Statistics and Probabilities for Computing (CEN1602)

Unit-2

Probability

Dr. Ahmad Alzahrani

Content adapted from
Montgomery, D. C., & Runger, G. C. (2018).
Applied Statistics and Probability for Engineers (7th ed.). Wiley.

Outline

1. **Random Experiments and Sample Spaces**
   – Identifying possible outcomes
   – Discrete vs. continuous
2. **Events and Basic Probability Concepts**
   – Events, complements, unions, intersections
   – Interpreting probability values
3. **Axioms and Basic Rules of Probability**
   – Fundamental probability rules
   – Addition rule
4. **Counting Methods**
   – Multiplication rule
   – Permutations
   – combinations
5. **Conditional Probability**
   – Probability given information
   – Realistic computing examples
6. **Independence of Events**
   – Meaning of independence
   – Difference between independence and mutual exclusivity
7. **Bayes' Theorem**
   – Updating probabilities with evidence
   – Simple applications (e.g., alerts, classification intuition)
8. **Introduction to Random Variables (Bridge to Unit-3)**
   – Discrete vs. continuous random variables

Learning Objectives for Unit-2

After this unit, you should be able to:

1. **Define** random experiment, sample space, event.
2. **Compute** probabilities using basic rules (complement, union, intersection).
3. **Count** outcomes using multiplication / permutations / combinations.
4. **Use conditional probability** to study relationships among events**.**
5. **Decide independence** and use it to simplify calculations.
6. **Apply Bayes' Theorem** to find a conditional probability in one direction from the conditional probability in the opposite direction.
7. **Prepare for next unit:** understand what a random variable is.

Random Experiment

- A **random experiment** is an action that can be repeated but whose outcome is not known in advance.
- The random experiment has a known set of all **possible outcomes**.
- Each **trial** of the random experiment results in a single outcome.
- Repeating the same random experiment can lead to different results (or outcomes) each time.
- The random experiment is repeated in the same way each time.
- The possible outcomes are known, but the actual outcome is uncertain.
- **Probability** is used to study this **uncertainty** (**randomness**) .

Sample Spaces

- A **sample space** is the set of all possible outcomes of a random experiment.
- The sample space is usually denoted by $\boldsymbol{S}$.
- Each outcome of the experiment is included in the sample space.
- A **discrete sample space** has outcomes that can be listed.
- A **continuous sample space** has outcomes that lie in an interval of real numbers.

Example: Defining Sample Spaces

- We define the sample space using set notation:

$$S = \{all\ possible\ outcomes\ of\ the\ random\ experiment\}$$

Example 1: Consider a random experiment of tossing a coin **once** (discrete). The possible outcomes are Head (H) or Tail (T).

$$S = \{H, T\}$$

Example 2: Consider a random experiment of rolling a six-sided die once (discrete). The possible outcomes are the numbers 1 through 6.

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

Example 3: Consider a random experiment of measuring the time (in seconds) to complete a task (continuous). The outcome can be any positive real number.

