Experiment

- An experiment:

  a process or action that produces an outcome
- Example:

  Tossing a coin, rolling a die, running an algorithm, reaction of hydrogen and oxygen

Two types of Experiments

Deterministic

and

Random

Deterministic Experiment

- Outcome is certain and predictable
- Same input → same output
- No element of chance

Examples of Deterministic Experiments

- Arithmetic operations (e.g., 2 + 3 = 5)
- Running a sorting algorithm (input list determines exact sorted list)
- Compiling a program (source code → fixed machine code)

Random Experiment

- Outcome cannot be predicted with certainty
- Involves chance or probability
- Each run may give a different result

Examples of Random Experiments

- Tossing a coin
- Rolling a die
- Network packet arrival times
- Randomized algorithms (e.g., QuickSort pivot selection)

Key Differences

- Predictability: Deterministic → Always predictable | Random → Not predictable
- Input–Output Relation: Deterministic → Fixed | Random → Variable
- Example: Deterministic → Sorting algorithm | Random → Coin toss
- Use in CS: Deterministic → Compiler, algorithm analysis | Random → AI, simulations, cryptography

Why Important in Computer Science?

- Deterministic:

  Ensures reliability and correctness

- Random:

  Used in cryptography, simulations, randomized algorithms, machine learning

Real-World Computer Science Applications

- Deterministic:
  - Operating system scheduling (in some policies)
  - Hash functions (same input → same hash)
- Random:
  - Monte Carlo simulations
  - Randomized load balancing
  - AI training with random initialization

Quick Activity

- Which of the following are deterministic or random?

1. Execution time of a sorting algorithm
2. Output of SHA-256 hash function
3. Tossing a fair coin
4. Arrival time of emails in your inbox

Summary

- Deterministic: Predictable, repeatable, no chance
- Random: Unpredictable, chance-driven
- Both are fundamental in computer science: algorithms, simulations, cryptography, ML

Outcomes, events, and the sample space

- A collection of all elementary results, or outcomes of an experiment, is called a sample space.
- Any set of outcomes is an event. Thus, events are subsets of the sample space.

A sample space of N possible outcomes yields 2^N possible events.

- Example 2.5.:

Consider a football game between the Dallas Cowboys and the New York Giants.

- The sample space consists of 3 outcomes,
- S = { Cowboys win, Giants win, they tie }
- Combining these outcomes in all possible ways, we obtain the following 23 = 8 events:

Set Operations

- Discuss from section 2.1.2
- Union, Intersection, Complement, Difference of events
- Venn diagram and Tree diagram
- Mutually exclusive events, Exhaustive events

Sigma-algebra

A collection M of events is a sigma-algebra on sample space Ω if

(a) it includes the sample space, $\Omega \in M$

(b) every event in M is contained along with its complement; that is, $E \in M \Rightarrow E \in M$

(c) every finite or countable collection of events in M is contained along with its union; that is,

$E_1, E_2, \ldots \in M \Rightarrow E_1 \cup E_2 \cup \ldots \in M.$

Degenerate sigma-algebra

- By conditions (a) and (b) in the above definition every sigma-algebra has to contain the sample space Ω and the empty event $\emptyset$.
- This minimal collection $M = \{\Omega, \emptyset\}$ forms a sigma-algebra that is called degenerate.

Power set

- On the other extreme, what is the richest sigma-algebra on a sample space ? It is the collection of all the events,
- $M = 2^\Omega = \{E, E \subset \Omega\}$.
- As we know, there are 2^N events on a sample space of N outcomes. This explains the notation 2^Ω.
- This sigma-algebra is called a power set.

Borel sigma-algebra

- Now consider an experiment that consists of selecting a point on the real line.
- Then, each outcome is a point $x \in R$, and the sample space is $\Omega = R$.

Do we want to consider a probability that the point falls in a given interval?

- Define a sigma-algebra B to be a collection of all the intervals, finite and infinite, open and closed, and all their finite and countable unions and intersections.
- This sigma-algebra is very rich, but apparently, it is much less than the power set 2^ Ω.
- This is the Borel sigma-algebra.
- In fact, it consists of all the real sets that have length.

