Chapter 1

Probabilistic Models and Sample Space

Peng-Hua Wang

Graduate Institute of Communication Engineering

National Taipei University

Chapter Contents

1.0 Introduction

Probability

- We use the concept of probability to discuss an uncertain situation. Try to express it in quantity and to make it measurable.
- One approach to define probability is in terms of frequency of occurrence (or called the relative frequency).

Probability

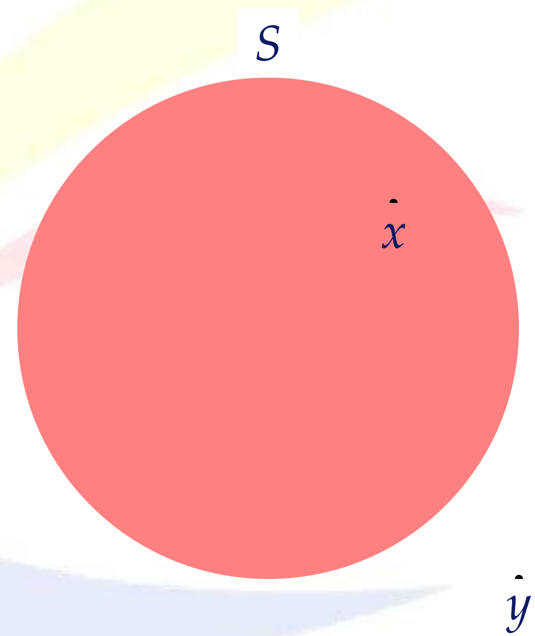
- We can also define probability by "axioms". This mathematical approach makes probability theory strick.
  - axiom: *"An axiom or postulate is a proposition that is not proved or demonstrated but considered to be either self-evident, or subject to necessary decision. "* (From Wiki)
- Probability is a number assigned to a set. Therefore, we begin in a short review of set theory.

1.1 Sets

Sets

A **set** is a collection of objects. These objects are called the **elements** of the set. Let S denote a set and x be an object.

- "$x \in S$" means that x is an element of S
- "$y \notin S$" means that y is not an element of S
- "$\emptyset$" is the symbol of a set that has no elements. That is, the **empty set**.
- Sets and associated operations can be visualized by **Venn graphs.**



Venn graph

Example

Let S be the set of all objects in your pencil case.

- Pen $\in S$
- Ruler $\in S$
- Mirror $\in S$ (?)
- Perfume $\in S$ (??)
- $S = \emptyset$ (Why do you bring an empty pencil case ?)

Specification of a Set

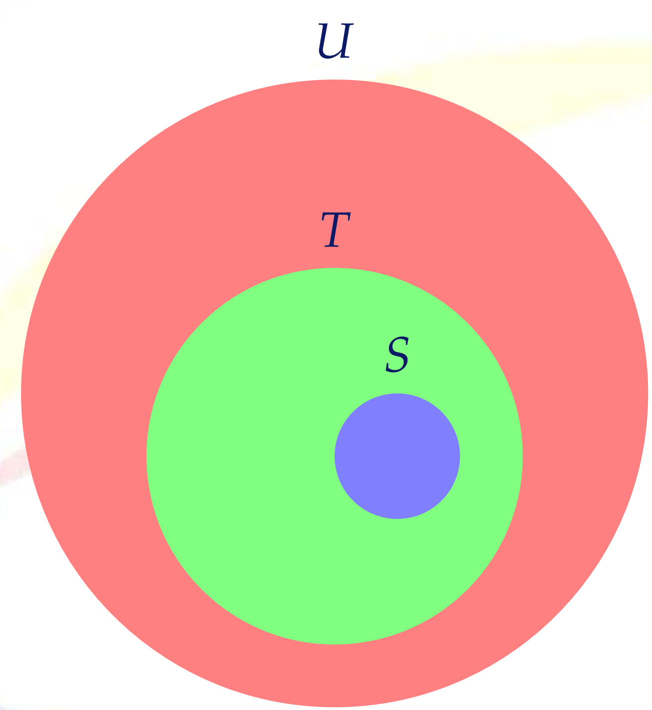
- We can specify a set in a variety ways.
  - "$S = \{x_1, x_2, ..., x_n\}$" means that S contains a finite number of elements $x_1, x_2, ..., x_n$. A set is a list of elements in braces.
  - "$S = \{x_1, x_2, ...\}$" means that S contains infinitely many elements $x_1, x_2, ...$ which can be enumerated. We say S is **countably infinite.**
  - "$S = \{x | x \text{ satisfies } P.\}$" means that all of the elements in the set S satisfy a certain property P.