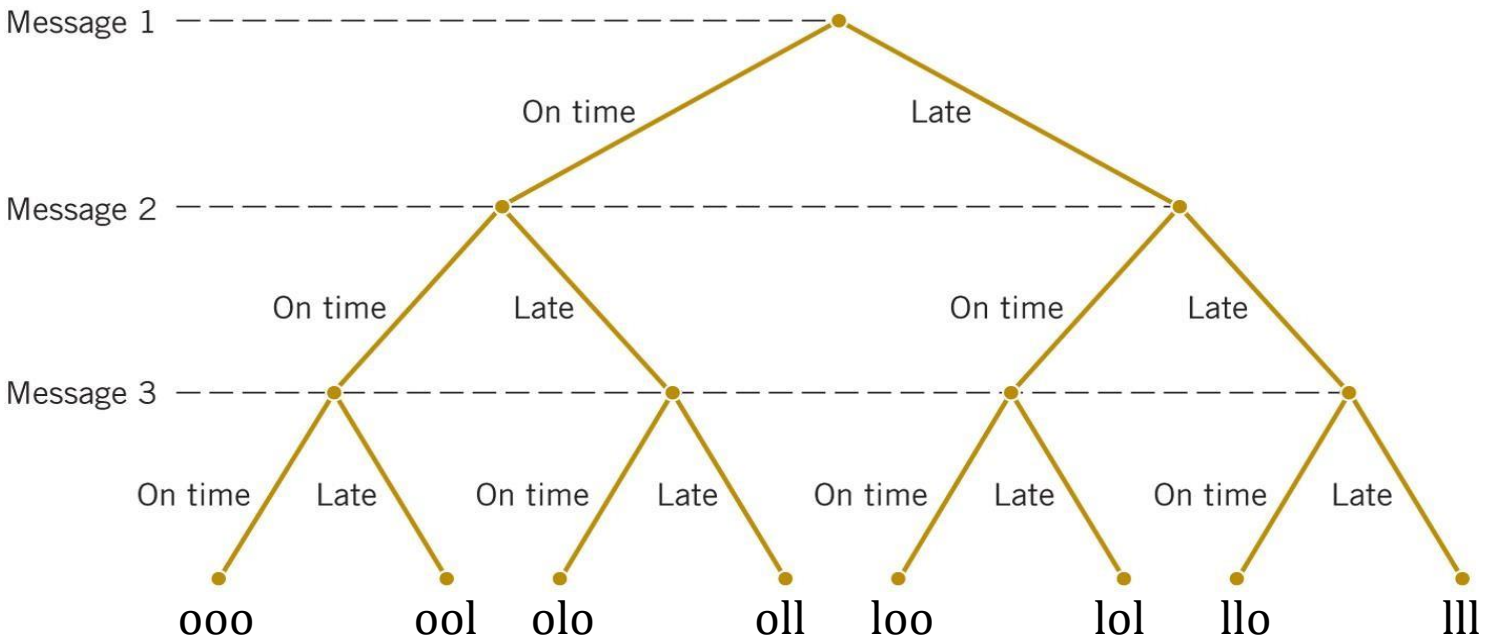
$$S = \{x \mid x > 0\}$$

Sample Space Defined By A Tree Diagram

- A tree diagram is a graphical way to describe all possible outcomes of a random experiment.
- Each complete path in the tree represents one possible outcome.

Example: Consider three messages sent in order. Each message is either on-time (o) or late (l). The sample space is:

$$S = \{ooo, ool, olo, oll, loo, lol, llo, lll\}$$



Events are Sets of Outcomes

- An **event (*E*)** is a subset of the sample space of a random experiment.
- Event combinations:
  1. The **Union** of two events consists of all outcomes that are contained in one event **or** the other, denoted as $E_1 \cup E_2$
  2. The **Intersection** of two events consists of all outcomes that are contained in one event **and** the other, denoted as $E_1 \cap E_2$
  3. The **Complement** of an event is the set of outcomes in the sample space that are **not** contained in the event, denoted as E^c (or E').

Example: Discrete Sample Space and Events

Suppose we check the connection status of **two devices** to a Wi-Fi network. Each device is either **connected (c)** or **not connected (n)**. The sample space is:

$$S = \{cc, cn, nc, nn\}$$

Definition of Events:

- Let E_1 be the event that at least one device is connected.
$$E_1 = \{cc, cn, nc\}$$
- Let E_2 be the event that no device is connected.
$$E_2 = \{nn\}$$
- Let E_3 be the event that at least one device is not connected.
$$E_3 = \{cn, nc, nn\}$$

Event Combinations:

- **Union:** $E_1 \cup E_2 = S$
- **Intersection:** $E_1 \cap E_3 = \{cn, nc\}$
- **Complement:** $E_1^c = \{nn\}$

Example: Events in a Continuous Sample Space

Suppose the response time (in milliseconds) of a computer system is measured for incoming requests. The response time is a continuous variable. The sample space is:

$$S = \{x \mid x > 0\}$$, basically $S = \mathrm{R}^{+}$ (positive real numers)

Definition of Events

- Let E_1 be the event that the response time is **between 100 ms and 300 ms**.

$$E_1 = \{x \mid 100 \leq x < 300\}$$

- Let E_2 be the event that the response time is **between 200 ms and 500 ms**.