Axioms of probability

- Assume a sample space Ω and a sigma-algebra of events M on it.
- Probability $P : M \to [0, 1]$ is a function of events with the domain M and the range [0, 1] that satisfies the following two conditions,

Axioms of probability

- (Unit measure) The sample space has unit probability, $P(\Omega) = 1$.
- (Sigma-additivity) For any finite or countable collection of mutually exclusive events

  $E_1, E_2, \ldots \in M$,

  $P\{E_1 \cup E_2 \cup \ldots\} = P(E_1) + P(E_2) + \ldots.$

Extreme cases

- A sample space consists of all possible outcomes, therefore, it occurs for sure.
- On the contrary, an empty event $\emptyset$ never occurs.
- So, $P\{\Omega\} = 1$ and $P\{\emptyset\} = 0$

$P\{\Omega\} = 1$ and $P\{\emptyset\} = 0$.

- Proof: Probability of Ω is given by the definition of probability.
- By the same definition,

  $$P\{\Omega\} = P\{\Omega \cup \emptyset\} = P\{\Omega\} + P\{\emptyset\},$$

  because Ω and $\emptyset$ are mutually exclusive.
- Therefore, $P\{\emptyset\} = 0$.

Union

- Consider an event that consists of some finite or countable collection of mutually exclusive outcomes, $E = \{\omega_1, \omega_2, \omega_3, \ldots\}$.
- Summing probabilities of these outcomes, we obtain the probability of the entire event,

$$P\{E\} = \sum P\{\omega_k\} = P\{\omega_1\} + P\{\omega_2\} + P\{\omega_3\} + \ldots$$

Example 2.13.

- If a job sent to a printer appears first in line with probability 60%, and second in line with probability 30%,
- then with probability 90% it appears either first or second in line.
- It is crucial to notice that only mutually exclusive events (those with empty intersections) satisfy the sigma-additivity.
- If events intersect, their probabilities cannot be simply added.

Example 2.14.

- During some construction, a network blackout occurs on Monday with probability 0.7 and
- on Tuesday with probability 0.5.
- Then, does it appear on Monday or Tuesday with probability 0.7+0.5 = 1.2?
- Obviously not, because probability should always be between 0 and 1!
- Probabilities are not additive here because blackouts on Monday and Tuesday are not mutually exclusive.
- In other words, it is not impossible to see blackouts on both days.

Probability of a union

- $P\{A \cup B\} = P\{A\} + P\{B\} - P\{A \cap B\}$
- For mutually exclusive events,

  $P\{A \cup B\} = P\{A\} + P\{B\}$
- Generalization of this formula is not straightforward. For 3 events,

$P\{A \cup B \cup C\} = P\{A\} + P\{B\} + P\{C\} - P\{A \cap B\} - P\{A \cap C\} - P\{B \cap C\} + P\{A \cap B \cap C\}$

Example 2.15.

- In Example 2.14, suppose there is a probability 0.35 of experiencing network blackouts on both Monday and Tuesday.
- Then the probability of having a blackout on Monday or Tuesday equals

  $0.7 + 0.5 - 0.35 = 0.85.$

Complement rule

- Recall that events A and A^c are exhaustive, hence $A \cup A^c = \Omega$.
- Also, they are disjoint, hence

  $P\{A\} + P(A^c) = P(A \cup A^c) = P\{\Omega\} = 1.$
- Solving this for P(A), we obtain a rule that perfectly agrees with the common sense,

  $P(A) = 1 - P\{A^c\}.$

Example 2.16.

- A system appears protected against a new computer virus with probability 0.7.
- Then it is exposed to it with probability

  $1 - 0.7 = 0.3.$

Example 2.17.

- Suppose a computer code has no errors with probability 0.45.
- Then, it has at least one error with probability 0.55.

Intersection of independent events

- Events $E_1, \ldots, E_n$ are independent if they occur independently of each other.
- Occurrence of one event does not affect the probabilities of others.
- $P\{E_1 \cap \ldots \cap E_n\} = P\{E_1\} \cdot \ldots \cdot P\{E_n\}$

Example 2.18 (Reliability of backups)

- There is a 1% probability for a hard drive to crash. Therefore, it has two backups, each having a 2% probability to crash, and all three components are independent of each other. The stored information is lost only in an unfortunate situation when all three devices crash. What is the probability that the information is saved?

