \cap B) = P(A) \cdot P(B) \Rightarrow$ Independent

How to Identify in Problems

- **Mutually Exclusive:** Check if events can happen together. If not, they are mutually exclusive.
- **Independent:** Compute $P(A \cap B)$ and compare it to $P(A) \cdot P(B)$. If equal, the events are independent.
- *Important:* Mutually exclusive events are never independent (unless one event has probability 0).

3.2 Conditional Probability

Conditional Probability

Conditional probability is the probability of one event occurring with some relationship to one or more other events. If A and B are two events and $P(B) > 0$, then the conditional probability of A given B is defined as:

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

This formula is read as "the probability of A given B equals the probability of A and B occurring together divided by the probability of B."

Key Concepts

- **Multiplication Rule:** $P(A \cap B) = P(A \mid B) \cdot P(B) = P(B \mid A) \cdot P(A)$
- If A and B are **independent**, then $P(A \mid B) = P(A)$ and $P(B \mid A) = P(B)$.
- **Bayes' Theorem:** $P(B \mid A) = \frac{P(A \mid B)P(B)}{P(A)}$

Example 1

A box contains 5 red balls and 3 green balls. One ball is drawn at random. Given that it is not green, what is the probability that it is red?
Solution: Total outcomes: 8 (5 red + 3 green).
Let event A: ball is red, event B: ball is not green = red.
Then $P(A|B) = \frac{P(A \cap B)}{P(B)}$
Since red ∩ not green = red: $P(A \cap B) = \frac{5}{8}$, $P(B) = \frac{5}{8}$
Thus, $P(A|B) = \frac{\frac{5}{8}}{\frac{5}{8}} = 1$

Example 2

A card is drawn from a deck of 52 cards. Given that it is a face card, what is the probability it is a king?
Solution: There are 12 face cards (Jack, Queen, King from each suit), and 4 Kings.
$P(King|Face) = \frac{P(King \cap Face)}{P(Face)} = \frac{4/52}{12/52} = \frac{1}{3}$

Example 3

Two dice are thrown. What is the probability that the sum is 9, given that the sum is odd?
Solution: Total odd sums: 3,5,7,9,11 → Count = 18.
Sum = 9 possible pairs: (3,6), (4,5), (5,4), (6,3) → 4 ways.
Only those with odd sum: All 4 valid.
So, $P(9|odd) = \frac{4}{18} = \frac{2}{9}$

Example 4

A family has two children. Given that one is a girl, what is the probability that both are girls?
Solution: Possible combinations: BB, BG, GB, GG → remove BB → 3 cases.
Only one case is GG, so $P(GG|at\ least\ one\ G) = \frac{1}{3}$

Example 5

A student passes a subject with 0.6 probability. What is the probability he passes both subjects if passing second subject given first is 0.5?
Solution: $P(A) = 0.6,\ P(B|A) = 0.5$
$P(A \cap B) = P(A) \cdot P(B|A) = 0.6 \times 0.5 = 0.3$

Example 6

A test is positive in 95% of people with disease, and 2% of people without disease. 1% have disease. What is the probability that a person has the disease if the test is positive?
Solution: Let D = has disease, T = test positive
$P(D) = 0.01,\ P(T|D) = 0.95,\ P(T|\neg D) = 0.02,\ P(\neg D) = 0.99$
$P(T) = 0.01 \cdot 0.95 + 0.99 \cdot 0.02 = 0.0095 + 0.0198 = 0.0293$
$P(D|T) = \frac{0.0095}{0.0293} \approx 0.324$

Example 7

A machine has 3 parts: A, B, and C. Failure probability of each is 0.1. Given that one failed, what is the probability it was part A?
Solution: Failures equally likely: $P(A|one\ fails) = \frac{1}{3}$

Example 8

Two cards drawn without replacement. Find probability that second card is king given first is king.
Solution: $P(K_2|K_1) = \frac{3}{51}$ (Only 3 kings left out of 51)

Example 9

A number is chosen from 1 to 100. Given it is divisible by 5, what is probability it is also divisible by 10?
Solution: Divisible by 5: 20 numbers.
Divisible by 10: 10 numbers → All divisible by 5 too.
So $P(10|5) = \frac{10}{20} = \frac{1}{2}$

Example 10

A box contains 2 white and 3 black balls. Two balls are drawn one after another without replacement. Find the probability that the second is black given that the first is white.
Solution: Total: 5 balls. If first is white (2/5), then left: 1 white, 3 black = 4 balls.
$P(B_2|W_1) = \frac{3}{4}$

3.3 Bayes Theorem

Definition of Bayes' Theorem

Bayes' theorem allows us to update the probability of an event based on new information. It relates the conditional and marginal probabilities of random events.
Formula:

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

Where:

- $P(A|B)$ is the probability of event A given that event B has occurred.
- $P(B|A)$ is the probability of event B given that event A has occurred.
- $P(A)$ is the prior probability of event A.
- $P(B)$ is the marginal probability of event B.

Law of Total Probability

Let $E_1, E_2, \ldots, E_n$ be a partition of the sample space S:

$$E_i \cap E_j = \varnothing \ (i \neq j), \quad \bigcup_{i=1}^{n} E_i = S, \quad \mathbb{P}(E_i) > 0.$$

Then for any event A,

$$\mathbb{P}(A) = \sum_{k=1}^{n} \mathbb{P}(A \mid E_k)\, \mathbb{P}(E_k).$$

Bayes' Rule for Multiple Events

Given $E_1, \ldots, E_n$ a partition with $\mathbb{P}(E_k) > 0$ and $\mathbb{P}(A) > 0$, Bayes' theorem states:

$$\mathbb{P}(E_i \mid A) = \frac{\mathbb{P}(A \mid E_i)\, \mathbb{P}(E_i)}{\sum_{k=1}^{n} \mathbb{P}(A \mid E_k)\, \mathbb{P}(E_k)}, \quad i = 1, \ldots, n.$$

Examples

Example 1

A disease affects 1% of the population. A test for the disease is 95% accurate, meaning:

- If someone has the disease, the test is positive with probability 0.95.

- If someone does not have the disease, the test is negative with probability 0.95.

What is the probability that someone who tests positive actually has the disease?
Solution:
Let:

- D = person has disease $\Rightarrow P(D) = 0.01$
- $\neg D$ = person doesn't have disease $\Rightarrow P(\neg D) = 0.99$
- T = test is positive
- $P(T|D) = 0.95$ (true positive)
- $P(T|\neg D) = 0.05$ (false positive)

Use Bayes' Theorem:

$$\begin{aligned} P(D|T) &= \frac{P(T|D)\cdot P(D)}{P(T|D)\cdot P(D) + P(T|\neg D)\cdot P(\neg D)} \\ &= \frac{0.95\cdot 0.01}{0.95\cdot 0.01 + 0.05\cdot 0.99} \\ &= \frac{0.0095}{0.0095 + 0.0495} \\ &= \frac{0.0095}{0.059} \\ &\approx 0.161 \end{aligned}$$

So, the probability is approximately **16.1%**.

Example 2

In a certain city, 60% of taxis are green and 40% are blue. A witness identifies a blue taxi involved in an accident. The witness is correct 80% of the time. What is the probability that the taxi was actually blue?
Solution:
Let:

- B = taxi is blue $\Rightarrow P(B) = 0.4$
- G = taxi is green $\Rightarrow P(G) = 0.6$
- W = witness says blue
- $P(W|B) = 0.8,\ P(W|G) = 0.2$

Using Bayes' Theorem:

$$\begin{aligned} P(B|W) &= \frac{P(W|B) \cdot P(B)}{P(W|B) \cdot P(B) + P(W|G) \cdot P(G)} \\ &= \frac{0.8 \cdot 0.4}{0.8 \cdot 0.4 + 0.2 \cdot 0.6} \\ &= \frac{0.32}{0.32 + 0.12} \\ &= \frac{0.32}{0.44} \\ &\approx 0.727 \end{aligned}$$

So, the probability that the taxi was actually blue is approximately **72.7%**.

Example 3

A factory has 3 machines:

- Machine A produces 30% of items and has a 2% defect rate.
- Machine B produces 45% of items and has a 3% defect rate.
- Machine C produces 25% of items and has a 1% defect rate.

If an item is defective, what is the probability it was produced by Machine B?
Solution:
Let D = item is defective, and B = produced by Machine B.

$$P(D) = (0.3 \cdot 0.02) + (0.45 \cdot 0.03) + (0.25 \cdot 0.01) = 0.006 + 0.0135 + 0.0025 = 0.022$$

$$P(B|D) = \frac{P(D|B) \cdot P(B)}{P(D)} = \frac{0.03 \cdot 0.45}{0.022} = \frac{0.0135}{0.022} \approx 0.614$$

So, the probability that a defective item came from Machine B is approximately **61.4%**.

3.4 GATE PYQs

PYQ1: GATE CSE 2025 - Set 1: Probability of Error-Free Transmission

Question: Suppose a 5-bit message is transmitted from a source to a destination through a noisy channel. The probability that a bit of the message gets flipped during transmission is 0.01. Flipping of each bit is independent of one another. The probability that the message is delivered error-free to the destination is __________. (rounded off to three decimal places)

Solution:

- Total number of bits in the message $= 5$
- Probability that a single bit is received **correctly** $= 1 - 0.01 = 0.99$

- Since bit flips are **independent**, the probability that all 5 bits are received correctly is:

$$P(\text{All correct}) = 0.99^5$$

- Now compute:

$$0.99^5 = (0.99)^5 = 0.95099 \quad \text{(approximately)}$$

Final Answer: **0.951**

PYQ2: GATE CSE 2025 - Set 1: Coin Toss and Bayes' Theorem

Question: A box contains 5 coins: 4 regular coins and 1 fake coin. When a regular coin is tossed, the probability $P(\text{head}) = 0.5$, and for a fake coin, $P(\text{head}) = 1$. You pick a coin at random and toss it twice, and get two heads. The probability that the coin you have chosen is the fake coin is ________ (rounded off to two decimal places).

Solution:

- Let:
  - F: Event that the chosen coin is fake.
  - R: Event that the chosen coin is regular.
  - H_2: Event of getting two heads.
- We need to find:

$$P(F \mid H_2) = \frac{P(H_2 \mid F) \cdot P(F)}{P(H_2)}$$

- Prior probabilities:

$$P(F) = \frac{1}{5}, \quad P(R) = \frac{4}{5}$$

- Likelihoods:

$$P(H_2 \mid F) = 1 \quad \text{(Fake coin always gives head)}$$
$$P(H_2 \mid R) = 0.5 \times 0.5 = 0.25$$

- Total probability of getting two heads:

$$P(H_2) = P(H_2 \mid F) \cdot P(F) + P(H_2 \mid R) \cdot P(R)$$

$$P(H_2) = 1 \cdot \frac{1}{5} + 0.25 \cdot \frac{4}{5} = \frac{1}{5} + \frac{1}{5} = \frac{2}{5}$$

- Apply Bayes' Theorem:

$$P(F \mid H_2) = \frac{1 \cdot \frac{1}{5}}{\frac{2}{5}} = \frac{1}{2} = 0.50$$

Final Answer: **0.50**

PYQ3: When six unbiased dice are rolled simultaneously, the probability of getting all distinct numbers (i.e., 1, 2, 3, 4, 5, and 6) is:

A. $\frac{1}{324}$

B. $\frac{5}{324}$

C. $\frac{7}{324}$

D. $\frac{11}{324}$

Solution:
The total number of outcomes when 6 dice are rolled:

$$\text{Total outcomes} = 6^6$$

To get all distinct numbers from 1 to 6, we must arrange the numbers $1, 2, 3, 4, 5, 6$ in some order. The number of such arrangements is:

$$\text{Favorable outcomes} = 6!$$

So the required probability is:

$$P = \frac{6!}{6^6} = \frac{720}{46656} = \frac{5}{324}$$

Final Answer

Correct Option: (B) $\frac{5}{324}$

PYQ4: GATE CSE 2024 SET-1: Conditional Probability with Drawing Without Replacement

Question:
A bag contains 10 red balls and 15 blue balls. Two balls are drawn randomly without replacement. Given that the first ball drawn is red, the probability (rounded off to 3 decimal places) that both balls drawn are red is

Solution:
There are: - 10 red balls, - 15 blue balls, - Total balls = 25
We're given that the **first ball is red**, so we condition on this event.
Now, we compute:

$$P(\text{Second is red} \mid \text{First is red}) = \frac{\text{Remaining red balls}}{\text{Remaining total balls}} = \frac{9}{24} = \frac{3}{8} = 0.375$$

Hence, the probability that both balls drawn are red **given that the first is red** is:

$$\boxed{0.375}$$

Final Answer

Answer: 0.375

PYQ5: GATE CSE 2024 SET-1: Events and Probability Relationships

Question: Let A and B be two events in a probability space with:

$$P(A) = 0.3, \quad P(B) = 0.5, \quad P(A \cap B) = 0.1$$

Which of the following statements is/are TRUE?

(A) The two events A and B are independent

(B) $P(A \cup B) = 0.7$

(C) $P(A \cap B^c) = 0.2$, where B^c is the complement of B

(D) $P(A^c \cap B^c) = 0.4$, where A^c and B^c are the complements of A and B

Solution:
(A) Independence check:
Two events A and B are independent if:

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

$$\Rightarrow 0.1 \stackrel{?}{=} 0.3 \cdot 0.5 = 0.15 \quad \text{(False)}$$

So, **(A) is False**.

(B) Union:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.3 + 0.5 - 0.1 = 0.7$$

(B) is True

(C) $P(A \cap B^c)$:
Use identity: $A = (A \cap B) \cup (A \cap B^c)$, disjoint events.

$$P(A \cap B^c) = P(A) - P(A \cap B) = 0.3 - 0.1 = 0.2$$

(C) is True

(D) $P(A^c \cap B^c)$:
Use De Morgan:

$$P(A^c \cap B^c) = 1 - P(A \cup B) = 1 - 0.7 = 0.3$$

(D) is False

Correct Answers

Correct Options: (B) and (C)

PYQ6: GATE CSE 2023 – Probability of Independent Events (Fair Coins)

Question: Consider a random experiment where two fair coins are tossed. Let A be the event that denotes HEAD on both the throws, B be the event that denotes HEAD on the first throw, and C be the event that denotes HEAD on the second throw. Which of the following statements is/are TRUE?

A. A and B are independent.

B. A and C are independent.

C. B and C are independent.

D. $\Pr(B \mid C) = \Pr(B)$

Solution: (C) (D)
Sample space when two fair coins are tossed:

$$S = \{HH, HT, TH, TT\}$$

Each outcome has probability $\frac{1}{4}$.

- $A = \{HH\}$
- $B = \{HH, HT\}$ (HEAD on first coin)
- $C = \{HH, TH\}$ (HEAD on second coin)

Now calculate probabilities:

$$\Pr(A) = \frac{1}{4}, \quad \Pr(B) = \frac{1}{2}, \quad \Pr(C) = \frac{1}{2}$$

A and B:

$$A \cap B = \{HH\} \Rightarrow \Pr(A \cap B) = \frac{1}{4}$$

$$\Pr(A) \cdot \Pr(B) = \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8} \neq \frac{1}{4} \Rightarrow \textbf{A and B are NOT independent}$$

A and C:

$$A \cap C = \{HH\} \Rightarrow \Pr(A \cap C) = \frac{1}{4}$$

$$\Pr(A) \cdot \Pr(C) = \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8} \neq \frac{1}{4} \Rightarrow \textbf{A and C are NOT independent}$$

B and C:

$$B \cap C = \{HH\} \Rightarrow \Pr(B \cap C) = \frac{1}{4}$$

$$\Pr(B) \cdot \Pr(C) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} \Rightarrow \textbf{B and C are independent}$$

$\Pr(B \mid C)$**:**

$$\Pr(B \mid C) = \frac{\Pr(B \cap C)}{\Pr(C)} = \frac{1/4}{1/2} = \frac{1}{2} = \Pr(B) \Rightarrow \textbf{Statement D is also TRUE}$$

Correct Answers: C and D

PYQ7: GATE CSE 2021 SET-2: Pólya's Urn Process

Question:
A bag has r red balls and b black balls. All balls are identical except for their colours. In a trial, a ball is randomly drawn from the bag, its colour is noted and the ball is placed back into the bag along with another ball of the same colour. Note that the number of balls in the bag increases by one after each trial. A sequence of four such trials is conducted. Which one of the following choices gives the probability of drawing a red ball in the **fourth** trial?

A. $\frac{r}{r+b}$

B. $\frac{r}{r+b+3}$

C. $\frac{r+3}{r+b+3}$

D. $\frac{r}{r+b} \cdot \frac{r+1}{r+b+1} \cdot \frac{r+2}{r+b+2} \cdot \frac{r+3}{r+b+3}$

Solution:
This is a classic example of the **Pólya's Urn Model**, where each draw reinforces the drawn color by adding another of the same.
A key property of this process is:

$$\text{Probability of drawing red at any trial} = \frac{r}{r+b}$$

This remains **unchanged in expectation**, regardless of the number of trials.
So, the probability of drawing a red ball on the fourth trial is:

$$\boxed{\frac{r}{r+b}}$$

Final Answer

Answer: A. $\frac{r}{r+b}$

PYQ8: GATE CSE: Conditional Probability with Bayes' Theorem

Question:
Let the temperature in Guwahati (G) be high, medium, or low with the following probabilities:

$$P(H_G) = 0.2, \quad P(M_G) = 0.5, \quad P(L_G) = 0.3$$

The conditional probabilities of Delhi (D) temperatures given Guwahati's temperature are:

	H_D	M_D	L_D
H_G	0.40	0.48	0.12
M_G	0.10	0.65	0.25
L_G	0.01	0.50	0.49

Find $P(H_G \mid H_D)$, i.e., the probability that Guwahati has high temperature given that Delhi has high temperature.