Examples

- $S_1 = \{$⚀, ⚁, ⚂, ⚃, ⚄, ⚅$\}$ is the set of possible outcomes of a die roll.
- $S_2 = \{H, T\}$ is the set of possible outcomes of a coin toss where H standards for "heads" and T standards for "tails".
- $S_3 = \{0, 2, -2, 4, -4, ...\}$ is the set of all even integers. This is a countably infinite set. We can also write S_3 as $S_3 = \{k | k/2 \text{ is an integer.}\}$.
- $S_4 = \{x | 0 \leq x \leq 1\}$ is the set of all real numbers in the interval $[0, 1]$. It is an **uncountable** set.

Set Relations

- "$S \subset T$" means that S is a **subset** of T. That is, every elements of S is also an element of T.
- "$U \supset T$" means that U is a **superset** of T. That is, every elements of T is also an element of U.
- If $S \subset T$ but $S \neq T$, we say S is a **proper subset** of T.
- If $S \supset T$ but $S \neq T$, we say S is a **proper superset** of T.



$S \subset T, U \supset T$

Set Relations: Properties

- If $S \subset T$, then $T \supset S$.
- If $S \subset T$ and $T \subset U$, then $S \subset U$.
- If $S \subset T$ and $S \supset T$, then $S = T$.
- The empty set is a subset of any set: $\varnothing \subset S$ for all sets S.

Special Sets

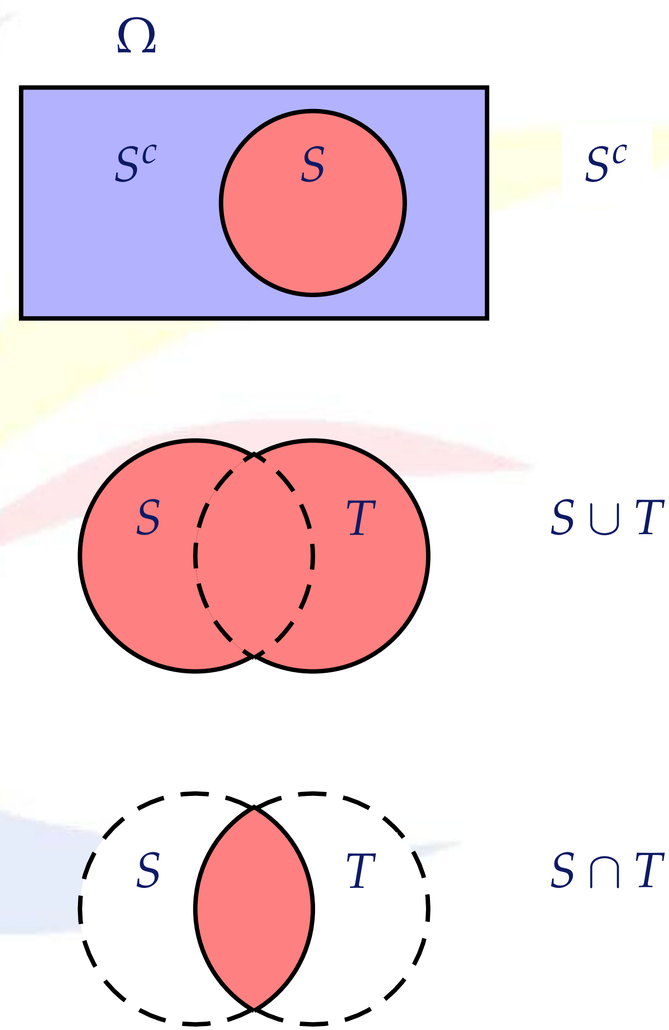
- The **universal set** Ω contains all elements that could be of interest in a particular context. By specifying the context, we can say that $S \subset \Omega$ for all sets S.
- The set of real numbers is denoted by $\Re$.
- The set of pairs of real numbers (i.e., the two-dimensional plane) is denoted by $\Re^2$. That is, $\Re^2 = \{(x, y) | x \in \Re, y \in \Re\}$