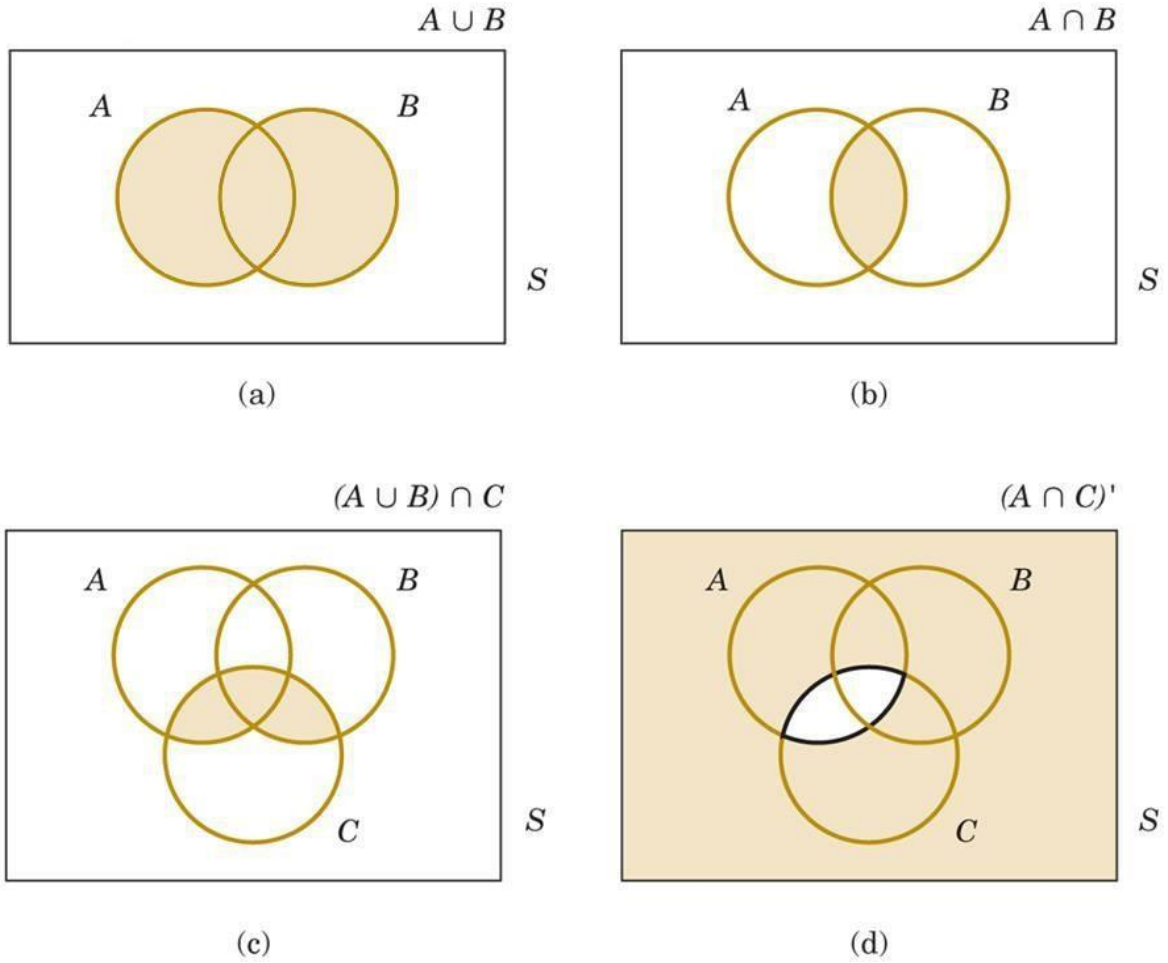
$$E_2 = \{x \mid 200 \leq x < 500\}$$

Event Operations

- **Union:** $E_1 \cup E_2 = \{x \mid 100 \leq x < 500\}$
- **Intersection:** $E_1 \cap E_2 = \{x \mid 200 \leq x < 300\}$
- **Complement of E_1:** $E_1^c = \{x \mid 0 < x < 100 \quad or \quad x \geq 300\}$

Venn Diagrams

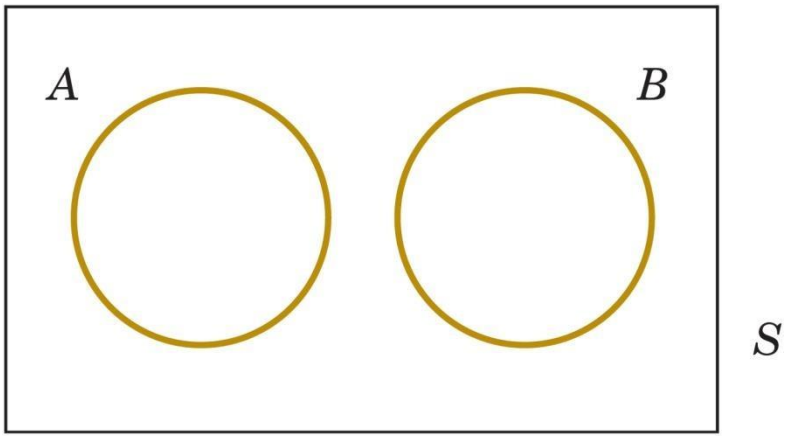
- Events *A* & *B* contain their respective outcomes. The shaded regions indicate the event relation of each diagram.