Solution:

- Denote the events, say,
- H = { hard drive crashes },

  B_1 = { first backup crashes } ,

  B2 = { second backup crashes } .
- It is given that H, B_1, and B_2 are independent,

  P {H} = 0.01, and

  P $\{B_1\}$ = P $\{B_2\}$ = 0.02.

Applying rules for complement and for intersection of independent events,

- $P\{\text{ saved }\} = 1 - P\{\text{ lost }\} = 1 - P\{H \cap B1 \cap B2\}$

  $= 1 - P\{H\}P\{B1\}P\{B2\} = 1 - (0.01)(0.02)(0.02)$

  $= 0.999996.$

- This is precisely the reason of having backups, isn't it?
- Without backups, the probability for information to be saved is only 0.99.

When computing probability for a system of several components to be functional

- When the system's components are connected in parallel, it is sufficient for at least one component to work in order for the whole system to function.
- Reliability of such a system is computed as in Example 2.18.
- Backups can always be considered as devices connected in parallel.

At the other end, consider a system whose components are connected in sequel.

- Failure of one component inevitably causes the whole system to fail.
- Such a system is more “vulnerable.”
- In order to function with a high probability, it needs each component to be reliable, as in the next example.

Example 2.19.

- Suppose that a shuttle's launch depends on three key devices that operate independently of each other and malfunction with probabilities 0.01, 0.02, and 0.02, respectively. If any of the key devices malfunctions, the launch will be postponed. Compute the probability for the shuttle to be launched on time, according to its schedule.

Solution

- $P\{\text{ on time }\} = P\{\text{ all devices function }\}$

  $= P\{ H^c \cap B_1^c\, 1 \cap B_2^c \}$

  $= P\{ H^c \} P\{B_1^c\} P\{B_2^c\}$ (independence)

$= (1 - 0.01)(1 - 0.02)(1 - 0.02)$ (complement rule)

  $= 0.9508.$

- Notice how with the same probabilities of individual components as in Example 2.18, the system's reliability decreased because the components were connected sequentially.

Many modern systems consist of a great number of devices connected in sequel and in parallel.

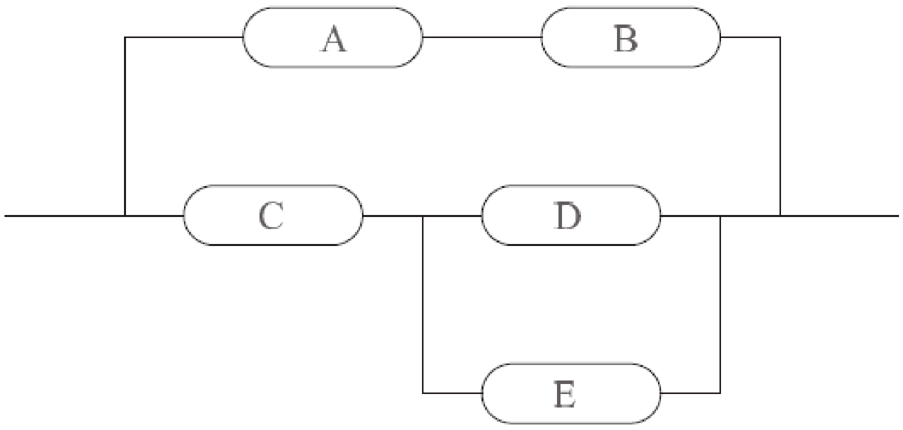

FIGURE 2.3: Calculate reliability of this system (Example 2.20).

Example 2.20 (Techniques for solving reliability problems).

- Calculate reliability of the system in Figure 2.3 if each component is operable with probability 0.92 independently of the other components.
- This problem can be simplified and solved “step by step.”

Solution.

- 1. The upper link A-B works if both A and B work, which has probability

  $P\{A \cap B\} = (0.92)^2 = 0.8464.$

- We can represent this link as one component F that operates with probability 0.8464.

2. By the same token,

- components D and E, connected in parallel, can be replaced by component G, operable with probability

  $P\{D \cup E\} = 1 - (1 - 0.92)^2 = 0.9936,$

as shown in Figure 2.4a.

- Note the formula,

  $P(D \cup E) = 1 - P(D \cup E)^c = 1 - \{ P(D^c) \cap P(E^c) \}$

Example 2.20 (Techniques for solving reliability problems).