Examples

- The universal set of possible outcomes of a die roll is $\Omega = \{$⚀, ⚁, ⚂, ⚃, ⚄, ⚅$\}$
- $\{$⚀, ⚁$\} \subset \{$⚀, ⚁, ⚂, ⚃$\}$
- $\{$⚀, ⚅$\} \not\subset \{$⚀, ⚁, ⚂, ⚃$\}$
- $\{$⚀, ⚁$\} = \{$⚁, ⚀$\}$
- ⚀ $\in \{$⚀, ⚁$\}$
- $\{$⚀$\} \subset \{$⚀, ⚁$\}$

Set Operations

- $S^c = \{x | x \in \Omega \text{ and } x \notin S\}$ is the **complement** set of S.
- $S \cup T = \{x | x \in S \text{ or } x \in T\}$ is the **union** of S and T. That is, the set of all elements that belongs to S or T (or both).
- $S \cap T = \{x | x \in S \text{ and } x \in T\}$ is the **intersection** of S and T. That is, the set of all elements that belongs to S and T .

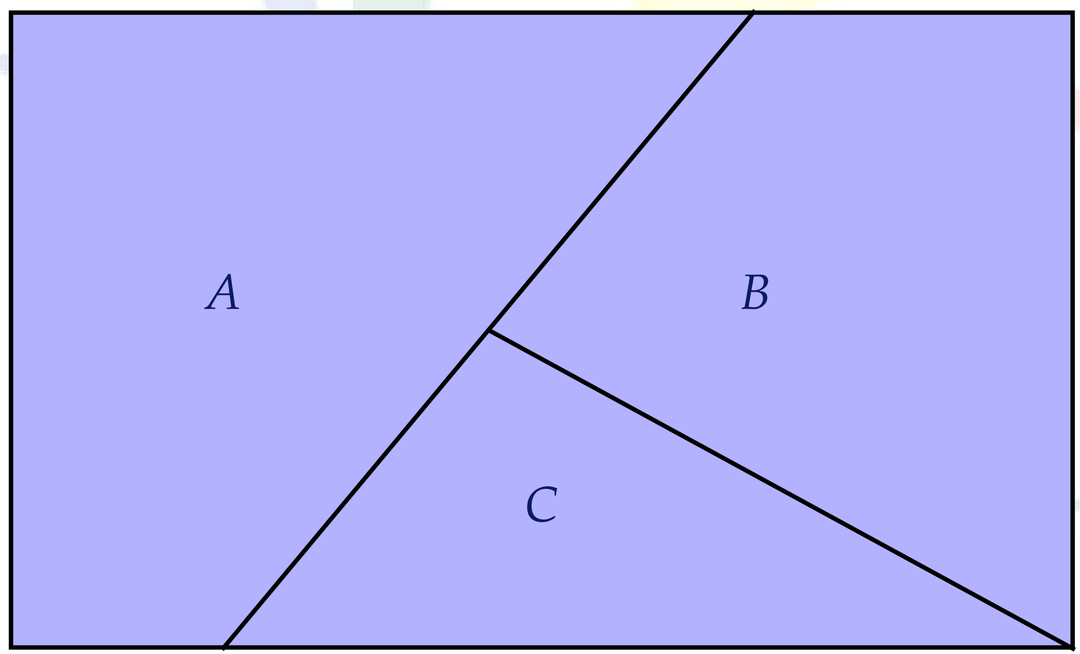


Set Operations

- $S^c = \{x | x \in \Omega \text{ and } x \notin S\}$ is the **complement** set of S.
- $$\bigcup_{n=1}^{\infty} S_n = S_1 \cup S_2 \cup \cdots = \{x | x \in S_n \text{ for some } n.\}$$
- $$\bigcap_{n=1}^{\infty} S_n = S_1 \cap S_2 \cap \cdots = \{x | x \in S_n \text{ for all } n.\}$$

Set Properties

- A and B is said to be **disjoint** if $A \cap B = \varnothing$.
- A collection of set is said to be a **partition** of a set S if these sets are disjoint and their union is S.