Solution:
Using **Bayes' Theorem**:

$$P(H_G \mid H_D) = \frac{P(H_D \mid H_G) \cdot P(H_G)}{P(H_D)}$$

We compute $P(H_D)$ using the **Law of Total Probability**:

$$P(H_D) = P(H_D \mid H_G)P(H_G) + P(H_D \mid M_G)P(M_G) + P(H_D \mid L_G)P(L_G)$$

$$P(H_D) = (0.40 \cdot 0.2) + (0.10 \cdot 0.5) + (0.01 \cdot 0.3)$$
$$P(H_D) = 0.08 + 0.05 + 0.003 = 0.133$$

Now apply Bayes' Theorem:

$$P(H_G \mid H_D) = \frac{0.40 \cdot 0.2}{0.133} = \frac{0.08}{0.133} \approx 0.6015$$

Final Answer

Answer: 0.60 (rounded to 2 decimal places)

PYQ9: GATE CSE 2018: Dice Rolling - Probability of Win on Third Trial

Question:
Two people, P and Q, independently roll two identical fair 6-faced dice. The person with the lower number wins. In case of a tie (i.e., both roll the same number), they repeat the trial (i.e., both roll again). A trial is defined as a single round of dice rolls by both players.
Assuming all die outcomes are equally likely and trials are independent, what is the probability (rounded to 3 decimal places) that one of them wins on the third trial?

Solution:
Each trial has three possible outcomes: - P wins - Q wins - Tie
Tie occurs when both roll the same number. For a die with faces 1 to 6:

$$P(\text{Tie}) = \frac{6}{36} = \frac{1}{6}$$

$$P(\text{Win}) = 1 - P(\text{Tie}) = \frac{5}{6}$$

To win on the **third trial**, the first two trials must be ties, and the third trial must be a win (by either P or Q):

$$P(\text{Win on 3rd trial}) = P(\text{Tie}) \times P(\text{Tie}) \times P(\text{Win}) = \left(\frac{1}{6}\right)^2 \times \frac{5}{6} = \frac{5}{216}$$

Convert to decimal:

$$\frac{5}{216} \approx 0.023$$

Final Answer

Answer: 0.023

PYQ10: GATE CSE 2017 (Set 2): Conditional Probability – Job Application

Question:
P and Q are considering applying for a job.

- Probability that P applies: $P(P) = \frac{1}{4}$
- Probability that P applies given Q applies: $P(P|Q) = \frac{1}{2}$
- Probability that Q applies given P applies: $P(Q|P) = \frac{1}{3}$

What is the probability that P does not apply given Q does not apply?

Solution:
Let's denote:
- $A =$ P applies - $B =$ Q applies
We are given:

$$P(A) = \frac{1}{4}, \quad P(A|B) = \frac{1}{2}, \quad P(B|A) = \frac{1}{3}$$

We need to find:

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

We know:

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \Rightarrow P(A \cap B) = P(B|A) \cdot P(A) = \frac{1}{3} \cdot \frac{1}{4} = \frac{1}{12}$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \Rightarrow P(B) = \frac{P(A \cap B)}{P(A|B)} = \frac{1/12}{1/2} = \frac{1}{6}$$

Now we compute: - $P(A \cap B^c) = P(A) - P(A \cap B) = \frac{1}{4} - \frac{1}{12} = \frac{1}{6}$ - $P(A^c \cap B^c) = 1 - P(A \cup B)$
First, calculate:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{4} + \frac{1}{6} - \frac{1}{12} = \frac{3+2-1}{12} = \frac{4}{12} = \frac{1}{3}$$

$$\Rightarrow P(A^c \cap B^c) = 1 - \frac{1}{3} = \frac{2}{3}$$

And

$$P(B^c) = 1 - P(B) = 1 - \frac{1}{6} = \frac{5}{6}$$

So finally:

$$P(A^c|B^c) = \frac{P(A^c \cap B^c)}{P(B^c)} = \frac{2/3}{5/6} = \frac{2/3 \times 6/5}{1} = \frac{12}{15} = \frac{4}{5}$$

Final Answer

Answer: (A) $\frac{4}{5}$

PYQ11: GATE CSE 2016 (Set 1): Probability – Coin Tossing with Looping Condition

Question:
Consider the following experiment.

1. Step 1: Flip a fair coin twice.
2. Step 2: If the outcomes are (TAILS, HEADS), then output Y and stop.
3. Step 3: If the outcomes are either (HEADS, HEADS) or (HEADS, TAILS), then output N and stop.
4. Step 4: If the outcomes are (TAILS, TAILS), then go to Step 1.

The probability that the output of the experiment is Y is (up to two decimal places) ____.

Solution:
Let's compute the probability of getting an output Y.
Let $P(Y)$ be the probability that the experiment ends with output Y.
There are 4 equally likely outcomes when flipping a fair coin twice:

- (H,H) $\rightarrow$ output N
- (H,T) $\rightarrow$ output N
- (T,H) $\rightarrow$ output Y
- (T,T) $\rightarrow$ repeat the process

So:

$$P(\text{Y in one trial}) = \frac{1}{4}, \quad P(\text{N in one trial}) = \frac{1}{2}, \quad P(\text{Repeat}) = \frac{1}{4}$$

Let P be the total probability that the experiment eventually outputs Y.
We can set up the recurrence:

$$P = \frac{1}{4} \cdot 1 + \frac{1}{4} \cdot P \Rightarrow P - \frac{1}{4}P = \frac{1}{4} \Rightarrow \frac{3}{4}P = \frac{1}{4} \Rightarrow P = \frac{1}{3}$$

So the required probability is:

Final Answer

Answer: $\boxed{0.33}$

PYQ12: GATE CSE 2014 (Set 3): Maximum of Product of Probabilities

Question:
Let S be a sample space and two mutually exclusive events A and B be such that $A \cup B = S$. If $P(\cdot)$ denotes the probability of an event, what is the maximum value of $P(A) \cdot P(B)$?

Solution:
Since A and B are mutually exclusive, $P(A \cap B) = 0$, and since $A \cup B = S$, we have:

$$P(A \cup B) = P(S) = 1 \Rightarrow P(A) + P(B) = 1$$

Let $P(A) = x \Rightarrow P(B) = 1 - x$
We need to maximize:

$$P(A) \cdot P(B) = x(1 - x)$$

This is a quadratic expression:

$$f(x) = x - x^2$$

Maximum of $f(x)$ occurs at $x = \frac{1}{2}$, and:

$$f\left(\frac{1}{2}\right) = \frac{1}{2} \cdot \left(1 - \frac{1}{2}\right) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

Final Answer

Answer: $\boxed{\frac{1}{4}}$

PYQ13: GATE CSE 2014 (Set 2): Probability Not Divisible by 2, 3, or 5

Question:
The probability that a given positive integer lying between 1 and 100 (both inclusive) is **NOT divisible** by 2, 3, or 5 is .

Solution:
Total numbers from 1 to 100 $=$ 100
Let us use the **inclusion-exclusion principle** to count numbers divisible by 2, 3, or 5.

Let:

$$A = \text{Set of numbers divisible by } 2 \Rightarrow \left\lfloor \frac{100}{2} \right\rfloor = 50$$
$$B = \text{Set of numbers divisible by } 3 \Rightarrow \left\lfloor \frac{100}{3} \right\rfloor = 33$$
$$C = \text{Set of numbers divisible by } 5 \Rightarrow \left\lfloor \frac{100}{5} \right\rfloor = 20$$

Now calculate overlaps:

$$A \cap B : \left\lfloor \frac{100}{6} \right\rfloor = 16$$
$$A \cap C : \left\lfloor \frac{100}{10} \right\rfloor = 10$$
$$B \cap C : \left\lfloor \frac{100}{15} \right\rfloor = 6$$
$$A \cap B \cap C : \left\lfloor \frac{100}{30} \right\rfloor = 3$$

By inclusion-exclusion:

$$|A \cup B \cup C| = 50 + 33 + 20 - 16 - 10 - 6 + 3 = 74$$

So, numbers **not divisible** by 2, 3, or 5:

$$100 - 74 = 26$$

Hence, required probability:

$$\frac{26}{100} = \boxed{0.26}$$

Final Answer

Answer: $\boxed{0.26}$

PYQ14: GATE CSE 2014 (Set 1): Dice Sum Probability

Question:
Four fair six-sided dice are rolled. The probability that the sum of the results is 22 is $\frac{X}{1296}$. Find the value of X.

Solution:
Each die has 6 outcomes, so total number of outcomes:

$$6^4 = 1296$$

We need to count the number of outcomes where the sum of four dice is 22.

Let the numbers on the four dice be x_1, x_2, x_3, x_4 such that:

$$1 \leq x_i \leq 6, \quad \text{and} \quad x_1 + x_2 + x_3 + x_4 = 22$$

We transform it into an integer partition problem with bounds:
Let $y_i = x_i - 1$, so that $y_i \geq 0$, and $y_i \leq 5$
Then the equation becomes:

$$y_1 + y_2 + y_3 + y_4 = 22 - 4 = 18 \quad \text{with } 0 \leq y_i \leq 5$$

We need to find the number of integer solutions to this bounded equation. This is a standard problem solved using **generating functions** or inclusion-exclusion.
Using brute-force or known enumeration, the number of combinations where the sum is 22 is:

$$\boxed{10}$$

So the probability is:

$$\frac{10}{1296}$$

Final Answer

Answer: $\boxed{X = 10}$

PYQ15: GATE CSE 2012: Conditional Dice Roll Probability

Question:
Suppose a fair six-sided die is rolled once. - If the value on the die is 1, 2, or 3, the die is rolled a second time. - If the value is 4, 5, or 6, it is not rolled again.
What is the probability that the **sum total** of values that turn up is **at least 6**?
Options: (A) $\frac{10}{20}$ (B) $\frac{5}{12}$ (C) $\frac{2}{3}$ (D) $\frac{1}{6}$

Solution:
Let's divide the sample space into two parts based on the first die roll:
1. First roll is 4, 5, or 6 $\rightarrow$ no second roll. - Prob $= \frac{3}{6} = \frac{1}{2}$ - The value is $\geq$ 4. - If it is 6 $\rightarrow$ satisfies sum $\geq$ 6 (Yes) - If it is 4 or 5 $\rightarrow$ does NOT satisfy sum $\geq$6 - So success only if first roll $=$ 6 $\rightarrow$ probability $= \frac{1}{6}$
2. First roll is 1, 2, or 3 $\rightarrow$ roll again. - Prob $= \frac{3}{6} = \frac{1}{2}$ - We need to find the number of successful outcomes where the sum of two rolls $\geq$ 6 - Enumerate:

- If first roll is 1:
  - To get sum $\geq$ 6, second roll $\geq$ 5 $\rightarrow$ 2 outcomes (5, 6)
  - Prob $= \frac{1}{6} \times \frac{2}{6} = \frac{2}{36}$
- If first roll is 2:
  - Need second roll $\geq$ 4 $\rightarrow$ 3 outcomes (4, 5, 6)
  - Prob $= \frac{1}{6} \times \frac{3}{6} = \frac{3}{36}$

- If first roll is 3:
  - Need second roll $\geq 3 \rightarrow$ 4 outcomes (3, 4, 5, 6)
  - Prob $= \frac{1}{6} \times \frac{4}{6} = \frac{4}{36}$

Total success from second roll branch:

$$\frac{2+3+4}{36} = \frac{9}{36} = \frac{1}{4}$$

Total probability of success:

$$P(\text{sum} \geq 6) = P(\text{first roll} = 6) + P(\text{two rolls with sum} \geq 6) = \frac{1}{6} + \frac{1}{4} = \frac{5}{12}$$

Final Answer

Correct Option: (B) $\boxed{\frac{5}{12}}$

PYQ16: GATE CSE 2011 - Probability of Sequential Cards

A deck of 5 cards (each carrying a distinct number from 1 to 5) is shuffled thoroughly. Two cards are then removed one at a time from the deck. What is the probability that the two cards are selected with the number on the first card being one higher than the number on the second card?
Solution

$$\text{First card number} = \text{Second card number} + 1$$

Total number of ways to choose 2 cards (in order): $5 \times 4 = 20$
Favorable outcomes: Check the pairs where first = second + 1:

$$(2, 1), (3, 2), (4, 3), (5, 4)$$

These are 4 favorable outcomes.

$$\text{Probability} = \frac{\text{Favorable outcomes}}{\text{Total outcomes}} = \frac{4}{20} = \frac{1}{5}$$

Correct Answer: A) $\frac{1}{5}$

PYQ17: GATE CSE 2011 - Conditional Probability with Coins

Problem: Two fair coins are flipped. It is known that **at least one** of them shows **Head**. What is the probability that **both are Heads**?
Sample space (S) of 2 coin tosses:

$$S = \{HH, HT, TH, TT\}$$

Given: At least one head $\rightarrow$ eliminate TT

$$\text{Reduced sample space} = \{HH, HT, TH\}$$

Favorable outcome = {HH}

$$\text{Required probability} = \frac{1}{3}$$

Correct Answer: A) $\frac{1}{3}$

PYQ18: GATE CSE 2010 – Probability of Declared Faulty Computer

Question:
Consider a company that assembles computers. The probability of a faulty assembly of any computer is p. The company therefore subjects each computer to a testing process. This testing process gives the correct result for any computer with a probability of q.
What is the probability that a computer is declared faulty?
Options:

A) $pq + (1-p)(1-q)$

B) $(1-q)p$

C) $(1-p)q$

D) pq

Solution:
A computer can be declared faulty in two cases:
Case 1: The computer is **actually faulty** and the test correctly identifies it.
Probability $= p \cdot q$
Case 2: The computer is **not faulty** but the test incorrectly declares it faulty.
Probability $= (1-p)\cdot(1-q)$
Total Probability:

$$P(\text{Declared faulty}) = pq + (1-p)(1-q)$$

Correct Answer: Option A

PYQ19: GATE CSE 2008

Question: Aishwarya studies either computer science (CS) or mathematics (Maths) every day. - If she studies CS on a day, then the probability that she studies Maths the next day is 0.6. - If she studies Maths on a day, then the probability that she studies CS the next day is 0.4.
Given that Aishwarya studies computer science on Monday, what is the probability that she studies computer science on Wednesday?
Options: A) 0.24 B) 0.36 C) 0.4 D) 0.6
Solution:
Answer: The probability that Aishwarya studies CS on Wednesday is 0.4.
Correct Option: C) 0.4

PYQ20: GATE

A pair of six-faced dice is rolled thrice. The probability that the sum of the outcomes in each roll equals 4 in exactly two of the three attempts is . (round off to three decimal places)

Solution:

$$E = \{(1,3),(3,1),(2,2)\}$$

$$n(E) = 3, \quad n(S) = 36$$

$$p = P(E) = \frac{n(E)}{n(S)} = \frac{3}{36} = \frac{1}{12}$$

$$q = P(\overline{E}) = 1 - \frac{1}{12} = \frac{11}{12}$$

$$P(x) = \binom{3}{2}(p^2)(q^1) = 3 \times \left(\frac{1}{12}\right)^2 \left(\frac{11}{12}\right) = 0.02$$

Set Theory Laws and Probability Counterparts

Law	Set Theory	Probability Form
Commutative	$A \cup B = B \cup A$ $A \cap B = B \cap A$	$\mathbb{P}(A \cup B) = \mathbb{P}(B \cup A)$ $\mathbb{P}(A \cap B) = \mathbb{P}(B \cap A)$ Variation: $\mathbb{P}(A \cup B \cup C) = \mathbb{P}(C \cup A \cup B)$
Associative	$(A \cup B) \cup C = A \cup (B \cup C)$ $(A \cap B) \cap C = A \cap (B \cap C)$	$\mathbb{P}((A \cup B) \cup C) = \mathbb{P}(A \cup (B \cup C))$ $\mathbb{P}((A \cap B) \cap C) = \mathbb{P}(A \cap (B \cap C))$
Distributive	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$	$\mathbb{P}(A \cup (B \cap C)) = \mathbb{P}((A \cup B) \cap (A \cup C))$ $\mathbb{P}(A \cap (B \cup C)) = \mathbb{P}((A \cap B) \cup (A \cap C))$
Identity	$A \cup \varnothing = A$ $A \cap S = A$	$\mathbb{P}(A \cup \varnothing) = \mathbb{P}(A)$ $\mathbb{P}(A \cap S) = \mathbb{P}(A)$
Complement	$A \cup A^c = S$ $A \cap A^c = \varnothing$	$\mathbb{P}(A \cup A^c) = 1$ $\mathbb{P}(A \cap A^c) = 0$ Variation: $\mathbb{P}(A^c \cap B^c) = 1 - \mathbb{P}(A \cup B)$
Idempotent	$A \cup A = A$ $A \cap A = A$	$\mathbb{P}(A \cup A) = \mathbb{P}(A)$ $\mathbb{P}(A \cap A) = \mathbb{P}(A)$
Domination	$A \cup S = S$ $A \cap \varnothing = \varnothing$	$\mathbb{P}(A \cup S) = 1$ $\mathbb{P}(A \cap \varnothing) = 0$
De Morgan	$(A \cup B)^c = A^c \cap B^c$ $(A \cap B)^c = A^c \cup B^c$	$\mathbb{P}((A \cup B)^c) = \mathbb{P}(A^c \cap B^c)$ $\mathbb{P}((A \cap B)^c) = \mathbb{P}(A^c \cup B^c)$ Variation: $(A \cup B \cup C)^c = A^c \cap B^c \cap C^c$
Difference	$A - B = A \cap B^c$	$\mathbb{P}(A - B) = \mathbb{P}(A \cap B^c) = \mathbb{P}(A) - \mathbb{P}(A \cap B)$
Double Complement	$(A^c)^c = A$	$\mathbb{P}((A^c)^c) = \mathbb{P}(A)$
Inclusion-Exclusion	–	2 events: $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ 3 events: $\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C)$

3.5 Problems

Problem 27 *A box contains 4 red and 6 black balls. Two balls are drawn at random without replacement. What is the probability that both are red?*

Problem 28 *An urn contains 5 white and 7 black balls. Two balls are drawn one after another with replacement. What is the probability that both balls are black?*

Problem 29 *In a class of 30 students, 18 like Mathematics, 12 like Physics, and 8 like both. What is the probability that a randomly selected student likes at least one of the two subjects?*

Problem 30 *A card is drawn from a pack of 52 cards. What is the probability that it is either a king or a heart?*

Problem 31 *A bag contains 3 red, 5 green, and 2 blue balls. Three balls are drawn without replacement. What is the probability that all are of different colors?*

Problem 32 *Two dice are thrown. What is the probability that the sum is a multiple of 4?*

Problem 33 *In a group of 100 people, 60 like tea, 50 like coffee, and 20 like both. What is the probability that a randomly selected person likes tea or coffee?*

Problem 34 *Three machines A, B, and C produce 20%, 30%, and 50% of items respectively. A produces 1% defective, B produces 2% defective, and C produces 3% defective. If an item is found defective, what is the probability it was produced by machine B?*

Problem 35 *A die is rolled twice. What is the probability that the first roll is greater than the second?*

Problem 36 *A speaks the truth 60% of the time, and B speaks the truth 75% of the time. What is the probability that they contradict each other on a fact?*

Problem 37 *Two cards are drawn from a well-shuffled pack of 52 cards. What is the probability that both are aces?*

Problem 38 *Three fair coins are tossed. What is the probability of getting at least two heads?*

Problem 39 *A committee of 3 is to be formed from 6 men and 4 women. What is the probability that the committee contains at least one woman?*

Problem 40 *If $P(A) = 0.5$, $P(B) = 0.6$, and $P(A \cap B) = 0.3$, find $P(A \cup B)$.*

Problem 41 *A student answers 5 multiple-choice questions. Each question has 4 choices with only one correct. What is the probability that exactly 2 answers are correct if she guesses all answers?*

Problem 42 *A lot has 3 defective and 7 non-defective bulbs. Two bulbs are selected at random. What is the probability that exactly one is defective?*

Problem 43 *In a certain college, 40% of the students are girls. 70% of the girls study science while only 50% of the boys study science. A student is selected at random and is found to be studying science. What is the probability that the student is a girl?*

Problem 44 *A box contains 10 bulbs out of which 3 are defective. Two bulbs are selected one by one without replacement. What is the probability that both are defective?*

Problem 45 *A person is known to hit a target 3 times out of 4. What is the probability that he hits the target at least once in 2 attempts?*

Problem 46 *An exam has 3 sections. A student has a 70% chance of passing section A, 60% chance for B, and 80% for C independently. What is the probability that he passes in at least one section?*

3.6 Try it Yourself

Exercise 22 *An environmental agency is testing water quality in three lakes.*

(a) *List the elements of a sample space S, using S for safe and U for unsafe.*

(b) *List the elements of S corresponding to event E that at least two lakes are safe.*

(c) *Define an event whose elements are $\{SSS, USS, SSU, SUS\}$.*

Exercise 23 *If $S = \{x \mid 0 < x < 15\}$, $A = \{x \mid 2 < x < 10\}$, and $B = \{x \mid 0 < x < 6\}$, find*

(a) $A \cup B$;

(b) $A \cap B$;

(c) $A' \cap B'$.