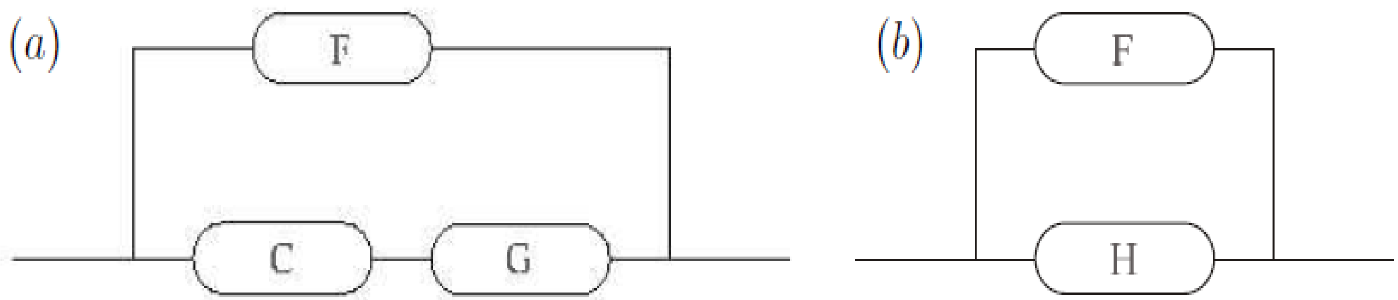

FIGURE 2.4: *Step by step solution of a system reliability problem.*

3. Components C and G, connected sequentially, can be replaced by component H, operable with probability

$$P\{C \cap G\} = (0.92)(0.9936) = 0.9141,$$

as shown in Figure 2.4b.

4. Last step. The system operates with probability

$$P\{F \cup H\} = 1 - (1 - 0.8464)(1 - 0.9141)$$

$$= 0.9868,$$

which is the final answer.

Equally likely outcomes

- When the sample space consists of n possible outcomes, $\omega_1, \ldots, \omega_n$, each having the same probability.
- Since $\sum_{1}^{n} P\{\omega_k\} = P\{\Omega\} = 1,$

we have in this case $P\{\omega_k\} = 1/n$ for all k.

A probability of any event E consisting of t outcomes, equals

$$\boldsymbol{P}\{E\} = \sum_{\omega_k \in E} \left(\frac{1}{n}\right) = t\left(\frac{1}{n}\right) = \frac{\text{number of outcomes in } E}{\text{number of outcomes in } \Omega}.$$

The outcomes forming event E are often called "favorable." Here index "F" means "favorable" and "T " means "total." Thus we have a formula

Equally likely outcomes

$$\boxed{\boldsymbol{P}\{E\} = \frac{\text{number of favorable outcomes}}{\text{total number of outcomes}} = \frac{\mathcal{N}_F}{\mathcal{N}_T}}$$

Example 2.21.

- Tossing a die results in 6 equally likely possible outcomes, identified by the number of dots from 1 to 6.
- We obtain, P {1} = 1/6,
- P { odd number of dots } = 3/6,
- P { less than 5 } = 4/6.

Example 2.22.

- A card is drawn from a bridge 52-card deck at random. Compute the probability that the selected card is a spade.
- **First solution.**
- The sample space consists of 52 equally likely outcomes—cards.
- Among them, there are 13 favorable outcomes—spades.
- Hence, P {spade} = 13/52 = 1/4.

Second solution

- The sample space consists of 4 equally likely outcomes—suits:
- clubs, diamonds, hearts, and spades.
- Among them, one outcome is favorable—spades.
- Hence, P {spade} = 1/4.

Example 2.23.

- A young family plans to have two children. What is the probability of two girls?
- **Solution 1 (wrong).**
- There are 3 possible families with 2 children: two girls, two boys, and one of each gender.
- Therefore, the probability of two girls is 1/3.

Solution 2 (right).

- Each child is (supposedly) equally likely to be a boy or a girl.
- Genders of the two children are (supposedly) independent. Therefore,
- P {two girls} =(1/2)(1/2) = 1/4.

In reality, business-related, sports-related, and political events are typically not equally likely.

- One outcome is usually more likely than another.
- For example, one team is always stronger than the other.
- Equally likely outcomes are usually associated with conditions of “a fair game” and “selected at random.”
- In fair gambling, all cards, all dots on a die, all numbers in a roulette are equally likely.