Set Algebra

- $S \cup T = T \cup S$
- $S \cup (T \cup U) = (S \cup T) \cup U = S \cup T \cup U$
- $S \cap (T \cup U) = (S \cap T) \cup (S \cap U)$
- $S \cup (T \cap U) = (S \cup T) \cap (S \cup U)$
- $(S^c)^c = S, S \cap S^c = \emptyset, S \cup \Omega = \Omega, S \cap \Omega = S$

De Morgan's Law

$$\left(\bigcup_n S_n\right)^c = \bigcap_n S_n^c, \quad \left(\bigcap_n S_n\right)^c = \bigcup_n S_n^c$$

Proof. Here is a proof of the left identity.

$$(1) x \in \left(\bigcup_n S_n\right)^c \Rightarrow x \notin \left(\bigcup_n S_n\right) \Rightarrow x \notin S_1 \text{ and } x \notin S_2 \text{ and } \ldots$$

$$\Rightarrow x \in S_1^c \text{ and } x \in S_2^c \text{ and } \cdots \Rightarrow x \in \bigcap_n S_n^c \Rightarrow \left(\bigcup_n S_n\right)^c \subset \bigcap_n S_n^c$$

$$(2) x \in \bigcap_n S_n^c \Rightarrow x \in S_1^c \text{ and } x \in S_2^c \text{ and } \cdots \Rightarrow x \notin S_1 \text{ and } x \notin S_2 \text{ and } \ldots$$

$$\Rightarrow x \notin \left(\bigcup_n S_n\right) \Rightarrow x \in \left(\bigcup_n S_n\right)^c \Rightarrow \bigcap_n S_n^c \subset \left(\bigcup_n S_n\right)^c$$