Exercise 24 *How many distinct permutations can be made from the letters of the word FINANCE?*

(a) *Total distinct permutations;*

(b) *How many start with the letter F?*

Exercise 25 *How many three-digit numbers can be formed from the digits $1, 2, 3, 4, 5, 6, 7$ if each digit is used only once?*

(a) *Total three-digit numbers;*

(b) *How many are even numbers;*

(c) *How many are greater than 450?*

Exercise 26 *A shelf contains 4 novels, 4 poetry books, and a dictionary. Three books are picked at random. Find the probability that*

(a) *the dictionary is selected;*

(b) *2 novels and 1 poetry book are selected.*

Exercise 27 *A letter is chosen at random from the English alphabet. Find the probability that it*

(a) *is a vowel (excluding y);*

(b) *comes before the letter L;*

(c) comes after the letter H.

Exercise 28 *An automobile has a 0.3 probability of needing an oil change and 0.5 probability of needing a new filter. If the probability that both are needed is 0.12, find*

(a) the probability the filter is needed given oil change;

(b) the probability the oil needs changing given the filter is needed.

Exercise 29 *A paint store sells latex and gloss paints. Probability a customer buys latex is 0.8. Of latex buyers, 50% buy rollers; of gloss buyers, 25% buy rollers. A customer buys a roller and paint. Find the probability the paint is latex.*

Exercise 30 *A box contains 5 black and 5 red balls. Three balls are drawn with replacement. Find the probability that*

(a) all are the same color;

(b) each color appears at least once.

Exercise 31 *A shipment of 10 TVs contains 2 defective units. In how many ways can a hotel purchase 4 TVs and get at least 1 defective unit?*

3.7 YouTube Links and QR codes

Lecture	Details	YouTube Link	QR Code
13	Chapter 3.1–3.2: Probability Theory — Axioms — Rules — Examples	https://youtu.be/S8TwqeHD_SI	
14	Chapter 3.3–3.4: Conditional Probability — Bayes Theorem — Multiple Events — Examples	https://youtu.be/yjgesOAq6pw	
15	Chapter 3.5: Conditional Probability — Bayes Theorem — GATE PYQs 1–10	https://youtu.be/kj_5VbAPtDQ	
16	Chapter 3.5: Conditional Probability — Bayes Theorem — GATE PYQs 11–20	https://youtu.be/qGpulGj4y_M	

17	Chapter 3.6: Probability Theory — Solutions to Problems 27–46	https://youtu.be/eB5s4QW4NDc	

Chapter 4

The Random Variables

4.1 Introduction

Random Variable

Definition: A **random variable (RV)** is a function that assigns a real number to each outcome in the sample space of a random experiment.
There are two types:

- **Discrete Random Variable:** Takes a countable number of values.
- **Continuous Random Variable:** Takes values in a continuous range.

Examples:

Example 1. Coin Tossing:

- Experiment: Toss a fair coin 3 times.
- Sample space: $\{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$
- Define X = number of heads. Then X is a discrete RV taking values $\{0, 1, 2, 3\}$.

Example 2. Dice Roll:

- Experiment: Roll a fair 6-sided die once.
- Sample space: $\{1, 2, 3, 4, 5, 6\}$
- Define Y = value shown on the die. Y is a discrete RV taking values from 1 to 6.

Example 3. Card Draw:

- Experiment: Draw one card from a deck.
- Define $Z = 1$ if red card drawn, $Z = 0$ otherwise.
- Z is a discrete RV with values $\{0, 1\}$.

Example 4. Measuring Height:

- Experiment: Measure the height of a randomly selected student.
- Define H = height in cm. H can take any real value in an interval, e.g., $[120, 200]$.
- H is a continuous RV.

Summary: A random variable converts qualitative outcomes into quantitative values for probabilistic analysis.

4.2 Overview of Discrete Probability Distributions

Discrete Probability Distribution

A **Discrete Random Variable (DRV)** is a variable that can take on a countable number of values, each with a corresponding probability.
A **Discrete Probability Distribution** describes the probability that a discrete random variable takes each of its possible values.
Probability Mass Function (PMF): Let X be a discrete random variable. The function $f(x)$ defined as:

$$f(x) = P(X = x)$$

is called the **Probability Mass Function** (PMF), which satisfies:

1. $f(x) \geq 0$ for all x
2. $\sum_x f(x) = 1$
3. $P(X = x) = f(x)$

Cumulative Distribution Function (CDF): The cumulative distribution function $F(x)$ gives the probability that X is less than or equal to x:

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t)$$

Example 1: Tossing a Coin 3 Times

Let X be the number of heads observed when a fair coin is tossed 3 times.
Sample Space: Each toss has 2 outcomes: H or T. Total outcomes: $2^3 = 8$ Enumerating

outcomes:

- 0 heads: TTT
- 1 head: HTT, THT, TTH
- 2 heads: HHT, HTH, THH
- 3 heads: HHH

PMF Table:

x	$P(X = x)$
0	$\frac{1}{8}$
1	$\frac{3}{8}$
2	$\frac{3}{8}$
3	$\frac{1}{8}$

CDF Table:

x	$F(x) = P(X \leq x)$
0	$\frac{1}{8}$
1	$\frac{4}{8}$
2	$\frac{7}{8}$
3	1

Example 2: Random Helmet Assignment

Let M be the number of employees getting back their own helmets when 3 helmets are randomly returned.
Possible values: $M = 0, 1, 3$
PMF:

m	$P(M = m)$
0	$\frac{1}{3}$
1	$\frac{1}{2}$
3	$\frac{1}{6}$

Check: $\frac{1}{3} + \frac{1}{2} + \frac{1}{6} = 1$
This is a valid discrete probability distribution.

Example 3: Defective Laptops in a Shipment

A small company buys 2 laptops at random from a shipment of 20, where 3 are defective. Let X be the number of defectives selected.
Total ways to choose 2 laptops out of 20: $\binom{20}{2} = 190$
Now, calculate:

- $P(X = 0) =$ Both are non-defective: $\binom{17}{2}/190 = 136/190$
- $P(X = 1) =$ One defective, one good: $\binom{3}{1}\binom{17}{1}/190 = 51/190$
- $P(X = 2) =$ Both defective: $\binom{3}{2}/190 = 3/190$

PMF Table:

x	$P(X = x)$
0	$\frac{136}{190}$
1	$\frac{51}{190}$
2	$\frac{3}{190}$

Check: $\frac{136+51+3}{190} = 1$

Example

Problem: A car dealership sells a particular foreign car, 50% of which are equipped with side airbags. Let X denote the number of cars with side airbags among the next 4 cars sold.

1. Determine a formula for the **probability distribution** of X.
2. Find the **cumulative distribution function (CDF)** of X.

Solution:
We are dealing with a **Binomial distribution** problem where:

- Number of trials: $n = 4$
- Probability of success (a car with airbags): $p = 0.5$
- Random variable X: number of cars with airbags among the next 4 sold

1. Probability Mass Function (PMF)
For a binomial random variable, the PMF is given by:

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}, \quad x = 0, 1, 2, 3, 4$$

Substituting $n = 4$ and $p = 0.5$:

$$P(X = x) = \binom{4}{x} (0.5)^x (0.5)^{4-x} = \binom{4}{x} \cdot (0.5)^4 = \binom{4}{x} \cdot \frac{1}{16}$$

2. PMF Values:

$$P(X=0)=\frac{1}{16}, \quad P(X=1)=\frac{4}{16},$$
$$P(X=2)=\frac{6}{16}, \quad P(X=3)=\frac{4}{16},$$
$$P(X=4)=\frac{1}{16}$$

3. Probability Distribution Table:

x	$P(X=x)$
0	$\frac{1}{16}$
1	$\frac{4}{16}$
2	$\frac{6}{16}$
3	$\frac{4}{16}$
4	$\frac{1}{16}$

4. Cumulative Distribution Function (CDF):

x	$F(x)=P(X\leq x)$
0	$\frac{1}{16}$
1	$\frac{1+4}{16}=\frac{5}{16}$
2	$\frac{5+6}{16}=\frac{11}{16}$
3	$\frac{11+4}{16}=\frac{15}{16}$
4	$\frac{15+1}{16}=1$

4.3 Overview of Continuous Probability Distributions

Continuous Random Variables and Probability Density Function (PDF)

A **continuous random variable** can assume an infinite number of values in a given range. The probability that a continuous random variable takes on any exact value is zero. This is because between any two real numbers, no matter how close, there are infinitely many values.

Example: Consider the heights of people over 21 years of age. The probability that a randomly chosen person is exactly 164 cm tall is **0**. However, the probability that the height lies within an interval, say between 163 and 165 cm, is greater than 0.

For a continuous random variable X, the probability is computed over an interval:

$$P(a < X < b) = \int_a^b f(x)\,dx$$

Here, $f(x)$ is called the **probability density function (pdf)** and must satisfy:

1. $f(x) \geq 0$ for all $x \in \mathbb{R}$
2. $\int_{-\infty}^{\infty} f(x)\,dx = 1$
3. $P(a < X < b) = \int_a^b f(x)\,dx$

Example: PDF Verification and Probability Calculation

Let X be a continuous random variable with pdf:

$$f(x) = \begin{cases} \frac{x^2}{3}, & \text{if } -1 < x < 2 \\ 0, & \text{otherwise} \end{cases}$$

(a) Verify that $f(x)$ is a valid pdf
We need to check:

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-1}^{2} \frac{x^2}{3}dx$$

$$= \frac{1}{3}\int_{-1}^{2} x^2 dx = \frac{1}{3}\left[\frac{x^3}{3}\right]_{-1}^{2} = \frac{1}{3}\left(\frac{8}{3} - \left(-\frac{1}{3}\right)\right) = \frac{1}{3} \cdot \frac{9}{3} = 1$$

So, $f(x)$ is a valid pdf.

(b) Find $P(0 < X \leq 1)$

$$P(0 < X \leq 1) = \int_0^1 \frac{x^2}{3}dx = \frac{1}{3}\int_0^1 x^2 dx = \frac{1}{3}\left[\frac{x^3}{3}\right]_0^1 = \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}$$

Answer: $P(0 < X \leq 1) = \frac{1}{9}$

Definition: Cumulative Distribution Function

The cumulative distribution function $F(x)$ of a continuous random variable X with density function $f(x)$ is given by:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt, \quad \text{for } -\infty < x < \infty.$$

As a consequence of this definition, we also have:

$$P(a < X < b) = F(b) - F(a),$$
$$f(x) = \frac{dF(x)}{dx}, \quad \text{if the derivative exists.}$$

Example: Finding CDF and Probability from PDF

Recall from previous example, the probability density function:

$$f(x) = \begin{cases} \frac{x^2}{3}, & \text{for } -1 < x < 2, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Finding the cumulative distribution function $F(x)$:
We compute $F(x) = \int_{-\infty}^{x} f(t)\,dt$ in different intervals:

- For $x \leq -1$:

$$F(x) = 0$$

- For $-1 < x < 2$:

$$F(x) = \int_{-1}^{x} \frac{t^2}{3}\,dt = \frac{1}{3}\int_{-1}^{x} t^2\,dt = \frac{1}{3}\left[\frac{t^3}{3}\right]_{-1}^{x} = \frac{1}{9}(x^3 + 1)$$

- For $x \geq 2$:

$$F(x) = \int_{-1}^{2} \frac{t^2}{3}\,dt = \frac{1}{3}\left[\frac{t^3}{3}\right]_{-1}^{2} = \frac{1}{9}(8 + 1) = 1$$

So, the cumulative distribution function is:

$$F(x) = \begin{cases} 0, & x \leq -1, \\ \frac{1}{9}(x^3 + 1), & -1 < x < 2, \\ 1, & x \geq 2. \end{cases}$$

(b) Use $F(x)$ to compute $P(0 < X \leq 1)$:

$$P(0 < X \leq 1) = F(1) - F(0) = \frac{1}{9}(1^3 + 1) - \frac{1}{9}(0^3 + 1) = \frac{2}{9} - \frac{1}{9} = \frac{1}{9}$$

Example 1: Identifying Random Variables (Question and Solution)

Question:
Determine whether each of the following random variables is **discrete** or **continuous**:

- X: the number of traffic tickets issued in a city per month.

- Y: the duration (in hours) of a marathon race.
- M: the yearly milk yield from a specific cow.
- N: the number of eggs a hen lays in a week.
- P: the monthly number of building inspections conducted in a town.
- Q: the weight of wheat harvested per acre.

Solution:
Classify each variable based on whether it assumes countable values (discrete) or can take any value within a range (continuous):

- X **is Discrete:** Tickets are counted as whole numbers (0, 1, 2, ...).
- Y **is Continuous:** Race duration can take any positive real value (e.g., 4.75 hours).
- M **is Continuous:** Milk yield varies continuously (e.g., 2135.6 liters/year).
- N **is Discrete:** Eggs are counted in integers.
- P **is Discrete:** Number of inspections is a countable quantity.
- Q **is Continuous:** Wheat weight can take any value within a range (e.g., 3.42 tons/acre).

Example 2: Random Variable for Selected Automobiles

Question:
A shipment contains 5 foreign cars, 2 of which are special edition models. An agency randomly selects 3 cars from this shipment.

1. List all possible selections of 3 cars, using S for special edition and R for regular models.
2. Define a random variable X as the number of special edition cars in the selection, and determine its value for each selection.

Solution:
We have 5 cars in total: 2 special edition (S) and 3 regular (R). The total number of ways to choose 3 cars is:

$$\binom{5}{3} = 10$$

Step 1: Sample space
All distinct subsets of 3 cars:

$$S = \{\{S_1, S_2, R_1\}, \{S_1, S_2, R_2\}, \{S_1, S_2, R_3\}, \{S_1, R_1, R_2\}, \{S_1, R_1, R_3\}, \{S_1, R_2, R_3\}, \{S_2, R_1, R_2\}, \{S_2, R_1, R$$

Step 2: Assign values of X
The random variable X counts the number of special edition cars in each subset:

Sample Point	Value of X
{S1, S2, R1}	2
{S1, S2, R2}	2
{S1, S2, R3}	2
{S1, R1, R2}	1
{S1, R1, R3}	1
{S1, R2, R3}	1
{S2, R1, R2}	1
{S2, R1, R3}	1
{S2, R2, R3}	1
{R1, R2, R3}	0

Hence, the random variable X can take the values $0, 1, 2$ with the sample space completely enumerated.

Example 3: Sample Space for Coin Tossing Until 3 Successive Heads

Question: A coin is flipped until 3 heads in succession occur. List only those elements of the sample space that require 6 or fewer tosses. Is this a discrete sample space? Explain.

Solution:
We are to record sequences of coin tosses that result in the first occurrence of three successive heads (HHH) and require at most 6 tosses.
Step 1: Understand the stopping condition: We stop the experiment as soon as 3 consecutive heads occur. So we only include sequences that end with HHH and do not have HHH earlier in the sequence.
Step 2: List all such sequences of length ≤ 6 that end with the first occurrence of HHH:

- Length 3: `HHH`
- Length 4: `THHH`
- Length 5: `TTHHH, HTHH, HTTHH`
- Length 6: `TTTHHH, THTHHH, HTTHHH, HHTTHH, HTHTHH`

(The list can be extended exhaustively by generating all valid sequences of length ≤ 6 that end with the `HHH` and have no earlier HHH. This process requires a state-based approach or tree traversal for accuracy.)

Step 3: Is the sample space discrete?
Yes, this is a **discrete sample space** because:

- The outcomes (coin sequences) can be listed individually.
- Each outcome corresponds to a finite sequence of tosses.
- There are a countable number of possible outcomes even though the experiment could theoretically continue beyond 6 tosses.

Conclusion: The sample space consists of all sequences that end with the first appearance of 3 successive heads, and the total set is countable. Therefore, it is a discrete sample space.