WR, WOR, Distinguishable, Indistinguishable

- Sampling with replacement means that every sampled item is replaced into the initial set
- Sampling without replacement means that every sampled item is removed from further sampling
- Objects are distinguishable if sampling of exactly the same objects in a different order yields a different outcome
- Indistinguishable objects arranged in a different order do not generate a new outcome.

Example 2.25 (Computer-generated passwords).

- When random passwords are generated, the order of characters is important because a different order yields a different password.
- Characters are distinguishable in this case.
- Further, if a password has to consist of different characters, they are sampled from the alphabet without replacement.

Example 2.26 (Polls).

- When a sample of people is selected to conduct a poll, the same participants produce the same responses regardless of their order.
- They can be considered indistinguishable.

Permutations with replacement

- Possible selections of k distinguishable objects from a set of n are called permutations.
- When we sample with replacement, each time there are n possible selections, and the total number of permutations is

Permutations with replacement

$$\boxed{P_r(n,k) = \overbrace{n \cdot n \cdot \ldots \cdot n}^{k \text{ terms}} = n^k}$$

Example 2.27 (Breaking passwords).

- From an alphabet consisting of 10 digits, 26 lower-case and 26 capital letters,
- one can create

  $P_r\ (62, 8) = 218{,}340{,}105{,}584{,}896$ (over 218 trillion) different 8-character passwords.
- At a speed of 1 million passwords per second, it will take a spy program almost 7 years to try all of them.
- Thus, on the average, it will guess your password in about 3.5 years.

At this speed, the spy program can test 604,800,000,000 passwords within 1 week.

- The probability that it guesses your password in 1 week is

$$\frac{\mathcal{N}_F}{\mathcal{N}_T} = \frac{\text{number of favorable outcomes}}{\text{total number of outcomes}} = \frac{604{,}800{,}000{,}000}{218{,}340{,}105{,}584{,}896} = 0.00277.$$

If capital letters are not used, the number of possible passwords is reduced to

- $P_r(36, 8) = 2{,}821{,}109{,}907{,}456$.
- On the average, it takes the spy only 16 days to guess such a password!
- The probability that it will happen in 1 week is 0.214.
- A wise recommendation to include all three types of characters in our passwords and to change them once a year is perfectly clear to us now...

Permutations without replacement

Permutations without replacement	$P(n,k) = \overbrace{n(n-1)(n-2)\cdot \ldots \cdot (n-k+1)}^{k \text{ terms}} = \dfrac{n!}{(n-k)!}$

- where n! = 1 · 2 · . . . n (n-factorial) denotes the product of all integers from 1 to n.
- The number of permutations without replacement also equals the number of possible allocations of k distinguishable objects among n available slots.

Example 2.28. In how many ways can 10 students be seated in a classroom with 15 chairs?

- Solution. Students are distinguishable, and each student has a separate seat.
- Thus, the number of possible allocations is the number of permutations without replacement,

$$P(15, 10) = 15 \cdot 14 \cdot \ldots \cdot 6 = (1.09) \cdot (10^{10}).$$

Combinations without replacement

- Possible selections of k indistinguishable objects from a set of n are called combinations.
- The total number of combinations is then

Combinations without replacement	$C(n,k) = \left(\begin{array}{c} n \\ k \end{array} \right) = \dfrac{P(n,k)}{P(k,k)} = \dfrac{n!}{k!(n-k)!}$

Example 2.29. An antivirus software reports that 3 folders out of 10 are infected. How many possibilities are there?

- Solution. Folders A, B, C and folders C, B, A represent the same outcome, thus, the order is not important.
- A software clearly detected 3 different folders, thus it is sampling without replacement.
- The number of possibilities is

$$\binom{10}{3} = \frac{10!}{3!\,7!} = \frac{10 \cdot 9 \cdot \ldots \cdot 1}{(3 \cdot 2 \cdot 1)(7 \cdot \ldots \cdot 1)} = \frac{10 \cdot 9 \cdot 8}{3 \cdot 2 \cdot 1} = 120.$$

Computational shortcuts

$$C(n,k) = \left(\begin{array}{c} n \\ k \end{array} \right) = \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{k \cdot (k-1) \cdot \ldots \cdot 1},$$

the top and the bottom of this fraction being products of k terms.