Mutually Exclusive Events

- Two events are **mutually exclusive** if they cannot occur at the same time.
- If one event occurs, the other cannot occur.
- Mutually exclusive events have no outcomes in common.
- The intersection of two mutually exclusive events is the empty set.
- Using symbols, if events A and B are mutually exclusive, then:

$$\boldsymbol{A} \cap \boldsymbol{B} = \emptyset$$



Mutually Exclusive Events - Laws

1. Commutative Law (order does not matter):

$$A \cup B = B \cup A \qquad \textit{and} \qquad A \cap B = B \cap A$$

2. Associative Law (grouping does not matter):

$$(A \cup B) \cup C = A \cup (B \cup C)$$
$$(A \cap B) \cap C = A \cap (B \cap C)$$

3. Distributive Law (like algebra):

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$
$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$

4. Complement Law:

$$(A^c)^c = A$$

5. De Morgan's Laws:

$$(A \cup B)^c = A^c \cap B^c$$
$$(A \cap B)^c = A^c \cup B^c$$

Probability

- **Probability** is a numerical measure of how likely an event from a random experiment is to occur.
- The probability is a number **between 0 and 1**.
- A probability of **0** means the event cannot occur.
- A probability of **1** means the event will certainly occur.
- Probabilities are assigned to events in the **sample space**.
- The probability of an event depends on the **outcomes included in that event**.
- In this unit, we consider only discrete (finite or countably infinite) sample spaces.

Types of Probability

1. Classical Probability

- Based on equally likely outcomes
- Used when all outcomes have the same chance of occurring.
- *Example*: Tossing a fair coin.

$$P(Head) = \frac{1}{2}$$

2. Empirical (Experimental) Probability

- Based on observed data or repeated experiments.
- Estimated using relative frequencies.
- *Example:* A server request is monitored over time.
  If 200 out of 1000 requests fail,

$$P(Failure) \approx \frac{200}{1000} = 0.2$$

3. Subjective Probability

- Based on judgment, experience, or belief.
- Used when data or equally likely outcomes are not available.
- *Example:* An engineer estimates there is a 70% chance that a system will meet a deadline, based on experience.

Classical Probability

- This type of probability is used when all outcomes have the same chance of occurring.
- Each outcome in the sample space is considered equally likely.
- The probability of an event (A) is calculated as:

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

- It is commonly used for theoretical models and simple experiments.
- Not applicable when outcomes are not equally likely.

Probability of an Event

- For a discrete sample space, the probability of an event E, denoted by $P(E)$,is the sum of the probabilities of the outcomes in E:

$$\boldsymbol{P(E) = \sum P(outcomes\ in\ E)}$$

Example: A computer system classifies each request into one of four types $S = \{\mathrm{A}, \mathrm{B}, \mathrm{C}, \mathrm{D}\}$, The probabilities are:

$$P(A) = 0.20,\ P(B) = 0.30,\ P(\mathrm{C}) = 0.40,\ P(\mathrm{D}) = 0.10$$

- Let E be the event that the request is **type B or C**: $E = \{\mathrm{B}, \mathrm{C}\}$
- Let F be the event that the request is **type C or D**: $F = \{\mathrm{C}, \mathrm{D}\}$

- $P(E \cup F) = P(\{B, C, D\}) = 0.30 + 0.40 + 0.10 = 0.80$
- $P(E \cap F) = p(\{C\}) = 0.40$
- $P(E^c) = P(\{A, D\}) = 0.20 + 0.10 = 0.30$

Axioms of Probability

- Probability is a number assigned to each event from a random experiment such that:

1. Non-negativity (Axiom I): The probability of any event, E, is a non-negative real number between 0 and 1.

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

2. Normalization (Axiom II): The probability of the entire sample space, S, is 1.

$$P(S) = 1$$

3. Additivity (Axiom III): If E_1 and E_2 are mutually exclusive events, then:

$$P(E_1 \cup E_2) = P(E_1) + P(E_2)$$

- **Immediate Consequences (from these axioms):**
  - $P(\emptyset) = 0$
  - $P(E^c) = 1 - P(E)$
  - If E_1 event is contained in E_2 , $then$ $P(E_1) \leq P(E_2)$

Addition Rules of Probability

- Joint events are formed by combining individual events using basic set operations, such as:

  1. **Union**: event A **or** event B occurs

  $$A \cup B$$

  2. **Intersection**: **both** events A and B occur

  $$A \cap B$$

  3. **Complement**: event A does **not** occur

  $$A^c$$

- The probability of a joint event depends on the probabilities of the individual events involved.
- Addition rules are used to compute probabilities of joint events, depending on whether the events **overlap or not**.