Example 4: Shelf Life Probability of Medicine

Question: The shelf life, in days, for bottles of a certain prescribed medicine is a random variable having the density function

$$f(x) = \begin{cases} \frac{20{,}000}{(x+100)^3}, & x > 0 \\ 0, & \text{elsewhere} \end{cases}$$

Find the probability that a bottle of this medicine will have a shelf life of (a) at least 200 days; (b) anywhere from 80 to 120 days.

Solution:
This is a continuous probability density function. To find probabilities, we integrate the density over the desired interval.

—

(a) Probability that shelf life is at least 200 days: We want $P(X \geq 200) = \int_{200}^{\infty} \frac{20000}{(x+100)^3} dx$
Let us make the substitution: Let $u = x + 100 \Rightarrow du = dx$, when $x = 200$, $u = 300$

$$\begin{aligned} P(X \geq 300) &= \int_{300}^{\infty} \frac{20000}{u^3}\, du \\ &= 20000 \int_{300}^{\infty} u^{-3}\, du \\ &= 20000 \left[\frac{u^{-2}}{-2} \right]_{300}^{\infty} \\ &= 20000 \left(0 - \left(\frac{-1}{2 \cdot 300^2} \right) \right) \\ &= \frac{20000}{2 \cdot 300^2} \\ &= \frac{20000}{180000} \\ &= \frac{1}{9} \end{aligned}$$

$$\boxed{P(X \geq 200) = \frac{1}{9}}$$

—

(b) Probability that shelf life is between 80 and 120 days: We want $P(80 \leq X \leq 120) = \int_{80}^{120} \frac{20000}{(x+100)^3} dx$
Use substitution $u = x + 100 \Rightarrow du = dx$ When $x = 80$, $u = 180$; When $x = 120$, $u = 220$

$$\begin{aligned}
P(180 \leq X \leq 220) &= \int_{180}^{220} \frac{20000}{u^3} \, du \\
&= 20000 \int_{180}^{220} u^{-3} \, du \\
&= 20000 \left[\frac{u^{-2}}{-2} \right]_{180}^{220} \\
&= 20000 \left(\frac{-1}{2 \cdot 220^2} + \frac{1}{2 \cdot 180^2} \right) \\
&= \frac{20000}{2} \left(\frac{1}{180^2} - \frac{1}{220^2} \right) \\
&= 10000 \left(\frac{1}{32400} - \frac{1}{48400} \right) \\
&= 10000 \left(\frac{48400 - 32400}{32400 \cdot 48400} \right) \\
&= 10000 \cdot \frac{16000}{1{,}568{,}160{,}000} \\
&= \frac{160{,}000{,}000}{1{,}568{,}160{,}000} \\
&= \boxed{\frac{10}{98.01} \approx 0.102}
\end{aligned}$$

So,

$$\boxed{P(80 \leq X \leq 120) \approx 0.102}$$

Example 5: Continuous Distribution Problem

Problem:
The total number of hours, measured in units of 100 hours, that a family runs a vacuum cleaner over a period of one year is a continuous random variable X with the density function:

$$f(x) = \begin{cases} x, & 0 < x < 1 \\ 2 - x, & 1 \leq x < 2 \\ 0, & \text{otherwise} \end{cases}$$

Find the probability that over a period of one year, a family runs their vacuum cleaner:
(a) less than 120 hours;
(b) between 50 and 100 hours.

Solution:
The given PDF $f(x)$ is a piecewise function over the domain $(0, 2)$. Since x is in 100-hour units:

- 120 hours $= 1.2$ units

- 50 hours = 0.5 units
- 100 hours = 1 unit

(a) Probability that $X < 1.2$:

$$P(X < 1.2) = \int_0^1 x\,dx + \int_1^{1.2} (2-x)\,dx$$

First integral:

$$\int_0^1 x\,dx = \left[\frac{x^2}{2}\right]_0^1 = \frac{1}{2}$$

Second integral:

$$\int_1^{1.2} (2-x)\,dx = \left[2x - \frac{x^2}{2}\right]_1^{1.2} = \left(2(1.2) - \frac{(1.2)^2}{2}\right) - \left(2(1) - \frac{(1)^2}{2}\right)$$

$$= (2.4 - 0.72) - (2 - 0.5) = 1.68 - 1.5 = 0.18$$

So,

$$P(X < 1.2) = 0.5 + 0.18 = \boxed{0.68}$$

(b) Probability that $0.5 < X < 1$:
This falls entirely in the first part of the PDF $f(x) = x$:

$$P(0.5 < X < 1) = \int_{0.5}^1 x\,dx = \left[\frac{x^2}{2}\right]_{0.5}^1 = \frac{1}{2} - \frac{(0.5)^2}{2} = \frac{1}{2} - \frac{0.25}{2} = \frac{1}{2} - 0.125 = \boxed{0.375}$$

Example 6: Probability Question with Solution

Problem:
The proportion of people who respond to a certain mails is a continuous random variable X that has the density function:

$$f(x) = \begin{cases} \dfrac{2(x+2)}{5}, & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

(a) Show that $P(0 < X < 1) = 1$.
(b) Find the probability that more than $\frac{1}{4}$ but less than $\frac{1}{2}$ of the people respond.

Solution:
(a) We verify that the total probability over the range of the PDF is 1:

$$P(0 < X < 1) = \int_0^1 \frac{2(x+2)}{5}\,dx = \frac{2}{5}\int_0^1 (x+2)\,dx$$

$$= \frac{2}{5}\left[\frac{x^2}{2} + 2x\right]_0^1 = \frac{2}{5}\left(\frac{1}{2} + 2\right) = \frac{2}{5} \cdot \frac{5}{2} = 1$$

Hence, the PDF is valid.

(b) We find the probability:

$$P\left(\frac{1}{4} < X < \frac{1}{2}\right) = \int_{1/4}^{1/2} \frac{2(x+2)}{5}\,dx = \frac{2}{5}\int_{1/4}^{1/2}(x+2)\,dx$$

$$= \frac{2}{5}\left[\frac{x^2}{2} + 2x\right]_{1/4}^{1/2}$$

Evaluating:

$$\left(\frac{1}{2}\right)^2/2 + 2\cdot\frac{1}{2} = \frac{1}{8} + 1 = \frac{9}{8}$$

$$\left(\frac{1}{4}\right)^2/2 + 2\cdot\frac{1}{4} = \frac{1}{32} + \frac{1}{2} = \frac{17}{32}$$

$$\Rightarrow \frac{9}{8} - \frac{17}{32} = \frac{36-17}{32} = \frac{19}{32}$$

$$\Rightarrow \frac{2}{5}\cdot\frac{19}{32} = \frac{38}{160} = \boxed{\frac{19}{80} = 0.2375}$$

Example 7: Probability Question with Solution

Problem:
A continuous random variable X that can assume values between $x = 2$ and $x = 5$ has a density function given by:

$$f(x) = \frac{2(1+x)}{27}, \quad 2 \le x \le 5$$

Find:

(a) $P(X < 4)$

(b) $P(3 \le X < 4)$

Solution:

(a) Compute $P(X < 4)$:

$$P(X < 4) = \int_2^4 \frac{2(1+x)}{27}\,dx = \frac{2}{27}\int_2^4 (1+x)\,dx$$

$$= \frac{2}{27}\left[x + \frac{x^2}{2}\right]_2^4 = \frac{2}{27}\left(\left(4+\frac{16}{2}\right) - \left(2+\frac{4}{2}\right)\right)$$

$$= \frac{2}{27}(4+8-2-2) = \frac{2}{27}\cdot 8 = \frac{16}{27}$$

$$\boxed{P(X < 4) = \frac{16}{27} \approx 0.5926}$$

(b) Compute $P(3 \le X < 4)$:

$$P(3 \le X < 4) = \int_3^4 \frac{2(1+x)}{27}\,dx = \frac{2}{27}\int_3^4 (1+x)\,dx$$
$$= \frac{2}{27}\left[x + \frac{x^2}{2}\right]_3^4 = \frac{2}{27}\left(\left(4 + \frac{16}{2}\right) - \left(3 + \frac{9}{2}\right)\right)$$
$$= \frac{2}{27}(4 + 8 - 3 - 4.5) = \frac{2}{27}\cdot 4.5 = \frac{9}{27} = \frac{1}{3}$$

$$\boxed{P(3 \le X < 4) = \frac{1}{3} \approx 0.3333}$$

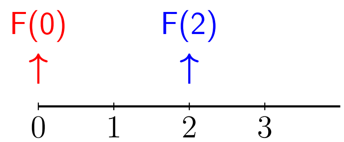
Discrete Random Variable Probability Rules

Rule 1: $P(X \leq a)$

Formula: $P(X \leq a) = F(a)$

Example: $P(X \leq 2) = F(2) = P(X = 0) + P(X = 1) + P(X = 2) = 0.875$

Visual:

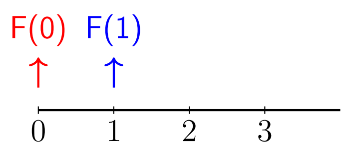


Rule 2: $P(X < a)$

Formula: $P(X < a) = F(a - 1)$

Example: $P(X < 2) = F(1) = P(X = 0) + P(X = 1) = 0.5$

Visual:

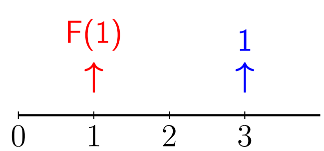


Rule 3: $P(X > a)$

Formula: $P(X > a) = 1 - F(a)$

Example: $P(X > 1) = 1 - F(1) = 0.5$

Visual:

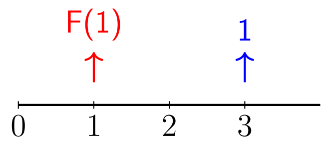


Rule 4: $P(X \geq a)$

Formula: $P(X \geq a) = 1 - F(a - 1)$

Example: $P(X \geq 2) = 1 - F(1) = 0.5$

Visual:

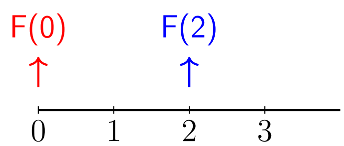


Rule 5: $P(a \leq X \leq b)$

Formula: $P(a \leq X \leq b) = F(b) - F(a-1)$

Example: $P(1 \leq X \leq 2) = F(2) - F(0) = 0.75$

Visual:

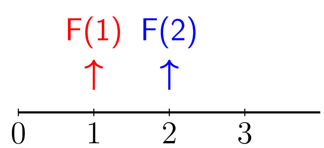


Rule 6: $P(a < X \leq b)$

Formula: $P(a < X \leq b) = F(b) - F(a)$

Example: $P(1 < X \leq 2) = F(2) - F(1) = 0.375$

Visual:

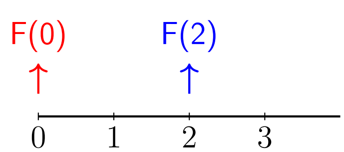


Rule 7: $P(a \leq X < b)$

Formula: $P(a \leq X < b) = F(b-1) - F(a-1)$

Example: $P(1 \leq X < 3) = F(2) - F(0) = 0.75$

Visual:

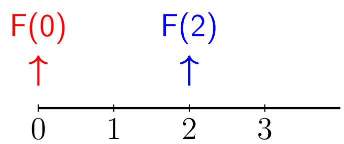


Rule 8: $P(a < X < b)$

Formula: $P(a < X < b) = F(b-1) - F(a)$

Example: $P(0 < X < 3) = F(2) - F(0) = 0.75$

Visual:

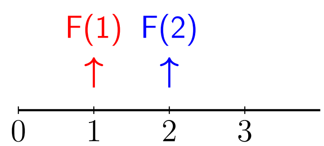


Rule 9: $P(X = a)$

Formula: $P(X = a) = F(a) - F(a-1)$

Example: $P(X = 2) = F(2) - F(1) = 0.375$

Visual:



Continuous Random Variable Probability Rules

Rule 1: CDF at a point

$$F(a) = P(X \leq a) = P(X < a)$$

- In continuous case, the probability at a single point is zero.
- So whether we write $X \leq a$ or $X < a$, both mean the same.
- Formally:

$$P(X = a) = \lim_{\epsilon \to 0} P(a \leq X \leq a + \epsilon) = 0$$

Rule 2: Right tail probability

$$P(X > a) = 1 - F(a)$$

- $F(a)$ already includes all probability up to a.
- Since the entire distribution sums to 1, the remaining area to the right is $1 - F(a)$.
- Geometrically: probability under density curve to the right of a.

Rule 3: Interval probability

$$P(a < X < b) = P(a \leq X \leq b) = F(b) - F(a)$$

- $F(b)$ counts probability up to b.
- $F(a)$ counts probability up to a.
- Subtraction gives probability between a and b.
- Since point probability is zero, choice of strict/weak inequalities doesn't matter.

4.4 Problems

Problem 47 *(a) Find the constant k that makes $f(x)$ a valid probability density function.*

(b) Compute the probability that the measurement error is less than $1/2$.

(c) It is considered unacceptable if the magnitude of the error exceeds 0.8. Determine the probability of this happening.

Problem 48 *A discrete random variable X takes values $1, 2, 3, 4$ with probabilities $p_X(1) = 0.2$, $p_X(2) = 0.3$, $p_X(3) = 0.1$, and $p_X(4) = 0.4$. Compute $P(X \leq 2)$ and $P(X > 3)$.*

Problem 49 *A continuous random variable Y has PDF*

$$f_Y(y) = \begin{cases} 3y^2, & 0 \le y \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

Find $P(0 \le Y \le 0.5)$ and $P(0.5 < Y \le 1)$.

Problem 50 *A discrete random variable X has PMF $p_X(x) = kx$ for $x = 1, 2, 3, 4$. Determine the value of k and compute $P(X \ge 3)$.*

Problem 51 *The PDF of a continuous random variable Z is*

$$f_Z(z) = \begin{cases} 2 - z, & 0 \le z \le 2, \\ 0, & \text{otherwise.} \end{cases}$$

Compute $P(0.5 \le Z \le 1.5)$ and $P(Z > 1)$.

Problem 52 *A discrete random variable X takes values $0, 1, 2, 3$ with PMF*

$$p_X(x) = \begin{cases} 0.1, & x = 0, \\ 0.2, & x = 1, \\ 0.4, & x = 2, \\ 0.3, & x = 3. \end{cases}$$

Compute $P(X \le 1)$ and $P(X > 2)$.

Problem 53 *A continuous random variable W has PDF*

$$f_W(w) = \begin{cases} 6w(1 - w), & 0 \le w \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

Find $P(0.25 \le W \le 0.75)$ and $P(W < 0.5)$.

4.5 Try it Yourself

Exercise 32 *Consider a family that uses a vacuum cleaner throughout the year. Let the total operating time X, measured in units of 100 hours, be a continuous random variable with density function*

$$f(x) = \begin{cases} x, & 0 < x < 1, \\ 2 - x, & 1 \le x < 2, \\ 0, & \text{otherwise.} \end{cases}$$

Determine the probabilities that in a year the vacuum cleaner is used:

(a) for less than 120 hours;

(b) between 50 and 100 hours.

Exercise 33 *The waiting time (in hours) between successive speeders observed by a radar is a continuous random variable X with cumulative distribution function*

$$F(x) = \begin{cases} 0, & x < 0, \\ 1 - e^{-8x}, & x \geq 0. \end{cases}$$

Calculate the probability that the waiting time is less than 12 minutes between successive speeders:

(a) using the CDF of X;

(b) using the corresponding PDF of X.

Exercise 34 *Let X be a continuous random variable with density*

$$f(x) = \begin{cases} k\sqrt{x}, & 0 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Determine the constant k that makes $f(x)$ a valid probability density function.

(b) Find the cumulative distribution function $F(x)$ and use it to compute $P(0.3 < X < 0.6)$.

Exercise 35 *A discrete random variable X takes values $1, 2, 3, 4, 5$ with PMF*

$$p_X(x) = \frac{x}{15}, \quad x = 1, 2, 3, 4, 5.$$

Compute $P(2 \leq X \leq 4)$ and $P(X$ is odd$)$.

Exercise 36 *A continuous random variable Y has PDF*

$$f_Y(y) = \begin{cases} c(1 - y^2), & -1 \leq y \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

1. *Determine the value of c.*
2. *Compute $P(-0.5 \leq Y \leq 0.5)$.*

Exercise 37 *A discrete random variable X takes values $0, 1, 2, 3$ with PMF*

$$p_X(x) = k(3 - x), \quad x = 0, 1, 2, 3.$$

1. *Determine the value of k.*
2. *Compute $P(X \geq 2)$.*
3. *Compute $P(1 \leq X \leq 3)$.*

4.6 YouTube Links and QR codes

Lecture	Details	YouTube Link	QR Code
18	Chapter 4.1: Random Variables, Discrete and Continuous	https://youtu.be/Xsor5zAOPVA	
19	Chapter 4.2: Overview of Discrete Probability Distribution and PMF	https://youtu.be/wlyMGM4kMn0	
20	Chapter 4.3: Overview of Continuous Probability Distribution and PDF	https://youtu.be/DMET1dE5tKo	

Chapter 5

Mean and Variance of a Random Variable

5.1 Mean

Expected Value - Discrete and Continuous Cases

Let X be a random variable with probability distribution $f(x)$. The **expected value** or **mean** of X is defined differently based on whether X is discrete or continuous:

- **Discrete case:**
$$\mu = \mathbb{E}(X) = \sum_x x f(x)$$
- **Continuous case:**
$$\mu = \mathbb{E}(X) = \int_{-\infty}^{\infty} x f(x)\, dx$$

Example - Discrete Expected Value

Let X be a discrete random variable with the following probability distribution:

x	1	2	3
$f(x)$	0.2	0.5	0.3

Then the expected value is:

$$\mathbb{E}(X) = 1 \cdot 0.2 + 2 \cdot 0.5 + 3 \cdot 0.3 = 0.2 + 1.0 + 0.9 = 2.1$$

Example - Continuous Expected Value

Let X be a continuous random variable with the probability density function:

$$f(x) = \begin{cases} 2x, & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

Then the expected value is:

$$\mathbb{E}(X) = \int_0^1 x \cdot 2x\,dx = \int_0^1 2x^2\,dx = \left[\frac{2x^3}{3}\right]_0^1 = \frac{2}{3}$$

Expected Commission Calculation

Problem:
A salesperson for a medical device company has two appointments on a given day.