By (1) and (2), we conclude that $(\bigcup_n S_n)^c = \bigcap_n S_n^c$

Fig1.1 in our Textbook

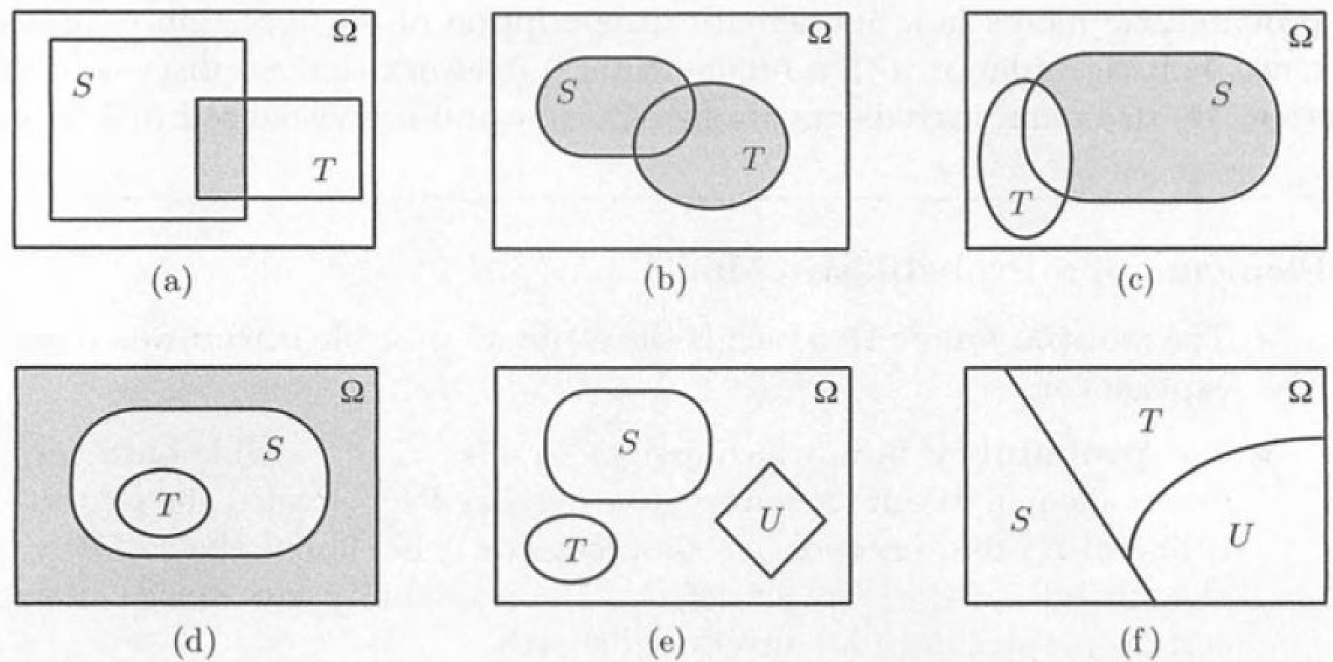


Figure 1.1: Examples of Venn diagrams. (a) The shaded region is $S \cap T$. (b) The shaded region is $S \cup T$. (c) The shaded region is $S \cap T^c$. (d) Here, $T \subset S$. The shaded region is the complement of S. (e) The sets S, T, and U are disjoint. (f) The sets S, T, and U form a partition of the set Ω.

1.2 Probabilistic Models

Probabilistic Models

- We try to describe an uncertain situation in mathematic by means of a **probabilistic model**.
- Elements of a Probabilistic Model
  - The **sample space** Ω, which is the set of all possible outcomes of an experiment.
  - The **probability law**, which maps a set A of possible outcomes (also called an **event**) to a nonnegative number $\mathrm{P}(A)$ (called the **probability** of A)
- In fact, only the problems that can be described in these two elements can be solved by probability theory.

Fig1.2 in our Textbook

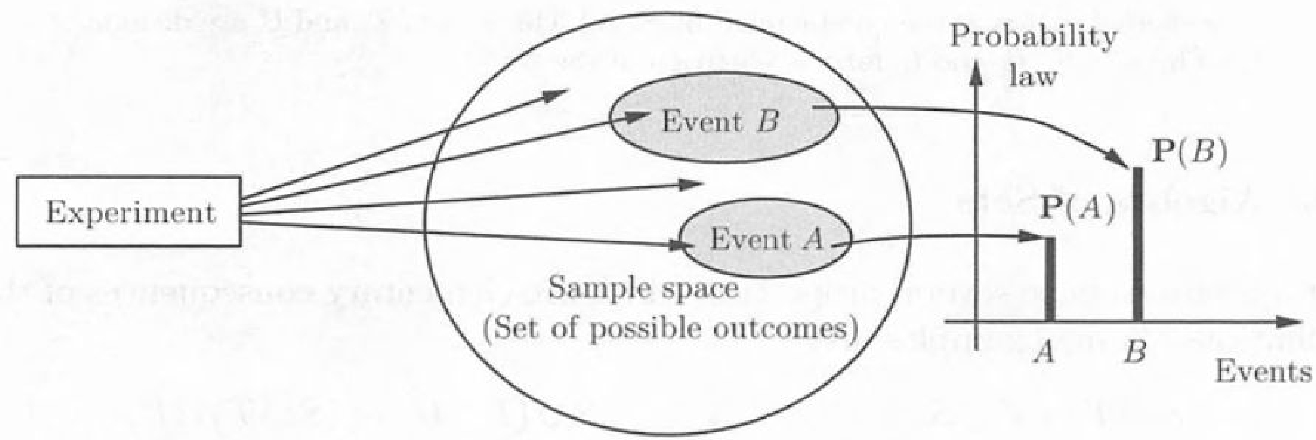


Figure 1.2: The main ingredients of a probabilistic model.

Sample space

- An **experiment** produces exactly one outcomes.
- The **sample space** (denoted by Ω) of the experiment is the set of all possible outcomes.
- An **event** is a collection of possible outcomes, i.e. a subset of the sample space.

Sample space

- Outcome must be:
  - Mutually exclusive
    - For example, a possible outcome of “1 or 3” and another outcome of “1 or 4” cannot be both contained in the sample space associated with a dice roll. Otherwise we cannot assign the probability to the outcome of “1”.
  - Collectively exhaustive
    - In each experiment, we always obtain an outcome in the sample space
  - At the “right” granularity
    - Outcome should distinguish for each other, and avoid unnecessary details.

Example

Sum of two die rolls.