Example: Webpage User Sessions (Addition Rule)

- A user session is randomly selected from a website.
- Each session is classified by device type and login status.
- The table below summarizes the session data used in this example.

Session Type	Mobile	Desktop	Total
Logged in	180	120	**300**
Guest	140	160	**300**
Total	**320**	**280**	**600**

- Let:
  - M = session is from a mobile device
  - L = session is from a logged-in user

$$P(M) = \frac{320}{600}, \qquad P(L) = \frac{300}{600}, \qquad P(M \cap L) = \frac{180}{600}$$

Using the **addition rule**:

$$P(M \cup L) = P(M) + P(L) - P(M \cap L)$$

$$= \frac{320 + 300 - 180}{600} = \frac{440}{600} \approx 0.73$$

Probability of a Union (Addition Rule)

- For any **two** events A and B:

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

- If A and B are **mutually exclusive**:

$$A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$$

- For **three** events A, B, and C:

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

- If events $E_1, E_2, \ldots, E_k$ are **pairwise mutually exclusive**:

$$P(E_1 \cup E_2 \cup \cdots \cup E_k) = \sum_{i=1}^{k} P(E_i)$$

Counting Techniques

- Counting techniques are methods used to determine the number of possible outcomes in a sample space or event.
- They are applied when listing all outcomes is difficult or impractical.
- Three main counting techniques are used:
    1. **Counting Multiplication Rule**
    2. **Permutation Rule**
    3. **Combination Rule**
- Each technique is used in a different situation.
- Choosing the correct technique is important for correct results.

(1) Counting Multiplication Rule

- The **counting multiplication rule** is used when an experiment consists of several steps (stages). Assuming:
  - Step 1 can be done in n_1 ways,
  - Step 2 can be done in n_2 ways,
  - …
  - Step k can be done in n_k ways,
- Then the total number of possible outcomes is:

$$\boldsymbol{n_1 \times n_2 \times \cdots \times n_k}$$

Example: A computer system generates a one-time verification code using:

- 1 letter chosen from {A, B, C, D} ($\boldsymbol{n_1 = 4}$)
- 2 digits selected from {0, 1, 2,….,9} ($\boldsymbol{n_2 = 10^2 = 100}$)
- 1 symbol from {#, @} ($\boldsymbol{n_3 = 2}$)

How many different verification codes are possible?

Using the counting multiplication rule:

$$| S | = 4 \times 100 \times 2 = 800$$

Therefore, the sample space contains 800 possible verification codes.

(2) Permutation Rule

- A **permutation** is an **ordered arrangement** of items.
- **Order matters** in permutations.
- If there are n distinct items, the number of possible permutations of all items is:

$$n!$$

- The factorial notation is defined as:

$$n! = n \times (n-1) \times (n-2) \times \cdots \times 2 \times 1$$

- By this definition: $0! = 1$

Example: Three different tasks A, B, and C must be completed. The order in which the tasks are performed **matters**. How many different orders are possible?

Using the permutation rule:

$$\mathbf{3! = 3 \times 2 \times 1 = 6}$$

The possible orders are:

$ABC,\ ACB,\ BAC,\ BCA,\ CAB,\ CBA$

Subset Permutations

- A **subset permutation** counts the number of ways to arrange r items chosen from n distinct items.
- Order matters.
- The number of subset permutations is:

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

Example: How many different 3-digit codes can be formed using the 5 digits {1, 2, 3, 4, 5}, if no digit is repeated and the order of digits matters?

Apply the formula: $n = 5$ digits, $r = 3$ positions

$$P_3^5 = \frac{5!}{(5-3)!} = \frac{5!}{2!} = 60$$

Why order matters here

- **123 ≠ 321**
- Both use the same digits, but represent **different codes**

Similar Item Permutations

- Sometimes, the items being arranged are **not all distinct**.
- When some items are **identical**, fewer different permutations are possible.
- If there are n total items, with:
  - n_1 identical items of one type,
  - n_2 identical items of another type, and so on,
- The number of different permutations is:

$$\frac{n!}{n_1!\ n_2!\ \cdots}$$

Example: How many distinct binary codes of length 5 contain exactly three 1s and two 0s?