- At the **first appointment**, there is a 70% chance ($P_1 = 0.7$) of making a deal with a commission of $1000.
- At the **second appointment**, there is a 40% chance ($P_2 = 0.4$) of making a deal with a commission of $1500.
- The two appointments are independent.

What is the salesperson's **expected total commission** based on their own probability belief?
Solution:
Let X_1 be the random variable representing the commission from the first appointment, and X_2 for the second.

$$\mathbb{E}[X_1] = (0.7)(1000) + (0.3)(0) = 700$$
$$\mathbb{E}[X_2] = (0.4)(1500) + (0.6)(0) = 600$$
$$\text{Total Expected Commission} = \mathbb{E}[X_1] + \mathbb{E}[X_2] = 700 + 600 = \boxed{1300}$$

So, the expected commission is **$1300**.

Expected Life of a Device

Problem:
Let X be the random variable that denotes the life in hours of a certain electronic device. The probability density function is

$$f(x) = \begin{cases} \frac{20,000}{x^3}, & x > 100 \\ 0, & \text{elsewhere} \end{cases}$$

Find the expected life of this type of device.
Solution:

The expected value (mean) of a continuous random variable X with PDF $f(x)$ is:

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x)\, dx$$

Since $f(x) = 0$ outside $x > 100$, we compute:

$$\mathbb{E}[X] = \int_{100}^{\infty} x \cdot \frac{20,000}{x^3}\, dx = \int_{100}^{\infty} \frac{20,000}{x^2}\, dx$$

Now compute the integral:

$$\begin{aligned}
\mathbb{E}[X] &= 20,000 \int_{100}^{\infty} x^{-2}\, dx \\
&= 20,000 \left[\frac{x^{-1}}{-1}\right]_{100}^{\infty} \\
&= 20,000 \left[0 - \left(-\frac{1}{100}\right)\right] \\
&= 20,000 \cdot \frac{1}{100} \\
&= \boxed{200}
\end{aligned}$$

Answer: The expected life of the device is **200 hours**.

Expected Value of a Function of a Random Variable

Let X be a random variable with probability distribution $f(x)$. The expected value of the random variable $g(X)$ is given by:

$$\mu_{g(X)} = \mathbb{E}[g(X)] = \sum_x g(x) f(x) \quad \text{if } X \text{ is discrete}$$

and

$$\mu_{g(X)} = \mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f(x)\, dx \quad \text{if } X \text{ is continuous.}$$

Example: Expected Value of a Function of a Random Variable

Consider a random variable X that denotes the number of cars arriving at a car wash in a day. The probability distribution of X is:

x	4	5	6	7	8	9
$P(X = x)$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$

Suppose the manager rewards the attendant according to the rule

$$g(X) = 2X - 1,$$

where $g(X)$ represents the amount (in dollars) paid.

Task: Compute the expected payment to the attendant during this one-hour period.
Solution: The expected reward is

$$\mathbb{E}[g(X)] = \sum_{x=4}^{9}(2x-1)\,P(X=x).$$

$$= (7)\left(\tfrac{1}{12}\right) + (9)\left(\tfrac{1}{12}\right) + (11)\left(\tfrac{1}{4}\right) + (13)\left(\tfrac{1}{4}\right) + (15)\left(\tfrac{1}{6}\right) + (17)\left(\tfrac{1}{6}\right).$$

$$= \tfrac{7+9}{12} + \tfrac{11+13}{4} + \tfrac{15+17}{6} = \tfrac{16}{12} + \tfrac{24}{4} + \tfrac{32}{6}.$$

$$= 1.33 + 6 + 5.33 = \boxed{\$12.67}.$$

Thus, on average, the attendant can expect to earn about \$12.67 during this period.

Example: Expected Value for a Continuous Random Variable

Let X be a random variable with density function

$$f(x) = \begin{cases} \frac{x^2}{3}, & -1 < x < 2 \\ 0, & \text{elsewhere} \end{cases}$$

Let $g(X) = 4X + 3$.
Find: The expected value of $g(X)$.
Solution: we have:

$$\mathbb{E}[g(X)] = \mathbb{E}(4X+3) = \int_{-1}^{2}(4x+3)\cdot\frac{x^2}{3}\,dx$$

$$= \frac{1}{3}\int_{-1}^{2}(4x^3+3x^2)\,dx = \frac{1}{3}\left[x^4+x^3\right]_{-1}^{2}$$

Now compute:

$$= \frac{1}{3}\left[(2^4+2^3) - ((-1)^4+(-1)^3)\right] = \frac{1}{3}[(16+8)-(1-1)] = \frac{1}{3}(24) = \boxed{8}$$

5.2 Variance

Variance and Standard Deviation

Let X be a random variable with probability distribution $f(x)$ and mean μ.
Variance of X is given by:

$$\sigma^2 = \mathbb{E}[(X-\mu)^2] = \sum_x (x-\mu)^2 f(x), \quad \text{if } X \text{ is discrete,}$$

$$\sigma^2 = \mathbb{E}[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx, \quad \text{if } X \text{ is continuous.}$$

Standard Deviation of X is the positive square root of the variance:

$$\sigma = \sqrt{\sigma^2}$$

Alternative Formula for Variance

The variance of a random variable X can also be computed using the formula:

$$\sigma^2 = \mathbb{E}(X^2) - \mu^2$$

where:

- $\mathbb{E}(X^2)$ is the expected value of X^2,
- $\mu = \mathbb{E}(X)$ is the mean of X,
- σ^2 is the variance of X.

Variance of a Function of a Random Variable

Let X be a random variable and $g(X)$ be a function of X. Then the variance of $g(X)$ is given by:

$$\sigma^2_{g(X)} = \mathbb{E}\left([g(X) - \mu_{g(X)}]^2\right)$$

where $\mu_{g(X)} = \mathbb{E}[g(X)]$ is the expected value of $g(X)$.

Variance of a Transformed Random Variable

Let X be a random variable with probability distribution $f(x)$. The variance of the random variable $g(X)$ is:

$$\sigma^2_{g(X)} = \mathbb{E}\left\{[g(X) - \mu_{g(X)}]^2\right\} = \sum_x [g(x) - \mu_{g(X)}]^2 f(x), \quad \text{if } X \text{ is discrete,}$$

$$\sigma^2_{g(X)} = \mathbb{E}\left\{[g(X) - \mu_{g(X)}]^2\right\} = \int_{-\infty}^{\infty} [g(x) - \mu_{g(X)}]^2 f(x)\,dx, \quad \text{if } X \text{ is continuous.}$$

Example 1: Standard Deviation of Discrete Random Variable

Let X be a random variable with the following probability distribution:

x	-2	3	5
$f(x)$	0.3	0.2	0.5

Step 1: Compute the mean $\mu = \mathbb{E}[X]$

$$\mu = (-2)(0.3) + 3(0.2) + 5(0.5) = -0.6 + 0.6 + 2.5 = 2.5$$

Step 2: Compute $\mathbb{E}[X^2]$

$$\mathbb{E}[X^2] = (-2)^2(0.3) + 3^2(0.2) + 5^2(0.5) = 4(0.3) + 9(0.2) + 25(0.5) = 1.2 + 1.8 + 12.5 = 15.5$$

Step 3: Compute variance and standard deviation

$$\sigma^2 = \mathbb{E}[X^2] - \mu^2 = 15.5 - (2.5)^2 = 15.5 - 6.25 = 9.25$$

$$\sigma = \sqrt{9.25} \approx 3.04$$

Example 2: Mean and Variance of Y

Let $Y = 3X - 2$, where X has the density function:

$$f(x) = \begin{cases} \frac{1}{4}e^{-x/4}, & x > 0 \\ 0, & \text{elsewhere} \end{cases}$$

Step 1: Mean and Variance of X (Exponential distribution)

$$\mathbb{E}[X] = 4, \quad \text{Var}(X) = 4^2 = 16$$

Step 2: Use linear transformation rules

$$\mathbb{E}[Y] = \mathbb{E}[3X - 2] = 3 \cdot \mathbb{E}[X] - 2 = 3(4) - 2 = 10$$

$$\text{Var}(Y) = \text{Var}(3X - 2) = 3^2 \cdot \text{Var}(X) = 9 \cdot 16 = 144$$

Example 3: Variance of g(X)

Given:

$$f(x) = \begin{cases} \frac{2(x+2)}{5}, & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases} \quad \text{and} \quad g(X) = 3X^2 + 4$$

Step 1: Compute $\mu_{g(X)} = \mathbb{E}[g(X)]$

$$\mathbb{E}[g(X)] = \int_0^1 (3x^2 + 4) \cdot \frac{2(x+2)}{5}\, dx$$

First simplify the integrand:

$$(3x^2 + 4)(2x + 4) = 6x^3 + 8x^2 + 24x + 32 \Rightarrow \frac{1}{5}(6x^3 + 8x^2 + 24x + 32)$$

$$\mu_{g(X)} = \frac{1}{5}\int_0^1 (6x^3 + 8x^2 + 24x + 32)\,dx$$
$$= \frac{1}{5}\left[\frac{6x^4}{4} + \frac{8x^3}{3} + \frac{24x^2}{2} + 32x\right]_0^1$$
$$= \frac{1}{5}\left(\frac{6}{4} + \frac{8}{3} + 12 + 32\right)$$
$$= \frac{1}{5}(1.5 + 2.\overline{6} + 12 + 32)$$
$$= \frac{48.\overline{166}}{5}$$
$$= \boxed{9.6334}$$

Step 2: Compute $\mathbb{E}[g(X)^2]$ (use expansion of $g(X)^2 = (3x^2+4)^2 = 9x^4 + 24x^2 + 16$)

$$\mathbb{E}[g(X)^2] = \int_0^1 (9x^4 + 24x^2 + 16)\cdot\frac{2(x+2)}{5}\,dx \quad \text{(you can expand, integrate term-wise similarly)}$$

Then compute:

$$\sigma^2_{g(X)} = \mathbb{E}[g(X)^2] - (\mu_{g(X)})^2$$

(This step left as exercise for integration; numerical computation follows same idea.)

5.3 Linear Combinations of Expectation and Variance

Linearity of Expectation

If a and b are constants, then:

$$E(aX+b) = aE(X) + b$$

Setting $a = 0$, we get $E(b) = b$.
Setting $b = 0$, we get $E(aX) = aE(X)$.

Example (Discrete)

Let X be a discrete random variable with $P(X=1) = 0.3$, $P(X=2) = 0.4$, $P(X=3) = 0.3$. Find $E(2X+3)$.
Solution:

$$E(X) = 1(0.3) + 2(0.4) + 3(0.3) = 2.0E(2X+3) = 2E(X) + 3 = 2(2.0) + 3 = 7$$

Example (Continuous)

Let X have pdf $f(x) = 2x$ for $0 < x < 1$. Find $E(5X-4)$.
Solution:

$$E(X) = \int_0^1 x\cdot 2x\,dx = \int_0^1 2x^2\,dx = \frac{2}{3}E(5X-4) = 5E(X) - 4 = 5\left(\frac{2}{3}\right) - 4 = \frac{10}{3} - 4 = -\frac{2}{3}$$

Expectation of Sum or Difference of Functions

The expected value of the sum or difference of two functions of a random variable X is the sum or difference of the expected values:

$$E[g(X) \pm h(X)] = E[g(X)] \pm E[h(X)]$$

Example (Continuous)

The weekly demand for a drink is $g(X) = X^2 + X - 2$, where X has the density function:

$$f(x) = \begin{cases} 2(x-1), & 1 < x < 2 \\ 0, & \text{elsewhere} \end{cases}$$

Solution: We compute:

$$E[g(X)] = \int_1^2 (x^2 + x - 2) \cdot 2(x-1)\, dx$$

Expand the integrand:

$$(x^2 + x - 2)(2x - 2) = 2x^3 - 2x^2 + 2x^2 - 2x - 4x + 4 = 2x^3 - 6x + 4$$

Now integrate:

$$\int_1^2 (2x^3 - 6x + 4)\, dx = \left[\frac{2x^4}{4} - 3x^2 + 4x\right]_1^2 = \left[\frac{2(16)}{4} - 3(4) + 8\right] - \left[\frac{2(1)}{4} - 3(1) + 4\right] = (8-12+8)-(0.5-3+4$$

$$\boxed{E[g(X)] = 2.5}$$

Expectation of Product of Independent Variables

If X and Y are independent random variables, then:

$$E(XY) = E(X) \cdot E(Y)$$

Example

Let X and Y be independent with:

$$E(X) = 2, \quad E(Y) = 5$$

Then:

$$E(XY) = 2 \cdot 5 = \boxed{10}$$

Covariance of Independent Variables

If X and Y are independent random variables, then:

$$\sigma_{XY} = \text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0$$

Example

Let X and Y be independent random variables such that:

$$E(X) = 1, \quad E(Y) = 3, \quad E(XY) = 3$$

Then:

$$\sigma_{XY} = E(XY) - E(X)E(Y) = 3 - (1)(3) = 0$$

$$\boxed{\text{Cov}(X, Y) = 0}$$

Linear Combinations of Expectation and Variance

Formula	**Condition / Example**
$E(aX + b) = aE(X) + b$	Any random variable X Example: $E(3X + 7) = 3E(X) + 7$
$E(aX + bY) = aE(X) + bE(Y)$	Any X, Y Example: $E(2X - 3Y) = 2E(X) - 3E(Y)$
$\text{Var}(aX + b) = a^2\text{Var}(X)$	Any X Example: $\text{Var}(5X + 2) = 25\text{Var}(X)$
$\text{Var}(aX + bY) = a^2\text{Var}(X) + b^2\text{Var}(Y) + 2ab\,\text{Cov}(X, Y)$	Any X, Y Example: $3X - 2Y \Rightarrow 9\text{Var}(X) + 4\text{Var}(Y) - 12\text{Cov}(X, Y)$
$\text{Var}(X \pm Y) = \text{Var}(X) + \text{Var}(Y)$	If X and Y independent Example: $\text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y)$
$\text{Cov}(X, Y) = 0$	If X and Y independent Example: $X \perp Y \Rightarrow \text{Cov}(X, Y) = 0$
$\rho_{XY} = \dfrac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$	Correlation coefficient Example: $\rho_{XY} = \dfrac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$

5.4 Problems

Problem 54 *The probability distribution of X, the number of scratches per 15 meters of a polished wooden plank, is given as follows:*

x	0	1	2	3	4
$f(x)$	0.35	0.40	0.18	0.06	0.01

Find the average (expected) number of scratches per 15 meters of the plank.

Problem 55 *A biased coin is such that heads is twice as likely as tails. The coin is tossed twice. Find the expected number of tails.*

Problem 56 *The density function of calibrated weights (in kilograms) of metallic rods is*

$$f(x) = \begin{cases} \frac{6}{1+x^2}, & 0 < x < 2 \\ 0, & \text{elsewhere} \end{cases}$$

Find the expected value of X.

Problem 57 *If a fruit vendor's profit, in units of \$2000, on a crate of apples is modeled as a random variable X with density*

$$f(x) = \begin{cases} 3(1-x^2), & 0 < x < 1 \\ 0, & \text{elsewhere} \end{cases}$$

find the expected profit per crate.

Problem 58 *A publishing company orders a variable number of new printers annually, depending on the maintenance record. Let X denote the number of printers ordered, with probability distribution:*

x	0	1	2	3
$f(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{2}$	0

If each printer costs \$1500 and a rebate of \$$30X^2$ is applied at year-end, find the expected annual expenditure.

Problem 59 *A continuous random variable X has the density function*

$$f(x) = \begin{cases} 2e^{-2x}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$$

Find the expected value of $g(X) = e^{X/2}$.

Problem 60 *A discrete random variable X takes values $1, 2, 3, 4$ with probabilities $p_X(1) = 0.1$, $p_X(2) = 0.2$, $p_X(3) = 0.3$, and $p_X(4) = 0.4$. Compute $E[X]$ and $Var(X)$.*

Problem 61 *A discrete random variable Y has PMF $p_Y(y) = ky$ for $y = 1, 2, 3, 4$.*

1. *Find the value of k.*

2. *Compute $E[Y]$ and Var(Y).*

Problem 62 *A continuous random variable Z has PDF*

$$f_Z(z) = \begin{cases} 2z, & 0 \leq z \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

Compute $E[Z]$ and Var(Z).

Problem 63 *A discrete random variable W takes values $0, 1, 2$ with probabilities $0.2, 0.5, 0.3$ respectively. Compute $E[W]$ and Var(W).*

Problem 64 *A continuous random variable V has PDF*

$$f_V(v) = \begin{cases} 3v^2, & 0 \leq v \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

Compute $E[V]$ and Var(V).

Problem 65 *A discrete random variable X has PMF $p_X(x) = \frac{1}{6}$ for $x = 1, 2, 3, 4, 5, 6$. Compute $E[X]$ and Var(X). (Hint: think of a fair die)*

5.5 GATE PYQs

Problem 66 *Consider a probability distribution given by the density function:*

$$P(x) = \begin{cases} Cx^2, & 1 \leq x \leq 4 \\ 0, & \text{otherwise} \end{cases}$$

*Find the value of $P(2 \leq x \leq 3)$, rounded off to three decimal places.**[GATE CSE 2021]***

Problem 67 *In an examination, a student can choose the order in which two questions (QuesA and QuesB) must be attempted.*

If the first question is answered wrong, the student gets zero marks.

If the first question is answered correctly and the second question is not answered correctly, the student gets the marks only for the first question.

If both the questions are answered correctly, the student gets the sum of the marks of the two questions.

The following table shows the probability of correctly answering a question and the marks of the question respectively.

Question	Probability of answering correctly	Marks
QuesA	0.8	10
QuesB	0.5	20

Assuming that the student always wants to maximize her expected marks in the examination, in which order should she attempt the questions and what is the expected marks for that order (assume that the questions are independent)?