- $\Omega_1 = \{(1,1),(1,2),...(6,6)\}$: too detailed
- $\Omega_2 = \{2,3,4,...,12\}$: good

Sequential Model, Fig 1.3 in our Textbook

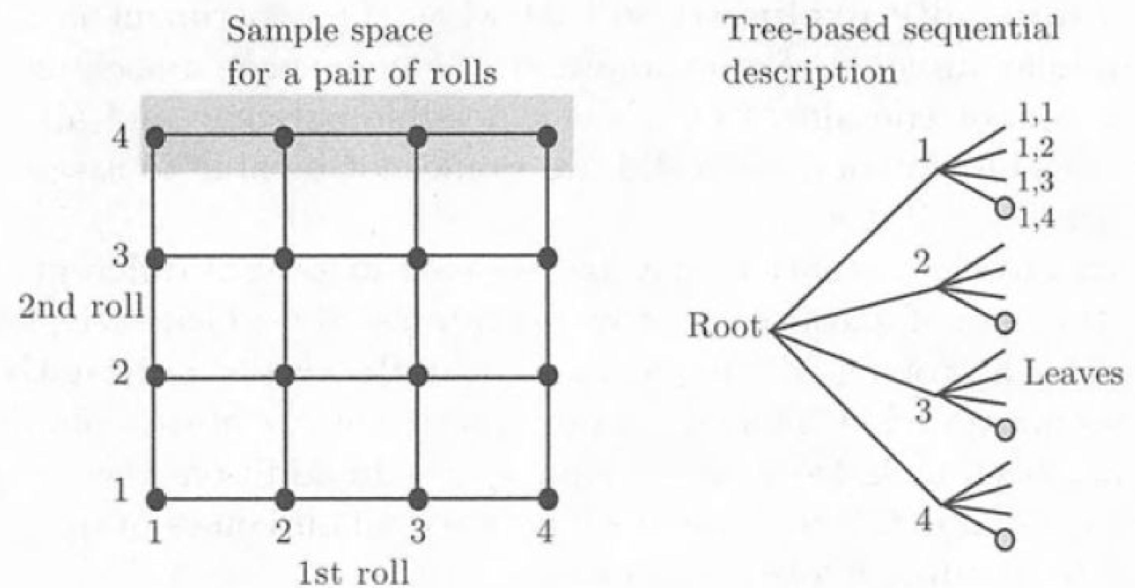


Figure 1.3: Two equivalent descriptions of the sample space of an experiment involving two rolls of a 4-sided die. The possible outcomes are all the ordered pairs

- It is helpful to evaluate the sequential model by tree-based description.

Axioms of probability

- Event A: a subset of the sample space
- Probability law: assign an nonnegative number $\mathrm{P}(A)$ (probability) to every events
- Probability law satisfies the following **probability axioms**:
  1. (Nonnegativity) $\mathrm{P}(A) \geq 0$
  2. (Normalization) $\mathrm{P}(\Omega) = 1$
  3. (Additivity) If $A \cap B = \varnothing$, then $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B)$

Example 1.2: A single coin toss

- There are two possible outcomes, heads (H) and tails (T).
- The sample space is $\Omega = \{H, T\}$
- The possible events are $\{H, T\}, \{H\}, \{T\}, \varnothing$
- Assume the coin is fair in the sense of equally likely outcomes. That is $\mathrm{P}(\{H\}) = \mathrm{P}(\{T\})$
- Solve the probability law. That is, find the probability of $\{H, T\}, \{H\}, \{T\}$, and $\varnothing$

Solution.

- $\mathrm{P}(\Omega) = 1 = \mathrm{P}(\{H, T\}) = \mathrm{P}(\{H\}) + \mathrm{P}(\{T\})$ since $\mathrm{P}(\{H\}) \cap \mathrm{P}(\{T\}) = \varnothing$
- Since $\mathrm{P}(\{H\}) = \mathrm{P}(\{T\})$, we have $\mathrm{P}(\{H\}) = \mathrm{P}(\{T\}) = 0.5$
- Since $\mathrm{P}(\Omega) \cap \varnothing) = \varnothing)$, we have $\mathrm{P}(\Omega) \cup \varnothing) = \mathrm{P}(\Omega) + \mathrm{P}(\varnothing)$. That is, $1 = 1 + \mathrm{P}(\varnothing)$ and $\mathrm{P}(\varnothing) = 0$.

Example 1.2: Three coin tosses

- The sample space is $\Omega =?$
- How many possible events ?
- Assume possible outcome has the same probability.
- What is the probability of exactly 2 heads occur?