- The number of different arrangements is:

$$\frac{5!}{3!\,2!} = 10$$

(3) Combination Rule

- A **combination** is a selection of items where **order does not matter**.
- Combinations are used when we care only about **which items are selected**, not their order.
- If there are n **distinct items** and r **items are selected**, the number of combinations is:

$$C(n,r) = C_r^n = \binom{n}{r} = \frac{n!}{r!\,(n-r)!}$$

Example: A class has 10 students. 3 students are selected to form a project group. How many different project groups are possible?

- $n = 10$,$r = 3$

$$C_3^{10} = \binom{10}{3} = \frac{10!}{3!\ (10-3)!} = 120$$

There are 120 different possible groups.

Combinations: With and Without Replacement

1. Combinations Without Replacement:

- An item **cannot** be selected again once chosen.
- Each selection is from a smaller remaining set.
- This is the standard case for combinations (combinations = C_r^n).

Example:

A box contains 5 different types of candies. A student selects 3 candies. The order does not matter. How many different selections of 3 candies are possible?

$$\text{Number } \textit{of selections} = C_r^n = C_3^5 = 10$$

2. Combinations With Replacement:

- An item can be selected more than once.
- Each selection is from the same original set.
- This is a special case and uses a different counting approach.

Example:

Consider the same example, but now each candy is returned to the box before the next selection. The order of selection does not matter. How many different selections of 3 candies are possible?

$$\text{Number } \textit{of selections} = C_r^{n+r-1} = C_3^7 = 35$$

Why? 

Conditional Probability

- **Conditional probability** measures the likelihood of an event given that another event has already happened.
- The notation $\boldsymbol{P(B \mid A)}$ means *the probability that event B occurs, given that event A has* happened.
- We only focus on cases where A is true, and then check how often B happens.
- **Example:**

  If a user logs into a system. Let:
  - A: the username is correct.
  - B: the password is correct.

  $P(B \mid A)$ here is the *probability* that password is correct only if the username is correct.
- Conditional probability is fundamental in dependence, independence, and Bayes' theorem (will discuss shortly).

Conditional Probability Rule

- The conditional probability of an event *B* given an event *A*, denoted as $P(B \mid A)$, is given as:

$$\boldsymbol{P(B \mid A) = \frac{P(A \cap B)}{P(A)}}, \qquad P(A) > 0$$

- This rule can be used only if $P(A) > 0$.
- Simple interpretation to this:
  - First: check how often A happens.
  - Then: among those cases, see how often B also happens.
- Note: $P(B \mid A)$is not the same as $P(A \mid B)$.
- If knowing that A happened changes the value of $P(B)$,then events A and B are **dependent**.

Simple Practical Example: Email Spam Filter

An email system classifies (marks) 400 incoming emails as spam or not spam. Some emails are actually spam, and some are not. The email classification results are given in the table below.

	Actually Spam (S)	Not Spam (S')	Total
Classified as Spam (M)	30	10	40
Not Classified (M')	20	340	360
Total	50	350	**400**

Unconditional Probabilities:

- $P(M) = \frac{40}{400}$
- $P(M') = \frac{360}{400}$
- $P(S) = \frac{50}{400}$
- $P(S') = \frac{350}{400}$

Conditional Probabilities (Using the Rule):

- $P(S \mid M) = \frac{P(S \cap M)}{P(M)} = \frac{30/400}{40/400} = \frac{30}{40}$
- $P(S' \mid M) = \frac{P(S' \cap M)}{P(M)} = \frac{10/400}{40/400} = \frac{10}{40}$
- $P(S \mid M') = \frac{P(S \cap M')}{P(M')} = \frac{20/400}{360/400} = \frac{20}{360}$
- $P(S' \mid M') = \frac{P(S' \cap M')}{P(M')} = \frac{340/400}{360/400} = \frac{340}{360}$

Random Sampling and Conditional Probability

- Random sampling means each item has the same chance of being selected.
- In probability, selection is assumed to be random.
- When one item is selected first, probabilities for the next selection may change.
- Conditional probability is used to describe what happens **after** a previous outcome.
- When sampling without replacement, selected items are not returned.
- Probabilities can be computed without listing all outcomes.
- This idea leads directly to the probability multiplication rule.

Probability Multiplication Rule

- The Probability multiplication rule is used to find the probability that two events A and B both occur, or $P(A \cap B)$.
- This usually happens when events occur one after another.
- This rule tells us how to combine this probability with a conditional probability.
- This rule is given as:

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

- The order of the events can be reversed:

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

Example: Two-Step System Check

A system performs two checks before granting access. If:

- Probability the **first check passes** is 0.90
- Probability the **second check passes**, given the first passed, is 0.95

What is the probability that **both checks pass**?