- It is handy to notice that

  C(n, k) = C(n, n – k) for any k and n

  C(n, 0) = 1

  C(n, 1) = n

Example 2.30.

- There are 20 computers in a store. Among them, 15 are brand new and 5 are refurbished. Six computers are purchased for a student lab. From the first look, they are indistinguishable, so the six computers are selected at random. Compute the probability that among the chosen computers, two are refurbished.

Solution.

- Compute the total number and the number of favorable outcomes.
- The total number of ways in which 6 computers are selected from 20 is

$$\mathcal{N}_T = \binom{20}{6} = \frac{20 \cdot 19 \cdot 18 \cdot 17 \cdot 16 \cdot 15}{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}.$$

For the number of favorable outcomes,

- 2 refurbished computers are selected from a total of 5, and the remaining 4 new ones are selected from a total of 15.
- There are

$$\mathcal{N}_F = \binom{5}{2}\binom{15}{4} = \left(\frac{5 \cdot 4}{2 \cdot 1}\right)\left(\frac{15 \cdot 14 \cdot 13 \cdot 12}{4 \cdot 3 \cdot 2 \cdot 1}\right)$$

favorable outcomes.

$$P\{\text{ two refurbished computers }\} = \frac{\mathcal{N}_F}{\mathcal{N}_T} = \frac{7 \cdot 13 \cdot 5}{19 \cdot 17 \cdot 4} = 0.3522.$$

Conditional probability and independence

- Conditional probability of event A given event B is the probability that A occurs when B is known to occur.

$$\underline{\text{Notation}} \quad \Big\| \quad P\{A \mid B\} \;=\; \text{conditional probability of } A \text{ given } B \quad \Big\|$$

Formula for computing conditional probability

- The unconditional probability of A,

$$P\{A\} = \frac{\text{number of outcomes in } A}{\text{number of outcomes in } \Omega}$$

- The conditional probability of A given B,

$$P\{A \mid B\} = \frac{\text{number of outcomes in } A \cap B}{\text{number of outcomes in } B} = \frac{P\{A \cap B\}}{P\{B\}}$$

- The general formula:

Conditional probability

$$\boxed{P\{A \mid B\} = \frac{P\{A \cap B\}}{P\{B\}}}$$

Intersection, general case

$$\boxed{P\{A \cap B\} = P\{B\}\, P\{A \mid B\}}$$

Independence

- Events A and B are independent if occurrence of B does not affect the probability of A, i.e.,

$$P\{A \mid B\} = P\{A\}.$$

- According to this definition, *conditional* probability equals *unconditional* probability in case of independent events.
- Substituting this into intersection general case yields,

$$P\{A \cap B\} = P\{A\}P\{B\}.$$

Example 2.31.

- Ninety percent of flights depart on time. Eighty percent of flights arrive on time. Seventy-five percent of flights depart on time and arrive on time.
  - (a) You are meeting a flight that departed on time. What is the probability that it will arrive on time?
  - (b) You have met a flight, and it arrived on time. What is the probability that it departed on time?
  - (c) Are the events, departing on time and arriving on time, independent?

Solution.

- Denote the events,

  A = {arriving on time} ,

  D = {departing on time} .

- We have:

  $P\{A\} = 0.8$, $P\{D\} = 0.9$, $P\{A \cap D\} = 0.75$.

Solution.

- (a) $P\{A \mid D\} = P\{A \cap D\}/P\{D\}$

  $= 0.75/0.9 = 0.8333$

- (b) $P\{D \mid A\} = P\{A \cap D\}/P\{A\}$

  $= 0.75/0.8 = 0.9375.$

- (c) Events are not independent because

  $P\{A \mid D\} \neq P\{A\}$, $P\{D \mid A\} \neq P\{D\}$,

  $P\{A \cap D\} \neq P\{A\}\, P\{D\}$.

Comment on this example

- Actually, any one of these inequalities is sufficient to prove that A and D are dependent.
- Further, we see that $P\{A \mid D\} > P\{A\}$ and $P\{D \mid A\} > P\{D\}$.
- In other words, departing on time increases the probability of arriving on time, and vice versa.
- This perfectly agrees with our intuition.