Solution.

- The event of interest is $A =?$
- $\mathrm{P}(A) =?$

Example 1.3: Roll two 4-sided dice

- Assume the dice are fair. It means that each of the sixteen possible outcomes has the same probability .
- What is the probability that the sum of the rolls is even?
- What is the probability that the sum of the rolls is odd?
- What is the probability that the first roll is equal to the second?
- What is the probability that the first roll is larger than the second?
- What is the probability that at least one roll is equal to 4?

Example 1.3: Fig 1.4 in our Textbook

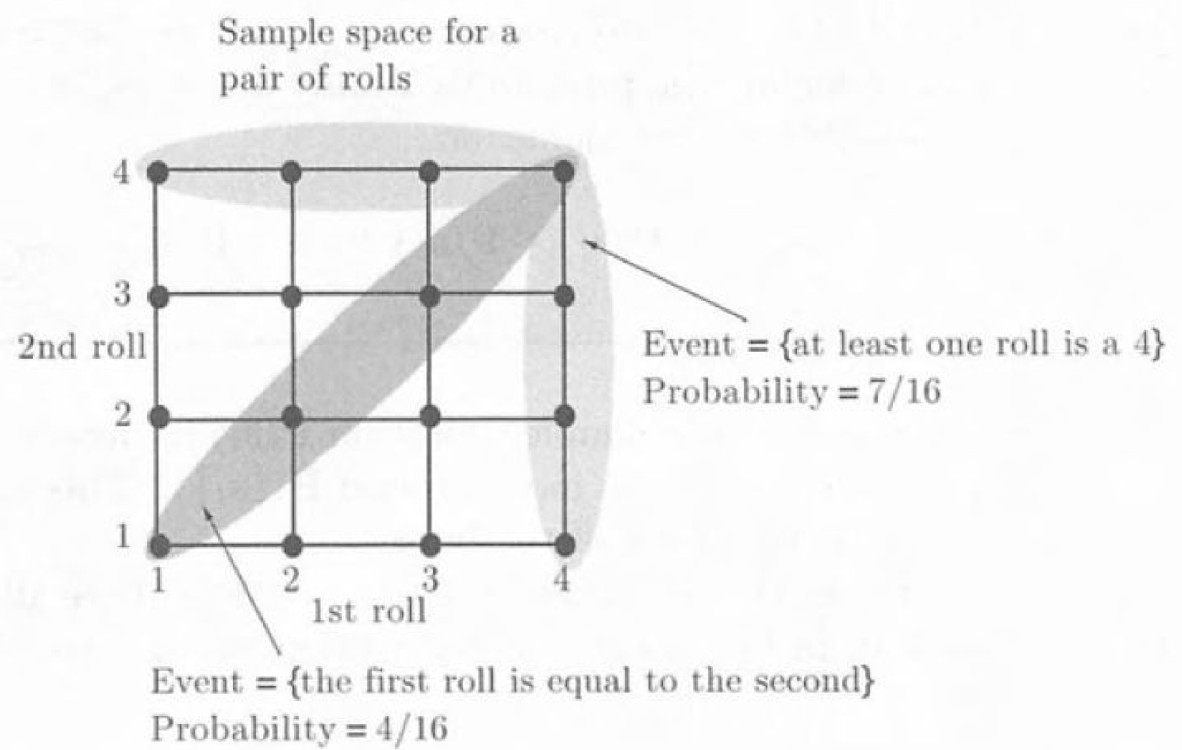


Figure 1.4: Various events in the experiment of rolling a pair of 4-sided dice, and their probabilities, calculated according to the discrete uniform law.

Discrete Probability Law

If the sample space consists of a finite number of possible outcomes $\{s_1, s_2, \ldots, s_n\}$, then the probability law is specified by the probabilities of the events that consist of a single outcome.

In particular, the probability of any event $E = \{s_i, s_j, s_k, \ldots, s_m\}$ is the sum of the probabilities of its outcomes:

$$\mathrm{P}(E) = \mathrm{P}(\{s_i\}) + \mathrm{P}(\{s_j\}) + \mathrm{P}(\{s_k\}) + \cdots + \mathrm{P}(\{s_m\})$$

Discrete Uniform Probability Law

If the sample space consists of N possible outcomes which are equally likely (i.e., all single-element events have the same probability), then the probability of any event A is given by

$$P(A) = \frac{\text{number of elements of } A}{N}$$

Continuous Model. Example 1.4

A number x uniformly comes from the interval $(0, 1)$.
What is the probability $\mathrm{P}(\{x = x_0\})$ of the event consisting of a single element?