Let

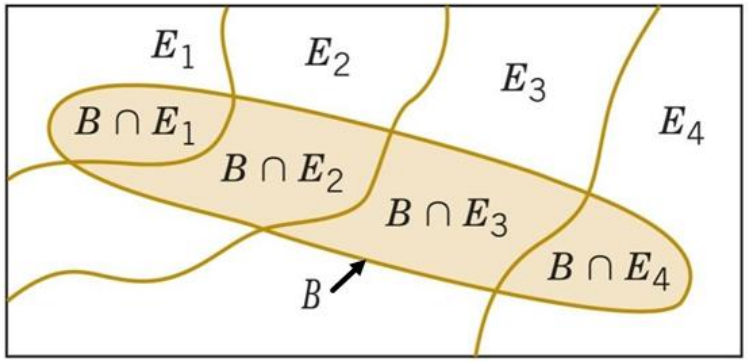
- A: first check passes, in this example $P(A) = 0.9$
- B: second check passes, in this example $P(B)$ is not needed.

Using the multiplication Rule:

$$P(both\ pass) = P(A \cap B) = P(A)P(B \mid A) = 0.90 \times 0.95 = 0.855$$

Total Probability Rule

- If events $E_1, E_2, \ldots, E_k$ cover the whole sample space *(exhaustive)*.
- This means that **only one** of these events can happen at a time *(mutually exclusive)*.
- If event B lies inside the sample space (as shown in the figure).



- From the figure, B is split into **separate, non-overlapping parts**:

$$B \cap E_1,\ B \cap E_2,\ \ldots,\ B \cap E_k$$

- All these parts together make up **all of** B.

- The probability of B is found by **adding the probabilities of these parts**.

$$P(B) = P(B \cap E_1) + P(B \cap E_2) + \cdots + P(B \cap E_k)$$

- Using conditional probability:

$$P(B) = P(B \mid E_1)P(E_1) + P(B \mid E_2)P(E_2) + \cdots + P(B \mid E_k)P(E_k)$$

Example: Cloud Server Request Failure

A cloud service processes user requests using **two** servers.

- A: the request is handled by Server 1
- A' :the request is handled by Server 2
- B: the request fails

Servers are used with different frequencies:

$$P(A) = 0.7 \quad \text{and} \quad P(A') = 0.3$$

Failure depends on which server is used:

$$P(B \mid A) = 0.02 \textit{ (failure rate on Server 1)}$$
$$P(B \mid A') = 0.08 \textit{ (failure rate on Server 2)}$$

What is the overall probability that a request fails?

Every request is handled by **exactly one server**, so:

$$P(B) = P(B \cap A) + P(B \cap A')$$
$$P(B) = P(B \mid A)P(A) + P(B \mid A')P(A')$$
$$P(B) = 0.02(0.7) + 0.08(0.3) = 0.014 + 0.024 = 0.038$$

Event Independence

- Two events are **independent** if knowing one event happens **does not change** the probability of the other.
- In independent events, the outcome of one event **has no effect** on the other.
- This means the conditional probability is the same as the original probability.

$$P(A \mid B) = P(A)$$
$$P(B \mid A) = P(B)$$

- An equivalent test for independence is:

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

- If this equation is not true, the events are dependent.
- For multiple events, independence means the multiplication rule holds for **every combination** of the events.

$$P(E_1 \cap E_2 \cap \cdots \cap E_r) = P(E_1)P(E_2)\cdots P(E_r)$$

Example: Software Error vs Network Delay

A monitoring system recorded **400 program runs**. Each run was classified by:

- S: software error occurred
- N: network delay occurred

The given table shows the frequency of software errors and network delays.

	Network Delay (N)	No Delay (N')	Total
Software Error (S)	20	60	80
No Error (S')	80	240	320
Total	100	300	**400**

Check independence of events S and N.

$$P(\mathrm{S}) = \frac{80}{400} = 0.20, \qquad P(\mathrm{N}) = \frac{100}{400} = 0.25, \qquad P(\mathrm{S} \cap \mathrm{N}) = \frac{20}{400} = 0.05$$

Now test:

$$P(\mathrm{S})P(\mathrm{N}) = 0.20 \times 0.25 = 0.05$$

Since:

$$P(\mathrm{S} \cap \mathrm{N}) = P(\mathrm{S})P(\mathrm{N})$$

Both events are **independent**. Thus, software errors and network delays occur independently in this system.