A. *First QuesA and then QuesB. Expected marks 14.*

B. *First QuesB and then QuesA. Expected marks 14.*

C. *First QuesB and then QuesA. Expected marks 22.*

D. *First QuesA and then QuesB. Expected marks 16.*

[GATE CSE 2021]

Problem 68 *A probability density function $f(x)$ is defined on the interval $[a, 1]$ as:*

$$f(x) = \frac{1}{x^2}$$

and $f(x) = 0$ outside this interval. The value of a such that $f(x)$ is a valid probability density function is: ***[GATE CSE 2016]***

Problem 69 *Each of the nine words in the sentence "The quick brown fox jumps over the lazy og" is written on a separate piece of paper. These nine pieces of paper are kept in a box. One of the pieces is drawn at random from the box. The expected length of the word drawn is . (The answer should be rounded to one decimal place)* ***GATE CSE 2014***

Problem 70 *Consider a random variable X that takes values $+1$ and -1 with probability 0.5 each. The values of the cumulative distribution function $F(x)$ at $x = -1$ and $x = +1$ are*

A) 0 *and* 0.5

B) 0 *and* 1

C) 0.5 *and* 1

D) 0.25 *and* 0.75

GATE CSE 2012

Problem 71 *Consider a finite sequence of random values $X = [x_1, x_2, \ldots, x_n]$. Let μ_x be the mean and σ_x be the standard deviation of X. Let another finite sequence Y of equal length be derived from this as $y_i = ax_i + b$, where a and b are positive constants. Let μ_y be the mean and σ_y be the standard deviation of this sequence.*

Which one of the following statements is ***INCORRECT****?*

(A) Index position of mode of X in X is the same as the index position of mode of Y in Y.

(B) Index position of median of X in X is the same as the index position of median of Y in Y.

(C) $\mu_y = a\mu_x + b$

(D) $\sigma_y = a\sigma_x + b$

GATE CSE 2011

Problem 72 *If the difference between the expectation of the square of a random variable, $E[X^2]$, and the square of the expectation of the random variable, $(E[X])^2$, is denoted by R, then:*

$$R = E[X^2] - (E[X])^2$$

What is the nature of R?
Options:

(A) $R = 0$

(B) $R < 0$

(C) $R > 0$

(D) $R \geq 0$

GATE CSE 2011

5.6 Try it Yourself

Exercise 38 *A discrete random variable X takes values $1, 2, 3, 4$ with probabilities $p_X(1) = 0.1$, $p_X(2) = 0.2$, $p_X(3) = 0.3$, $p_X(4) = 0.4$. Compute $E[X^2]$ and Var(X).*

Exercise 39 *A discrete random variable Y has PMF $p_Y(y) = ky$ for $y = 1, 2, 3, 4$.*

1. *Find the value of k.*
2. *Compute $E[Y]$, Var(Y), and $E[2Y + 3]$.*

Exercise 40 *A continuous random variable Z has PDF*

$$f_Z(z) = \begin{cases} 2z, & 0 \leq z \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

Compute $E[Z]$, Var(Z), and $E[Z^2]$.

Exercise 41 *A discrete random variable W takes values $0, 1, 2, 3$ with probabilities $0.1, 0.4, 0.3, 0.2$ respectively. Let $g(W) = W^2 + 1$. Compute $E[g(W)]$ and Var$(g(W))$.*

Exercise 42 *A continuous random variable V has PDF*

$$f_V(v) = \begin{cases} 3v^2, & 0 \leq v \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

Compute $E[V]$, $E[V^2]$, and Var(V).

Exercise 43 *A discrete random variable X has PMF $p_X(x) = 1/6$ for $x = 1, 2, 3, 4, 5, 6$. Compute $E[X]$, Var(X), and $E[(X - 3)^2]$.*

Exercise 44 *A discrete random variable Y takes values $1, 2, 3$ with PMF $p_Y(1) = 0.2$, $p_Y(2) = 0.5$, $p_Y(3) = 0.3$. Let $g(Y) = 2Y + 5$. Compute $E[g(Y)]$ and Var$(g(Y))$.*

Exercise 45 *A continuous random variable Z has PDF*

$$f_Z(z) = \begin{cases} 4z^3, & 0 \le z \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

Compute $E[Z^2]$, $E[1 - Z]$, and Var(Z).

Exercise 46 *A discrete random variable X has PMF $p_X(x) = k(5 - x)$ for $x = 1, 2, 3, 4$.*

1. *Find k.*
2. *Compute $E[X]$, Var(X), and $E[X^3]$.*

Exercise 47 *A continuous random variable W has PDF*

$$f_W(w) = \begin{cases} 6w(1 - w), & 0 \le w \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

Compute $E[W]$, $E[(1 - W)^2]$, and Var(W).

5.7 YouTube Links and QR codes

Lecture	Details	YouTube Link	QR Code
24	Chapter 5.1: Expectation of Random Variable and Function of Random Variable	https://youtu.be/kik2x2gr5d0	
26	Chapter 5.2: Variance and Examples	https://youtu.be/Ja5R2qvpMMA	
27	Chapter 5.3 - 5.4: Linear Combinations — Solutions to Problems 54–59	https://youtu.be/HPucLgiZ7hU	
28	Chapter 5.5: Solutions to GATE PYQs (Solutions to Problems 66–72)	https://youtu.be/n6DxfJXLh0o	

Chapter 6

Discrete Probability Distributions

Discrete probability distributions are needed because they assign probabilities to each possible outcome of a random variable, helping us understand how likely different events are. For example, a binomial distribution describing the number of heads in 10 coin tosses tells us exactly how likely it is to get 0, 1, 2,. . . , or 10 heads, and allows us to calculate the expected number of heads and how much variation to expect. This turns random, unpredictable events into quantifiable knowledge, which can guide decisions and predictions in real life.

6.1 Bernoulli trials and Binomial distribution

An experiment often involves repeated trials, each resulting in one of two possible outcomes, commonly labeled as **success** or **failure**. A typical example is testing items on an assembly line, where each item is classified as either **defective** or **nondefective**. Either outcome can be defined as a **success**, depending on the context. This type of experiment is known as a **Bernoulli process**, and each individual trial is referred to as a **Bernoulli trial**.

It is important to note that for a sequence of trials to qualify as **Bernoulli trials**, the probability of success must remain constant from one trial to the next.

For instance, consider tossing a fair coin multiple times. The probability of getting a head (which we define as **success**) is:

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

This probability remains the same for each toss, satisfying the condition for **Bernoulli trials**.

However, if we consider drawing balls from a bag **without replacement**, the probability changes with each draw. For example, suppose a bag contains 3 red balls and 2 blue balls. The probability of drawing a red ball on the first draw is:

$$P(\text{Red on first draw}) = \frac{3}{5}$$

If a red ball is drawn and not replaced, the probability of drawing another red ball on the second draw becomes:

$$P(\text{Red on second draw}) = \frac{2}{4} = \frac{1}{2}$$

Since the probability changes from one trial to the next, this process does **not** meet the criteria for a **Bernoulli process**.

Properties of a Bernoulli Process

A Bernoulli process must satisfy the following conditions:

1. The experiment consists of repeated trials.
2. Each trial results in an outcome that may be classified as a **success** or a **failure**.
3. The probability of success, denoted by p, remains constant from trial to trial.
4. The repeated trials are independent.

Binomial Distribution

A **Bernoulli trial** results in:

- **Success** with probability p
- **Failure** with probability $q = 1 - p$

The probability distribution of the **Binomial random variable** X, representing the number of successes in n independent trials, is given by:

$$b(x; n, p) = \binom{n}{x} p^x (1-p)^{n-x}, \quad \text{for } x = 0, 1, 2, \ldots, n$$

Mean and Variance of Binomial Distribution

The mean and variance of the binomial distribution $b(x; n, p)$ are given by:

$$\mu = np \quad \text{and} \quad \sigma^2 = npq$$

where $q = 1 - p$.

Tips to Identify Binomial Distribution Problems

A probability distribution problem is likely a **Binomial Distribution** if the following conditions (and keywords) are present:

1. **Fixed number of trials (n)**: Look for phrases like:
   - "out of 10 attempts", "for 20 customers", "tested 5 items"
2. **Each trial has two possible outcomes (Success/Failure)**: Keywords may include:
   - "defective/non-defective", "pass/fail", "survive/die", "hit/miss"

3. **Constant probability of success (p) in each trial**: Phrases like:
   - "with probability 0.3", "each has a 60% chance"
4. **Independent trials**: Often stated or implied as:
   - "each trial is independent", "randomly chosen"
5. **Questions asking "exactly x successes", "at least x successes", or "between a and b successes"** often indicate binomial distribution.

Hint: If you're counting the number of times something happens in a fixed number of identical and independent trials with a fixed success probability, you're likely dealing with a **binomial distribution**.

Example 1: Tossing a Fair Coin

A fair coin is tossed 5 times. What is the probability of getting exactly 3 heads? Also compute the mean and variance of the distribution.
Solution:
Here, $n = 5, p = 0.5, q = 0.5$
Let X be the number of heads. Then,

$$P(X = 3) = \binom{5}{3}(0.5)^3(0.5)^2 = 10 \times 0.125 \times 0.25 = 0.3125$$

Mean: $\mu = np = 5 \times 0.5 = 2.5$
Variance: $\sigma^2 = npq = 5 \times 0.5 \times 0.5 = 1.25$

Example 2: Defective Bulbs

A machine produces bulbs with a 10% defect rate. If 8 bulbs are chosen at random, what is the probability that exactly 2 are defective? Also find the mean and variance.
Solution:
Here, $n = 8, p = 0.1, q = 0.9$

$$P(X = 2) = \binom{8}{2}(0.1)^2(0.9)^6 = 28 \times 0.01 \times 0.531441 = 0.1488$$

Mean: $\mu = np = 8 \times 0.1 = 0.8$
Variance: $\sigma^2 = npq = 8 \times 0.1 \times 0.9 = 0.72$

Example 3: Multiple Choice Test

A student guesses on a multiple choice question with 4 options. If there are 10 questions, what is the probability of getting exactly 4 correct answers by guessing?
Solution:

Here, $n = 10, p = \frac{1}{4} = 0.25, q = 0.75$

$$P(X = 4) = \binom{10}{4}(0.25)^4(0.75)^6 = 210 \times 0.00390625 \times 0.1779785 \approx 0.146$$

Mean: $\mu = np = 10 \times 0.25 = 2.5$
Variance: $\sigma^2 = npq = 10 \times 0.25 \times 0.75 = 1.875$

Example 4: Pass Rate in Exam

The probability of a student passing an exam is 0.6. If 7 students appear for the exam, find the probability that exactly 5 pass. Also calculate mean and variance.
Solution:
Here, $n = 7, p = 0.6, q = 0.4$

$$P(X = 5) = \binom{7}{5}(0.6)^5(0.4)^2 = 21 \times 0.07776 \times 0.16 = 0.2613$$

Mean: $\mu = np = 7 \times 0.6 = 4.2$
Variance: $\sigma^2 = npq = 7 \times 0.6 \times 0.4 = 1.68$

Example 5: Product Quality Check

In a quality control check, the probability that an item is of good quality is 0.85. If 12 items are tested, find the probability that exactly 10 are of good quality. Also find the mean and variance.
Solution:
Here, $n = 12, p = 0.85, q = 0.15$

$$P(X = 10) = \binom{12}{10}(0.85)^{10}(0.15)^2 = 66 \times 0.19687 \times 0.0225 \approx 0.292$$

Mean: $\mu = np = 12 \times 0.85 = 10.2$
Variance: $\sigma^2 = npq = 12 \times 0.85 \times 0.15 = 1.53$

Example 6

The probability that a student passes a final exam is 0.6. If 12 students take the exam, what is the probability that

(a) at least 8 pass,

(b) between 4 and 9 pass, and

(c) exactly 6 pass?

Solution:
Let X be the number of students who pass. Then X follows a binomial distribution with parameters $n = 12$, $p = 0.6$.

(a) Probability that at least 8 pass:

$$P(X \geq 8) = 1 - P(X \leq 7) = 1 - \sum_{x=0}^{7} b(x; 12, 0.6)$$

Using cumulative binomial values:

$$P(X \geq 8) = 1 - 0.7831 = \boxed{0.2169}$$

(b) Probability that between 4 and 9 pass:

$$P(4 \leq X \leq 9) = \sum_{x=4}^{9} b(x; 12, 0.6) = \sum_{x=0}^{9} b(x; 12, 0.6) - \sum_{x=0}^{3} b(x; 12, 0.6)$$

$$P(4 \leq X \leq 9) = 0.9450 - 0.0635 = \boxed{0.8815}$$

(c) Probability that exactly 6 pass:

$$P(X = 6) = b(6; 12, 0.6) = \sum_{x=0}^{6} b(x; 12, 0.6) - \sum_{x=0}^{5} b(x; 12, 0.6)$$

$$P(X = 6) = 0.6100 - 0.4070 = \boxed{0.2030}$$

6.2 Poisson Distribution

Experiments that result in **numerical values of a random variable** X, which counts the **number of occurrences** during a specific **time interval** or within a designated **region**, are referred to as **Poisson experiments**. The **time interval** can be of any duration — for instance, a **minute**, **hour**, **day**, **week**, **month**, or even a **year**.

Examples:

- X: Number of phone calls received per hour in an office.
- Number of school closures due to snow in a winter season.
- Number of baseball games postponed due to rain over a season.

The **region** could also be spatial:

- Number of field mice per acre.
- Number of bacteria in a culture.
- Number of typing errors per page.

Such problems are modeled using a **Poisson process**, and the experiment satisfies the following properties:

Properties of the Poisson Process

1. **Independence Across Disjoint Intervals or Regions:**
   The number of events occurring in one time interval or region is **independent** of those in any **non-overlapping** interval or region. The process has **no memory**.

2. **Proportionality:**
   The **probability** of a **single event** occurring in a very **small time interval** or **tiny region** is:

   - **Proportional** to the length/size of the interval or region.
   - **Independent** of events outside that interval or region.

3. **Negligible Multiple Events in Small Intervals:**
   The probability that **more than one event** occurs in a very **short interval** or **small region** is **negligible**.

Poisson Random Variable and Distribution

The random variable X, representing the **number of outcomes** in the given interval or region, is called a **Poisson random variable**. The associated probability distribution is known as the **Poisson distribution**.

Poisson Distribution

The probability distribution of a **Poisson random variable** X, representing the number of outcomes occurring in a given time interval or specified region of size t, is given by:

$$P(X = x; \lambda t) = \frac{e^{-\lambda t}(\lambda t)^x}{x!}, \quad x = 0, 1, 2, \ldots$$

Where:

- $\lambda =$ average number of occurrences per unit time, area, or volume.
- $t =$ time interval or region size.
- $\lambda t =$ expected number of events in interval t.
- $e \approx 2.71828.$

Mean and Variance:

Mean: $\mu = \lambda t$ Variance: $\sigma^2 = \lambda t$

Poisson Distribution Example

Problem: During a laboratory experiment, the average number of radioactive particles passing through a counter in 1 millisecond is 4. What is the probability that exactly 6 particles enter the counter in a given millisecond?

Solution:
This is a Poisson distribution problem with:

$$\lambda = 4, \quad x = 6$$

$$P(X = 6) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-4} \cdot 4^6}{6!} = \frac{e^{-4} \cdot 4096}{720}$$

$$e^{-4} \approx 0.0183 \quad \Rightarrow \quad P(X = 6) \approx \frac{0.0183 \cdot 4096}{720} \approx \frac{74.9568}{720} \approx 0.1041$$

Answer: The probability is approximately $\boxed{0.1041}$.

Tips to Identify Poisson Distribution in Word Problems

Look for the following keywords and features:

- **Fixed time/space/area/volume:** Mentions like "per minute", "per km", "per hour", "per unit", or "within an area".
- **Rare/Random Events:** Events that occur randomly, such as arrival of customers, system failures, or radioactive emissions.
- **Average rate is known:** Given as a mean number (e.g., "on average 3 per hour").
- **Discrete counts:** The variable counts how many events occur (not how long or how big).
- **Independence of events:** The occurrence of one event does not affect the others.

Common phrases in problems:

- "The average number of ..."
- "Occurs randomly over time/area"
- "Find the probability that exactly x events happen ..."
- "In a given time frame/region ..."

Key Formula:

$$P(X = x) = \frac{e^{-\lambda t}(\lambda t)^x}{x!}$$

Remember: Poisson is ideal when you're dealing with counts of rare events in a continuous domain.