Bayes Rule

- Two conditional probabilities, P {A | B} and P {B | A}, are not the same, in general.

Example 2.32 (Reliability of a test).

- There exists a test for a certain viral infection (including a virus attack on a computer network).
- It is 95% reliable for infected patients and 99% reliable for the healthy ones.
- That is, if a patient has the virus (event V), the test shows that (event S) with probability

  $P\{S \mid V\} = 0.95$, and

  if the patient does not have the virus, the test shows that with probability

  $P(S^c \mid V^c) = 0.99$.

Consider a patient whose test result is positive (i.e., the test shows that the patient has the virus).

- Knowing that sometimes the test is wrong, naturally, the patient is eager to know the probability that he or she indeed has the virus.
- However, this conditional probability, $P\{V \mid S\}$, is not stated among the given characteristics of this test.

The problem is to connect the given P {S | V } and the quantity in question, P {V | S}.

- This was done in the eighteenth century by English minister Thomas Bayes (1702–1761) in the following way.
- Notice that A ∩ B = B ∩ A.
- Therefore, using multiplication rule,

  $P\{B\}\, P\{A \mid B\} = P\{A\}\, P\{B \mid A\}$.
- Solve for P {B | A} to obtain

Bayes Rule

$$\boxed{P\{B \mid A\} = \frac{P\{A \mid B\}\, P\{B\}}{P\{A\}}}$$

Example 2.33 (Situation on a midterm exam).

- On a midterm exam, students X, Y, and Z forgot to sign their papers. Professor knows that they can write a good exam with probabilities 0.8, 0.7, and 0.5, respectively. After the grading, he notices that two unsigned exams are good and one is bad. Given this information, and assuming that students worked independently of each other, what is the probability that the bad exam belongs to student Z?

Solution.

- Denote good and bad exams by G and B.
- Also, let GGB denote two good and one bad exams,
- XG denote the event "student X wrote a good exam," etc.
- We need to find $P\{ZB \mid GGB\}$
- Given that $P\{G \mid X\} = 0.8$, $P\{G \mid Y\} = 0.7$, and $P\{G \mid Z\} = 0.5$.

By the Bayes Rule,

$$\boldsymbol{P}\{ZB \mid GGB\} = \frac{\boldsymbol{P}\{GGB \mid ZB\}\,\boldsymbol{P}\{ZB\}}{\boldsymbol{P}\{GGB\}}.$$

- Given ZB, event GGB occurs only when both X and Y write good exams.
- Thus, P {GGB | ZB} = (0.8)(0.7).
- Event GGB consists of three outcomes depending on the student who wrote the bad exam.

Adding their probabilities, we get

- $P\{GGB\} = P\{XG \cap YG \cap ZB\} + P\{XG \cap YB \cap ZG\} + P\{XB \cap YG \cap ZG\}$

  $= (0.8)(0.7)(0.5) + (0.8)(0.3)(0.5) + (0.2)(0.7)(0.5)$

  $= 0.47.$

- Then $P\{ZB \mid GGB\} = (0.8)(0.7)(0.5)/0.47$

  $= 0.5957.$

Law of Total Probability

- This law relates the unconditional probability of an event A with its conditional probabilities.
- It is used every time when it is easier to compute conditional probabilities of A given additional information.
- Consider some partition of the sample space with mutually exclusive and exhaustive events $B_1, \ldots, B_k$.