Bayes' Theorem

- We know how to compute $P(B \mid A)$: how likely an outcome (B) is after we observe a **condition** (A).
- **Bayes' Theorem** is used to **reverse** a conditional probability.
- It tells us how likely the condition is after we observe the outcome.
- Bayes' Theorem is given by:

$$\boldsymbol{P(A \mid B) = \frac{P(B \mid A)\, P(A)}{P(B)} \quad , P(B) > 0}$$

- Bayes' theorem is widely used in real systems to infer hidden conditions from observed outcomes.
- Example: when a spam filter flags an email (result), Bayes helps estimate how likely it is truly spam (condition).
- This allows the system to make more accurate filtering decisions.

Example: Spam Filter

An email system classifies messages as spam. Given that About 30% of emails are actually spam. If an email is spam, the filter flags it 80% of the time. Overall, 40% of emails are flagged by the filter.

Let:

- **Condition (A):** the email is spam
- **outcome (B):** the filter flags the email

Then:

$$P(A) = 0.3 \qquad P(B \mid A) = 0.8 \qquad P(B) = 0.4$$

Using Bayes' Theorem:

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)} = \frac{0.8 \times 0.3}{0.4} = 0.60$$

- After observing the result (flag), there is a 60% chance the email is actually spam (condition).
- The terms **prior** and **posterior** are commonly used with Bayes' theorem. In this example spam probability is called the **prior**, and the updated probability after observing the flag is called the **posterior**.

Bayes' Theorem with Total Probability

- **Bayes' theorem** can be combined with the **Total Probability Rule** to compute reverse conditional probabilities as:

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

Example: An email system classifies messages as spam. Given that:

- 30% of emails are spam
- If spam, the filter flags it **80%** of the time
- If not spam, the filter flags it **10%** of the time

If the filter flags an email, what is the probability that the email is actually spam?

Total probability of a flag

$$P(B) = 0.8(0.30) + 0.1(0.70) = 0.31$$

Apply Bayes

$$P(A \mid B) = \frac{0.80(0.30)}{0.31} \approx 0.77$$

If an email is flagged, there is about a 77% chance it is truly spam.

Random Variable and its Notation

- A **random variable** assigns a **numerical value** to each outcome of a random experiment.
- It allows us to work with outcomes using **numbers instead of descriptions**.
- Random variables are usually denoted by **capital letters** such as X, Y, or Z.
- The value taken by a random variable is written in **lowercase**, such as x.
- We write $P(X = x)$ to represent the probability that the variable equals a specific value.
- Random variables turn experiment outcomes into numbers we can analyze.

Discrete & Continuous Random Variables

A random variable can be **discrete** or **continuous**, depending on the type of values it can take.

Discrete Random Variable:

- Takes **countable values**.
- Values are usually integers.
- We can list all possible outcomes.

Examples:

- Number of login attempts before success
- Number of packets lost in a transmission
- Number of defective components

Continuous Random Variable:

- Takes **any value within an interval**.
- Values are measured, not counted.
- Infinitely many possible outcomes.

Examples:

- Time required to download a file
- CPU temperature
- Signal delay in a network

Core Ideas Summary for Unit-2

- A **random experiment** produces outcomes we cannot predict in advance.
- All possible outcomes form the **sample space**.
- An **outcome** is a single result ; an **event** is a group of outcomes.
- When outcomes are **equally likely**, probability is based on counting.
- The **axioms of probability** define the basic rules probabilities must follow.
- The **addition rule** adds probabilities to find the chance of events happening together or separately (∪).
- **Conditional probability** measures likelihood when information is known.
- The **probability multiplication rule** finds the probability of events happening together (∩).
- The **total probability rule** finds a total probability by adding intersection probabilities from different events.
- **Bayes' theorem** reverses conditional probabilities to find an unknown probability, $P(B|A) \rightarrow P(A|B)$.
- **Independence** describes events that do not affect each other.

Important Terms

Foundations

- Random experiment
- Outcome
- Sample space: discrete/continuous
- Event
- Equally likely outcomes
- Probability

Probability Rules

- Axioms of probability
- Addition rule
- Conditional probability
- Multiplication rule
- Total probability rule
- Bayes' theorem
- Independence
- Mutually exclusive events

Visualization

- Venn diagram
- Tree diagram

Counting & Sampling

- Permutation
- Combination:
  - With replacement
  - Without replacement

Random Variables

- Random variable: discrete / continuous