Example 1: Emails Received

On average, a person receives 3 emails per hour. What is the probability that they receive exactly 5 emails in an hour?
Given: $\lambda = 3, x = 5$
Mean: $\mu = \lambda = 3$, **Variance:** $\sigma^2 = \lambda = 3$

$$P(X = 5) = \frac{e^{-3} \cdot 3^5}{5!} = \frac{e^{-3} \cdot 243}{120} \approx 0.1008$$

Example 2: Typos in a Page

An editor finds on average 2 typos per printed page. What is the probability that a page contains no typos?
Given: $\lambda = 2, x = 0$
Mean: $\mu = 2$, **Variance:** $\sigma^2 = 2$

$$P(X = 0) = \frac{e^{-2} \cdot 2^0}{0!} = e^{-2} \approx 0.1353$$

Example 3: Website Hits

A website gets 10 hits per minute on average. What is the probability of getting exactly 7 hits in a minute?
Given: $\lambda = 10, x = 7$
Mean: $\mu = 10$, **Variance:** $\sigma^2 = 10$

$$P(X = 7) = \frac{e^{-10} \cdot 10^7}{7!} = \frac{e^{-10} \cdot 10000000}{5040} \approx 0.0901$$

Example 4: Machine Breakdowns

A machine breaks down on average 1.5 times per month. What is the probability that it breaks down 3 times this month?
Given: $\lambda = 1.5, x = 3$
Mean: $\mu = 1.5$, **Variance:** $\sigma^2 = 1.5$

$$P(X = 3) = \frac{e^{-1.5} \cdot 1.5^3}{3!} = \frac{e^{-1.5} \cdot 3.375}{6} \approx 0.1255$$

Example 5: Car Accidents at a Junction

On average, 0.8 accidents occur at a junction daily. What is the probability that exactly 2 accidents occur tomorrow?
Given: $\lambda = 0.8, x = 2$
Mean: $\mu = 0.8$, **Variance:** $\sigma^2 = 0.8$

$$P(X = 2) = \frac{e^{-0.8} \cdot 0.8^2}{2!} = \frac{e^{-0.8} \cdot 0.64}{2} \approx 0.1144$$

6.3 Uniform Distribution

Discrete Uniform Distribution

A discrete uniform distribution has a finite number of outcomes that are equally likely.
PMF:

$$P(X = x) = \frac{1}{n}, \quad x = x_1, x_2, \ldots, x_n$$

Number of outcomes:

$$n = \frac{b-a}{d} + 1, \quad d = \text{spacing between consecutive numbers}$$

Mean:

$$E[X] = \frac{a+b}{2}$$

Variance:

$$\text{Var}(X) = \frac{(n^2-1)d^2}{12}$$

Tips to Identify Discrete Uniform Problems

- Problem states all outcomes are "equally likely"
- Finite number of possible outcomes
- Often includes phrases like:
  - "Choose a number at random from 1 to n"
  - "Fair die / coin / card deck"
  - "Equally probable outcomes"

Example 1: Die Roll Probability

Problem: A fair 12-sided die is rolled. Compute the expected value, variance, and the probability that the outcome is greater than 8.
Solution:

- Here $a = 1, b = 12, d = 1, n = 12$ (numbers 1 to 12)
- Mean: $E[X] = \frac{a+b}{2} = \frac{1+12}{2} = 6.5$
- Variance: Var$(X) = \frac{(n^2-1)d^2}{12} = \frac{(12^2-1)\cdot 1^2}{12} = \frac{143}{12} \approx 11.92$
- Probability $X > 8$: favorable outcomes $\{9, 10, 11, 12\} \Rightarrow 4$ out of 12

$$P(X > 8) = \frac{4}{12} = \frac{1}{3} \approx 0.333$$

Problem: A number is chosen randomly from integers 10 to 25. Compute the expected value, variance, and the probability that the number is ≤ 15.
Solution:

- $a = 10, b = 25, d = 1, n = b - a + 1 = 16$
- Mean: $E[X] = \frac{a+b}{2} = \frac{10+25}{2} = 17.5$
- Variance: $\text{Var}(X) = \frac{(n^2-1)d^2}{12} = \frac{(16^2-1)\cdot 1^2}{12} = \frac{255}{12} = 21.25$
- Probability $X \le 15$: numbers 10 to 15 $\Rightarrow 6$ out of 16

$$P(X \le 15) = \frac{6}{16} = \frac{3}{8} = 0.375$$

Example 3: Colored Card Draw

Problem: A deck contains 6 cards numbered 1 to 6. Compute the expected value, variance, and probability that the card number is > 4.
Solution:

- $a = 1, b = 6, d = 1, n = 6$
- Mean: $E[X] = \frac{1+6}{2} = 3.5$
- Variance: $\text{Var}(X) = \frac{(6^2-1)\cdot 1^2}{12} = \frac{35}{12} \approx 2.92$
- Probability $X > 4$: favorable outcomes $\{5, 6\} \Rightarrow 2/6$

$$P(X > 4) = \frac{2}{6} = \frac{1}{3} \approx 0.333$$

Example 4: Raffle Ticket Draw

Problem: A raffle ticket is randomly selected from tickets numbered 20 to 60. Compute expected value, variance, and probability ticket ≥ 50.
Solution:

- $a = 20, b = 60, d = 1, n = b - a + 1 = 41$
- Mean: $E[X] = \frac{20+60}{2} = 40$
- Variance: $\text{Var}(X) = \frac{(41^2-1)\cdot 1^2}{12} = \frac{1680}{12} = 140$
- Probability $X \ge 50$: numbers 50 to 60 $\Rightarrow 11/41$

$$P(X \ge 50) = \frac{11}{41} \approx 0.268$$

Example 5: Even Numbers Selection

Problem: A number is picked randomly from $\{2, 4, 6, 8, 10, 12\}$. Compute expected value, variance, probability < 7.
Solution:

- $a = 2, b = 12, d = 2, n = \frac{b-a}{d} + 1 = \frac{12-2}{2} + 1 = 6$
- Mean: $E[X] = \frac{2+12}{2} = 7$
- Variance: $\text{Var}(X) = \frac{(6^2-1)\cdot 2^2}{12} = \frac{35\cdot 4}{12} \approx 11.667$
- Probability $X < 7$: numbers 2,4,6 $\Rightarrow 3/6$

$$P(X < 7) = \frac{3}{6} = 0.5$$

6.4 Problems

Problem 73 *A radio telescope detects on average 12 cosmic rays per second. What is the probability that exactly 15 rays are detected in one second?*

Problem 74 *A drone has 10 propellers. Each propeller fails independently with probability 0.1. What is the probability that exactly 2 propellers fail?*

Problem 75 *A gamer plays 20 rounds of a guessing game with 4 possible moves per round. Choosing moves at random, what is the probability of guessing correctly in exactly 5 rounds?*

Problem 76 *A food delivery app receives an average of 5 orders per minute. What is the probability that it receives at most 3 orders in a given minute?*

Problem 77 *A student rolls a fair 50-sided die numbered 1 to 50. What is the probability that the result is divisible by 3 or 5?*

Problem 78 *The number of raindrops hitting a sensor per square centimeter of glass follows a Poisson distribution with mean 1.8. What is the probability that no raindrop hits in one square centimeter?*

Problem 79 *From a box of 100 numbered lottery tickets, 3 are drawn at random. What is the probability that a chosen ticket number lies between 41 and 60?*

Problem 80 *A message is broadcast to 8 satellites. Each satellite independently receives the signal with probability 0.25. What is the probability that at least 2 satellites receive it?*

Problem 81 *In DNA sequencing, each of 10,000 bases has a 0.0005 chance of being read incorrectly. What is the probability that exactly 3 errors occur?*

Problem 82 *A library has 10 new books, of which a student is familiar with 6. If 5 books are borrowed at random, what is the probability that the student knows at least 3 of them?*

Problem 83 *A robot makes on average 2.5 navigation errors per hour. What is the probability that it makes more than 3 errors in a given hour?*

Problem 84 *A shelf has 15 labeled jars (1 to 15). One jar is selected at random. Find the probability that the chosen label is a prime number.*

Problem 85 *A cluster contains 7 autonomous vehicles. Each vehicle independently malfunctions with probability 0.2. What is the probability that exactly 5 remain functional?*

Problem 86 *Fish enter a tank following a Poisson process with mean 3 per 2 minutes. What is the probability that no fish enters in the first minute?*

Problem 87 *A box contains 50 rechargeable batteries. Each has a 2% chance of being defective. What is the probability that exactly 1 defective battery is found?*

Problem 88 *A research kit contains 8 experimental resistors with values uniformly spaced from 100 to 180 ohms in steps of 10. What is the expected resistance value?*

Problem 89 *A coin is flipped 12 times in a simulation. What is the probability that the number of heads is between 4 and 8 (inclusive)?*

Problem 90 *An epidemic study records on average 1 new case every 4 hours. What is the probability that fewer than 2 cases are observed in an 8-hour period?*

Problem 91 *A ticket machine issues numbers from 1 to 30 uniformly at random. What is the variance of the distribution?*

Problem 92 *A space station has 20 sensors, each with a 0.05 probability of generating a false signal. What is the probability that at least one sensor generates a false signal?*

Problem 93 ***(GATE IT 2006)*** *A biased coin is tossed where the probability of landing Heads is p, with $0 < p < 1$. Let N represent the number of tosses until the first Head appears (including the toss where it appears). Assuming independence, the expected value of N is*

A. $\frac{1}{p}$

B. $\frac{1}{1-p}$

C. $\frac{1}{p^2}$

D. $\frac{1}{1-p^2}$

Problem 94 *The arrivals of trains at a station follow a Poisson distribution with mean $\lambda = 7$ per hour. (a) What is the probability that more than 10 trains arrive in a 2-hour interval? (b) What is the mean number of trains expected in 2 hours?*

6.5 Try it Yourself

Exercise 48 *A mountain observatory records meteors at an average rate of 4 meteors per hour. What is the probability that exactly 6 meteors are recorded in one hour?*

Exercise 49 *A data center has 12 hard disks. Each disk independently fails in a year with probability* 0.15*. What is the probability that exactly 3 disks fail this year?*

Exercise 50 *A quiz has 30 multiple-choice questions with 5 choices each. A contestant guesses every answer. What is the probability of getting exactly 7 correct answers?*

Exercise 51 *A café receives on average 8 customers in every 15-minute interval. What is the probability that at most 2 customers arrive in a particular 15-minute interval?*

Exercise 52 *An integer is chosen uniformly at random from* $\{1, 2, \ldots, 60\}$*. What is the probability that the chosen integer is divisible by* 4 *or* 6*?*

Exercise 53 *A textile inspection machine reports defects per metre that follow a Poisson distribution with mean* 0.9*. What is the probability that there are no defects in one metre?*

Exercise 54 *A raffle draws one ticket from tickets numbered* 1 *to* 80*. What is the probability the drawn ticket number lies between* 21 *and* 40 *(inclusive)?*

Exercise 55 *A marketing firm sends an SMS to 10 leads. Each lead independently responds with probability* 0.3*. What is the probability that at least 4 leads respond?*

Exercise 56 *A memory array has* 5000 *bits and each bit flips independently with probability* 0.001*. Using the Poisson approximation, estimate the probability that exactly 2 bits flip.*

Exercise 57 *An exam contains 12 questions. A student knows the answers to 7 of them. If 6 questions are selected at random for grading, what is the probability that the student knows at least 3 of those selected questions?*

6.6 YouTube Links and QR codes

Lecture	Details	YouTube Link	QR Code
29	Chapter 6.1: Binomial Distribution and Examples	https://youtu.be/LjVbhm2YTG0	
30	Chapter 6.2: Poisson Distribution and Examples	https://youtu.be/Pmn2NvEhXA8	
31	Chapter 6.3: Discrete Uniform Distribution and Examples	https://youtu.be/a8AyImPEkZ0	
32	Chapter 6.4: Solutions to Problems 73–94	https://youtu.be/BSfpHM7p1Cs	

Chapter 7

Continuous Probability Distributions

7.1 Uniform Distribution

What is Continuous Uniform Distribution?

A **continuous uniform distribution** is a probability distribution in which all values between two specified limits (a and b) are equally likely. The probability is uniformly spread over the interval $[a, b]$.

Probability Density Function (PDF):

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{otherwise} \end{cases}$$

Mean (Expected Value):

$$E(X) = \frac{a+b}{2}$$

Variance:

$$\text{Var}(X) = \frac{(b-a)^2}{12}$$

Cumulative Distribution Function (CDF):

$$F(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ 1, & x > b \end{cases}$$

How to Identify Continuous Uniform Distribution

- The variable is **continuous** and can take any real value within a specified range.
- Every value between a and b is **equally likely**.
- The problem may mention a **flat or constant probability density**.
- Usually used for time, distance, or measurement-based outcomes.

Keywords to look for:

- "equally likely between a and b"
- "randomly chosen between two continuous values"
- "uniformly distributed over an interval"

Examples

Example 1

A commuter arriving at a bus stop finds that the time between successive buss is uniformly distributed between 0 and 20 minutes. Let X be the waiting time (in minutes) from the commuter's arrival until the next bus. Find:

i. The PDF of X.

ii. The probability that the commuter waits more than 7 minutes.

iii. The expected waiting time and the variance.

Solution:
Since interarrival times are uniform on $[0, 20]$, we have $X \sim U(0, 20)$.

i. The PDF of a continuous uniform on $[a, b]$ is

$$f_X(x) = \begin{cases} \dfrac{1}{b-a}, & a \leq x \leq b, \\ 0, & \text{otherwise.} \end{cases}$$

Here $a = 0, b = 20$, so

$$f_X(x) = \frac{1}{20}, \quad 0 \leq x \leq 20.$$

ii. For $X \sim U(0, 20)$,

$$P(X > 7) = \frac{b-7}{b-a} = \frac{20-7}{20-0} = \frac{13}{20} = 0.65.$$

(Alternatively integrate: $\int_7^{20} \frac{1}{20}\, dx = \frac{13}{20}$.)

iii. The mean and variance of $U(a, b)$ are

$$E[X] = \frac{a+b}{2}, \qquad \text{Var}(X) = \frac{(b-a)^2}{12}.$$

Thus

$$E[X] = \frac{0+20}{2} = 10, \qquad \text{Var}(X) = \frac{20^2}{12} = \frac{400}{12} = \frac{100}{3} \approx 33.3333.$$

Example 2

A laboratory thermostat produces temperatures uniformly between $10°\text{C}$ and $50°\text{C}$ during a calibration process. Let X denote the recorded temperature. Compute:

i. The mean and variance of X.

ii. The probability that $X < 30°\text{C}$.

Solution:
Here $X \sim U(10, 50)$ with $a = 10$, $b = 50$.

i. Mean:

$$E[X] = \frac{a+b}{2} = \frac{10+50}{2} = 30.$$

Variance:

$$\text{Var}(X) = \frac{(b-a)^2}{12} = \frac{(50-10)^2}{12} = \frac{40^2}{12} = \frac{1600}{12} = \frac{400}{3} \approx 133.3333.$$

ii. Since the distribution is uniform,

$$P(X < 30) = \frac{30-10}{50-10} = \frac{20}{40} = \frac{1}{2} = 0.5.$$

(Equivalently, integrate $f_X(x) = 1/40$ from 10 to 30: $\int_{10}^{30} \frac{1}{40}\,dx = \frac{20}{40}$.)

Example 3

A customer support call lasts uniformly between 2 and 12 minutes. Let X denote the call duration (minutes). Find the probability that a randomly selected call lasts between 4 and 9 minutes. Also give the PDF and expected call duration.
Solution:
$X \sim U(2, 12)$ so $a = 2$, $b = 12$.

- PDF:

$$f_X(x) = \frac{1}{b-a} = \frac{1}{10}, \quad 2 \le x \le 12.$$

- Probability:

$$P(4 \le X \le 9) = \frac{9-4}{12-2} = \frac{5}{10} = 0.5.$$

(Integration: $\int_4^9 \frac{1}{10}\,dx = \frac{5}{10}$.)

- Expected value:

$$E[X] = \frac{2+12}{2} = 7 \text{ minutes.}$$

(Variance for completeness: $\mathrm{Var}(X) = \frac{10^2}{12} = \frac{100}{12} = \frac{25}{3} \approx 8.3333$.)

Example 4

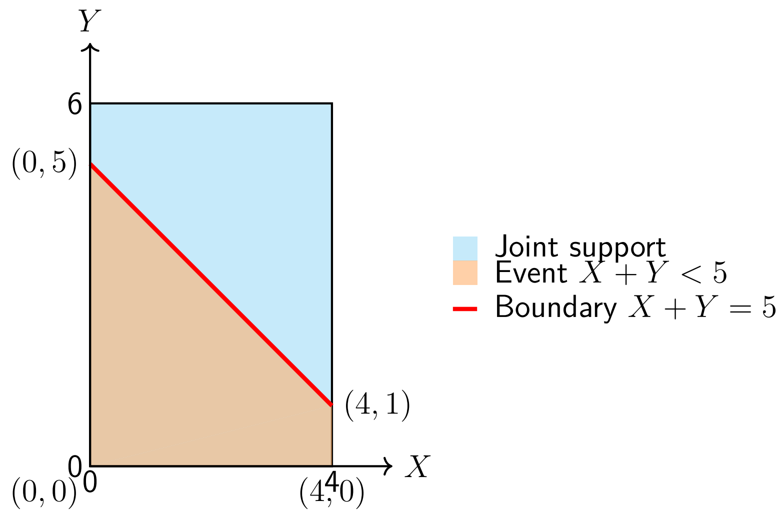
Let $X \sim U(0,4)$ and $Y \sim U(0,6)$, independent random variables. Find:

i. The probability that $X + Y < 5$.

ii. The probability that $X > Y$.

Solution: Both X and Y are uniform on $[0,4] \times [0,6]$, so total area $= 24$.

(i) $P(X+Y<5)$**:**



Area $= \int_0^4 (5-x)\,dx = 12 \Rightarrow P(X+Y<5) = \frac{12}{24} = \frac{1}{2}$.

(ii) $P(X>Y)$**:**
Favourable region: triangle below $x = 4$ and above $y = x$. Area $= \int_0^4 (4-y)\,dy = 8 \Rightarrow P(X > Y) = \frac{8}{24} = \frac{1}{3}$.
Final answers: (i) $\frac{1}{2}$, (ii) $\frac{1}{3}$.

Example 5

Let X be uniformly distributed on $[-2, 5]$. Find the cumulative distribution function $F_X(x)$.
Solution:

For $X \sim U(a, b)$ with $a = -2$, $b = 5$, the CDF is

$$F_X(x) = \begin{cases} 0, & x < a = -2, \\ \dfrac{x-a}{b-a} = \dfrac{x+2}{7}, & -2 \le x \le 5, \\ 1, & x > b = 5. \end{cases}$$

This piecewise formula gives, for example, $F_X(0) = \dfrac{0+2}{7} = \dfrac{2}{7}$.

Median: For a continuous uniform the median equals the mean:

$$\text{median} = \frac{a+b}{2} = \frac{-2+5}{2} = \frac{3}{2} = 1.5.$$

7.2 Exponential Distribution

Exponential Distribution Overview

The Exponential Distribution is a continuous probability distribution used to model the time between successive events in a Poisson process. It is memoryless, meaning the probability of an event occurring in the future is independent of the past.

Probability Density Function (PDF):

$$f(x; \lambda) = \lambda e^{-\lambda x}, \quad x \ge 0, \lambda > 0$$

Cumulative Distribution Function (CDF):

$$F(x) = 1 - e^{-\lambda x}$$

Mean (Expected Value):

$$E(X) = \frac{1}{\lambda}$$

Variance:

$$Var(X) = \frac{1}{\lambda^2}$$

Memoryless Property:

$$P(X > s + t \mid X > s) = P(X > t)$$

Tips to Identify Exponential Distribution

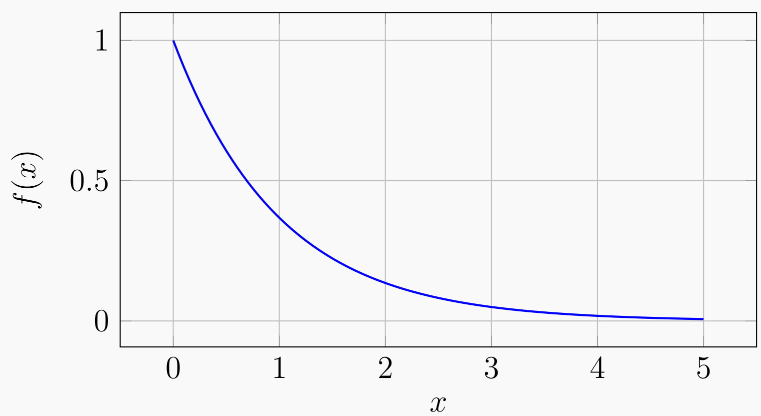
Look for the following keywords in the problem statement:

- "Time between arrivals"
- "Waiting time until the next event"
- "Lifetimes of devices or components"

- "Memoryless property"
- Events that happen **independently** and **continuously** over time

Such problems are often tied to Poisson processes, where the inter-arrival times are exponentially distributed.