Law of Total Probability

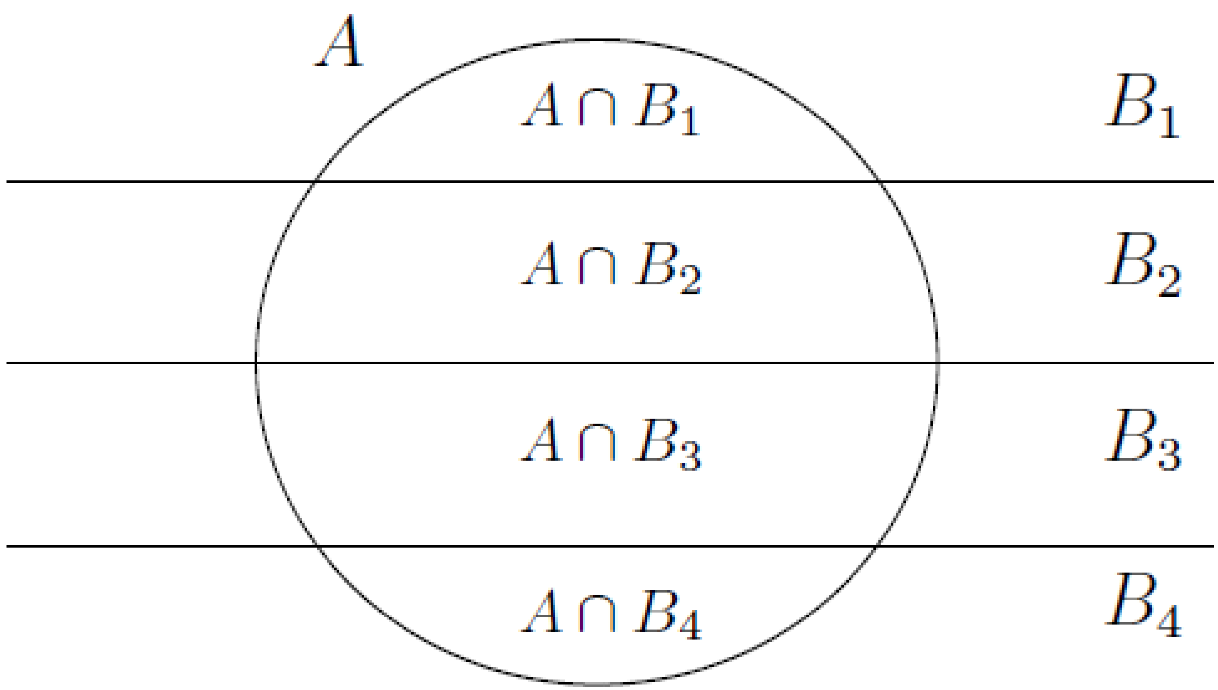

FIGURE 2.6: *Partition of the sample space* Ω *and the event* A.

It means that

- $B_i \cap B_j = \emptyset$ for any $i \neq j$ and $B_1 \cup \ldots \cup B_k = \Omega$.
- These events also partition the event A,

  $A = (A \cap B_1) \cup \ldots \cup (A \cap B_k)$,
- and this is also a union of mutually exclusive events (Figure 2.6). Hence,

$$\boldsymbol{P}\{A\} = \sum_{j=1}^{k} \boldsymbol{P}\{A \cap B_j\},$$

- and we arrive to the following rule.

Law of Total Probability and Bayes rule

Law of Total Probability

$$\boldsymbol{P}\{A\} = \sum_{j=1}^{k} \boldsymbol{P}\{A \mid B_j\} \boldsymbol{P}\{B_j\}$$

In case of two events ($k = 2$),

$$\boldsymbol{P}\{A\} = \boldsymbol{P}\{A \mid B\} \boldsymbol{P}\{B\} + \boldsymbol{P}\{A \mid \overline{B}\} \boldsymbol{P}\{\overline{B}\}$$

- Together with the Bayes Rule, it makes the following popular formula

Bayes Rule for two events

$$\boldsymbol{P}\{B \mid A\} = \frac{\boldsymbol{P}\{A \mid B\} \boldsymbol{P}\{B\}}{\boldsymbol{P}\{A \mid B\} \boldsymbol{P}\{B\} + \boldsymbol{P}\{A \mid \overline{B}\} \boldsymbol{P}\{\overline{B}\}}$$

Example 2.34 (Reliability of a test, continued). Continue Example 2.32.

- Suppose that 4% of all the patients are infected with the virus, $P\{V\} = 0.04$.
- Recall that $P\{S \mid V\} = 0.95$ and $P(S^c \mid V^c) = 0.99$.
- If the test shows positive results, the (conditional) probability that a patient has the virus equals

Example 2.34 (Reliability of a test, continued).

$$\begin{aligned} P\{V \mid S\} &= \frac{P\{S \mid V\} P\{V\}}{P\{S \mid V\} P\{V\} + P\{S \mid \overline{V}\} P\{\overline{V}\}} \\ &= \frac{(0.95)(0.04)}{(0.95)(0.04) + (1 - 0.99)(1 - 0.04)} = \underline{0.7983}. \end{aligned}$$