Shape of Exponential Distribution



The PDF is highest at $x = 0$ and decreases exponentially.

Examples

Example 1

The average time between calls at a call center is 3 minutes. What is the probability that the next call comes in more than 5 minutes?
Solution: $\lambda = \frac{1}{3}$.
We use: $P(X > 5) = e^{-\lambda x} = e^{-5/3} \approx 0.1889$.

Example 2

The time to failure of a light bulb follows an exponential distribution with $\lambda = 0.02$. What is the expected life of the bulb?
Solution: $E(X) = \frac{1}{\lambda} = \frac{1}{0.02} = 50$ hours.

Example 3

If a system fails with a mean time of 10 hours, find the probability that it fails before 4 hours.
Solution: $\lambda = 1/10$.
$P(X < 4) = 1 - e^{-\lambda x} = 1 - e^{-0.4} \approx 0.3297$.

Example 4

A machine breaks down on average every 15 days. What is the variance in the time between breakdowns?
Solution: $\lambda = 1/15$, so: $Var(X) = \frac{1}{\lambda^2} = 225$ days^2.

Example 5

A component lasts more than 8 years with probability 0.2. What is the failure rate λ?
Solution: $P(X > 8) = e^{-\lambda \cdot 8} = 0.2$.
Taking $\ln$ on both sides: $-8\lambda = \ln(0.2) \Rightarrow \lambda = -\frac{\ln(0.2)}{8} \approx 0.2012$.

7.3 Normal Distribution and Standard Normal

Normal Distribution

The **Normal Distribution** is a continuous probability distribution that is symmetric about its mean. It models many natural phenomena such as height, test scores, and measurement errors.
Probability Density Function (PDF):

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

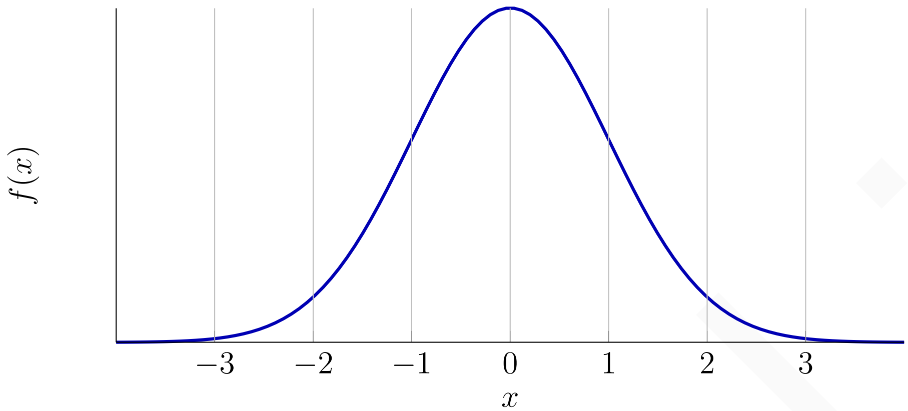
Where:

- $\mu =$ mean
- $\sigma =$ standard deviation
- $x =$ continuous random variable

Properties:

- Symmetric about the mean
- Mean = Median = Mode
- Total area under the curve = 1
- Empirical Rule:
  - 68% data within 1σ
  - 95% data within 2σ
  - 99.7% data within 3σ

Graph of Normal Distribution



Standard Normal Curve (Mean = 0, Std. Dev = 1)

Tips to Identify Normal Distribution Problems

Identification Tips

A problem likely involves a Normal Distribution if:

- It involves continuous data (e.g., heights, weights, time).
- You are given mean and standard deviation.
- The data is symmetric or bell-shaped.
- It mentions that data follows a "normal distribution" or "Gaussian distribution".
- It asks for probability between two values or more/less than a certain value.
- It involves a standardization step using Z-scores.

Standard Normal Distribution

Definition

The **Standard Normal Distribution** is a special case of the Normal Distribution where the mean (μ) is 0 and the standard deviation (σ) is 1.
A random variable Z follows a standard normal distribution if it has the following probability density function (pdf):

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < z < \infty$$

This distribution is symmetric about $z = 0$.

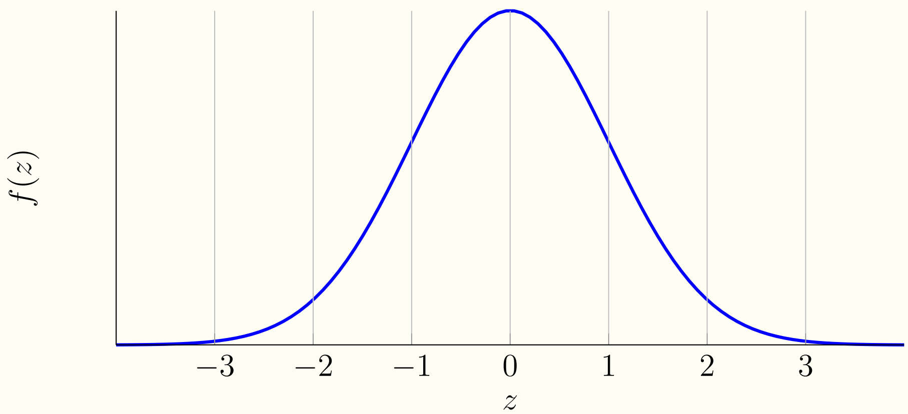
Z-Score Formula

To convert a normal variable $X \sim N(\mu, \sigma^2)$ to a standard normal variable Z, we use:

$$Z = \frac{X - \mu}{\sigma}$$

The Z-score tells us how many standard deviations a value is from the mean.

Graph of Standard Normal Distribution



When to Use Standard Normal Distribution

Use the Standard Normal Distribution when:

- You are dealing with a normally distributed variable and have been given the mean and standard deviation.
- You are asked to compute probabilities or quantiles using Z-scores.
- The problem involves symmetry around the mean.
- Values are standardized (i.e., Z-scores are used).

Examples

Example 1

The weights of apples in an orchard are normally distributed with a mean of 150 g and a standard deviation of 12 g. What is the probability that a randomly chosen apple weighs more than 165 g?

$$Z = \frac{165 - 150}{12} = 1.25$$

$$P(X > 165) = P(Z > 1.25) = 1 - 0.8944 = 0.1056$$

Answer: 10.56%

Example 2

The lifespan of LED bulbs follows a normal distribution with mean 2000 hours and standard deviation 200 hours. What is the probability that a bulb lasts between 1800 and 2200 hours?

$$Z_1 = \frac{1800 - 2000}{200} = -1, \quad Z_2 = \frac{2200 - 2000}{200} = 1$$

$$P(1800 < X < 2200) = P(-1 < Z < 1) = 0.8413 - 0.1587 = 0.6826$$

Answer: 68.26%

Example 3

The volume of milk in a carton is normally distributed with mean 1000 ml and standard deviation 8 ml. What percentage of cartons will contain more than 1015 ml?

$$Z = \frac{1015 - 1000}{8} = 1.875 \Rightarrow P(Z > 1.875) = 1 - 0.9699 = 0.0301$$

Answer: 3.01%

Example 4

Let $X \sim N(60, 16)$. Find the value x such that $P(X < x) = 0.975$.
From Z-tables, $P(Z < 1.96) = 0.975$

$$x = \mu + Z\sigma = 60 + 1.96 \cdot 4 = 60 + 7.84 = 67.84$$

Answer: 67.84

Example 5

The exam scores in a statistics class are normally distributed with mean 70 and standard deviation 6. What is the probability that a student scores between 65 and 80?

$$Z_1 = \frac{65 - 70}{6} = -0.83, \quad Z_2 = \frac{80 - 70}{6} = 1.67 \Rightarrow P = P(-0.83 < Z < 1.67)$$

$$= 0.9525 - 0.2033 = 0.7492$$

Answer: 74.92%

Example 6

If $X \sim N(45, 9)$, find $P(42 < X < 50)$.

$$Z_1 = \frac{42 - 45}{3} = -1, \quad Z_2 = \frac{50 - 45}{3} = 1.67$$

$$P = P(-1 < Z < 1.67) = 0.9525 - 0.1587 = 0.7938$$

Answer: 79.38%

Example 7

Find x such that $P(X < x) = 0.20$ for $X \sim N(100, 25)$.

$$P(Z < z) = 0.20 \Rightarrow z = -0.84$$
$$x = \mu + z\sigma = 100 - 0.84 \cdot 5 = 95.8$$

Answer: 95.8

Example 8

$X \sim N(500, 100^2)$. Find $P(X > 650)$.

$$Z = \frac{650 - 500}{100} = 1.5 \Rightarrow P(Z > 1.5) = 1 - 0.9332 = 0.0668$$

Answer: 6.68%

Example 9

For $X \sim N(0, 1)$, find $P(|X| < 1.5)$.

$$P(|X| < 1.5) = P(-1.5 < X < 1.5) = 0.9332 - 0.0668 = 0.8664$$

Answer: 86.64%

Example 10

The battery life of a smartphone is normally distributed with mean 8 hours and standard deviation 1.2 hours. What is the probability that a phone lasts more than 10 hours?

$$Z = \frac{10 - 8}{1.2} = 1.67 \Rightarrow P(Z > 1.67) = 1 - 0.9525 = 0.0475$$

Answer: 4.75%

7.4 Problems

Problem 95 *In a city, the heights of adult men follow a standard normal distribution with mean 0 and standard deviation 1. If the probability that a man's height is above 2.13 is p, the probability that a man's height is below -2.13 is ___.*

Problem 96 *The heights of adult men in a population follow a standard normal distribution with mean 0 and standard deviation 1. Determine the value k such that $\Phi(k) - 0.5 = 0.275$, where Φ is the standard normal CDF.*

Problem 97 *A factory produces rods with lengths normally distributed with mean 60 cm and standard deviation 12 cm. A new variable $Y = aX + b$ is defined so that Y has mean 0 and variance 1. The constants a and b are ___ and ___ respectively.*

Problem 98 *Rods manufactured by a machine have lengths normally distributed with mean 120 cm and standard deviation 8 cm. Two rods are selected independently. The probability that both rods are longer than 132 cm, given that* $\Phi(1.5) = 0.9332$*, is ___.*

Problem 99 *Scores of a quality test are normally distributed with mean 50 and standard deviation 10. Find two symmetric score limits around the mean such that the probability of scoring between them is 0.76, and the probability outside is equally split. Use* $\Phi(1.1503) = 0.874$*.*

Problem 100 *The lifetime of a rechargeable battery is normally distributed with mean 5 years and standard deviation 0.8 years. If a battery lasts more than 4 years, the probability that it lasts less than 6 years is ___.*

Problem 101 *Light bulbs have lifetimes normally distributed with mean 950 hours and standard deviation 60 hours. The probability that a randomly selected bulb lasts less than 890 hours or more than 1020 hours is ___.*

Problem 102 *Gears manufactured in a factory have diameters normally distributed with mean 5.0 cm and standard deviation 0.15 cm. The value* d *such that 90% of the gears have diameters within* $5 \pm d$ *cm (*$\Phi(1.645) = 0.95$*) is ___.*

Problem 103 *Resistors have resistances normally distributed with mean 100 ohms and standard deviation 5 ohms. If three resistors are selected independently, the probability that at least one exceeds 110 ohms (*$\Phi(2) = 0.9772$*) is ___.*

Problem 104 *Exam scores are normally distributed with mean 82 and standard deviation 6. Students whose scores lie in the middle 70% of the distribution are awarded a B grade. Using* $\Phi(1.036) = 0.85$*, the range of scores that correspond to a B grade is ___ to ___.*

Problem 105 *A juice machine produces daily output that is uniformly distributed between two unknown limits* a *and* b*. If* $P(X < 12) = 0.25$ *and it never produces more than 18 liters, find* a *and* b*.*

Problem 106 *A juice machine produces between 12 and 18 liters per day with equal probability. The probability that the daily production is between 13.5 and 17 liters is ___ and the probability that it is at least 15 liters is ___.*

Problem 107 *Consider two consecutive days of juice production from the same machine that produces between 12 and 18 liters per day with equal probability. The probability that the production exceeds 16 liters on both days is ___.*

Problem 108 *The lifetime of a component follows an exponential distribution with mean 2 years (*$\lambda = 0.5$*). The lifetime threshold* L *such that 70% of components last longer than* L *is ___.*

Problem 109 *Four independent components have lifetimes exponentially distributed with mean 2 years each (*$\lambda = 0.5$*). The probability that at least one component fails before the threshold* L *is ___.*

Problem 110 *Two independent rechargeable batteries have lifetimes normally distributed with mean 5 years and standard deviation 0.8 years. The expected lifetime of the first battery given that the sum of the lifetimes of both batteries is 12 years is ___.*

Problem 111 *A factory produces rods with lengths normally distributed with mean 120 cm and standard deviation 8 cm. Let* $Z = (X - 120)/8$. *The probability that* $Z^2 > 4$ *(*$\Phi(2) = 0.9772$*) is ___.*

Problem 112 *Two machines produce juice independently. The daily output of machine A is uniformly distributed between 12 and 18 liters, and machine B is uniformly distributed between 10 and 16 liters. The probability that the total output on a given day exceeds 25 liters is ___.*

Problem 113 *Two independent resistors have resistances normally distributed with mean 100 ohms and standard deviation 5 ohms. Let* Y *be the average resistance of the two resistors. The probability that* $Y > 102$ *ohms is ___.*

7.5 Try it Yourself

Exercise 58 *The daily production of widgets in a factory is normally distributed with mean 500 and standard deviation 25. Find the probability that a randomly chosen day produces more than 540 widgets.*

Exercise 59 *The weight of apples in a shipment is normally distributed with mean 150 g and standard deviation 20 g. Find the probability that an apple weighs between 130 g and 170 g.*

Exercise 60 *A machine fills sugar packets with amounts uniformly distributed between 950 g and 1050 g. What is the probability that a packet contains less than 980 g?*

Exercise 61 *The time taken by a cashier to process a customer is exponentially distributed with mean 3 minutes. Find the probability that a customer is processed in less than 2 minutes.*

Exercise 62 *Scores on a standardized test are normally distributed with mean 1200 and standard deviation 100. Find the score such that the top 10% of students exceed it.*

Exercise 63 *A random variable* X *follows a uniform distribution over* $[2, 8]$. *Find the probability that* X *is between 3 and 6.*

Exercise 64 *The lifetime of a light bulb is exponentially distributed with* $\lambda = 0.001$. *Find the probability that a bulb lasts more than 2000 hours.*

Exercise 65 *A batch of cookies has normally distributed weights with mean 50 g and standard deviation 5 g. What percentage of cookies weigh between 45 g and 55 g?*

Exercise 66 *The daily demand for a product is uniformly distributed between 30 and 70 units. What is the probability that demand on a given day is at least 60 units?*

Exercise 67 *A machine's component lifetimes follow an exponential distribution with mean 1000 hours. If 3 components are selected randomly, what is the probability that at least one lasts less than 800 hours?*

7.6 YouTube Links and QR codes

Lecture	Details	YouTube Link	QR Code
33	Chapter 7.1: Continuous Uniform Distribution and Examples	https://youtu.be/tZ8wBRTAq3w	
34	Chapter 7.2: Exponential Distribution and Examples	https://youtu.be/iQC4YbyrrrA	
35	Chapter 7.3: Normal and Standard Distribution and Examples	https://youtu.be/BOVw4KrDrSs	
37	Chapter 7.6: Solutions to Problems 95–104	https://youtu.be/9mxmhUoiKIQ	

38	Chapter 7.6: Solutions to Problems 105–113	https://youtu.be/uT8uy8ZuGpw	

Chapter 8

Solutions

Problems Covered	YouTube Link	QR Code
Solutions to Problems 1–5	https://youtu.be/Zcl7dzRsde8	
Solutions to Problems 6–10	https://youtu.be/gDD8iECtZRA	
Solutions to Problems 11–20	https://youtu.be/6P9hPGiYjbk	

Solutions to Problems 21–26	https://youtu.be/j0wNY7c2WEQ	
Solutions to Problems 27–46	https://youtu.be/eB5s4QW4NDc	
Solutions to Problems 54–59	https://youtu.be/HPucLgiZ7hU	
Solutions to GATE PYQs (Solutions to Problems 66–72)	https://youtu.be/n6DxfJXLh0o	
Solutions to Problems 73–94	https://youtu.be/BSfpHM7p1Cs	

Solutions to Problems 95–104	https://youtu.be/9mxmhUoiKIQ	
Solutions to Problems 105–113	https://youtu.be/uT8uy8ZuGpw	

Bibliography

[1] Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, and Keying E. Ye, *Probability & Statistics for Engineers& Scientists*, 9th Edition, Pearson Education, 2012.

[2] U. Dinesh Kumar, *Business Analytics: The Science of Data-Driven Decision Making*, First Edition, Wiley India, 2017.

Appendix

.1 Statistical Tables

Standard Normal Distribution Table (Z-Scores)
(Area to the left of the Z-score)

Z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
Z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.001	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.001	0.001
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.002	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.003	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.004	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.006	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048

Continued on next page

continued from previous page

Z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-2.4	0.0082	0.008	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.011
-2.1	0.0179	0.0174	0.017	0.0166	0.0162	0.0158	0.0154	0.015	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.025	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.063	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.102	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.123	0.121	0.119	0.117
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.166	0.1635	0.1611
-0.8	0.2119	0.209	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.242	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.305	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.281	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.33	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.352	0.3483

Continued on next page

continued from previous page

Z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-0.2	0.4207	0.4168	0.4129	0.409	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.0	0.5	0.496	0.492	0.488	0.484	0.4801	0.4761	0.4721	0.4681	0.4641
Z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633

Continued on next page

continued from previous page

Z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999

GateXAIML

Free GATE resources for Data Science & AI and CSE
Website: www.gatexaiml.in
Email: contact@gatexaiml.in