If $\mathrm{P}(\{x = x_0\}) = \epsilon$, then $\mathrm{P}(\{x\}) = \epsilon$ for every $x \in (0, 1)$ since x is uniformly distributed in $(0, 1)$. Now, we choose N numbers $x_1, x_2, \ldots x_N$ from $(0, 1)$ and consider the probability $P = \mathrm{P}(\{x = x_0\} \cup \{x = x_1\} \cup \cdots \cup \{x = x_N\})$.
From the axiom of additivity, we have $P = N\epsilon$. For any $\epsilon > 0$, we can always select a large N such that $P > 1$ and this would contradict the normalization axiom.
That is, $\mathrm{P}(\{x\}) = 0$.

Continuous Model. Example 1.4

If we assign probability $b - a$ to any subinterval $[a, b] \in [0, 1]$, then this assignment satisfies the three probability axioms.

Example 1.5

Romeo and Juliet have a date at a given time, and each will arrive at the meeting place with a delay between 0 and 1 hour, with all pairs of delays being equally likely. The first to arrive will wait for 15 minutes and will leave if the other has not yet arrived. What is the probability that they will meet?

- Let x be the delay of Romeo and y the delay of Juliet.
- The sample space $\Omega = \{(x, y) | 0 \leq x \leq 1, 0 \leq y \leq 1\}$
- They meet if the following event occurs

$$M = \{(x, y) || x - y| \leq 1/4, 0 \leq x \leq 1, 0 \leq y \leq 1\}$$

Example 1.5. Fig. 1.5 in the textbook

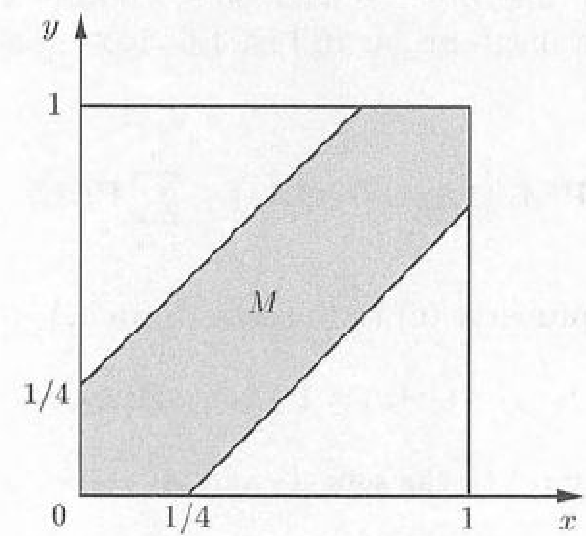


Figure 1.5: The event M that Romeo and Juliet will arrive within 15 minutes of each other (cf. Example 1.5) is

$$M = \left\{(x, y) \;\middle|\; |x - y| \le 1/4,\ 0 \le x \le 1,\ 0 \le y \le 1\right\},$$

The area of shaded area is $1 - (3/4) \times (3/4) = 7/16$. Thus, The probability that they will meet is $7/16$.

Some Properties of Probability Laws

1. If $A \subset B$, then $\mathrm{P}(A) \leq \mathrm{P}(B)$.
2. $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$.
3. $\mathrm{P}(A \cup B) \leq \mathrm{P}(A) + \mathrm{P}(B)$.
4. $\mathrm{P}(A \cup B \cup C) = \mathrm{P}(A) + \mathrm{P}(A^c \cap B) + \mathrm{P}(A^c \cap B^c \cap C)$.

These properties can be visualized and verified by using Venn diagrams.

1.3 Conditional Probability

Conditional Probability

- Reason about the outcome of an experiment based on partial information. For example,
  - It is rainy today. What is the probability that it will be sunny tomorrow?
  - In a word guessing game, the first letter of the word is a “t”. What is the likelihood that the second letter is an “h”?
  - In an experiment involving two successive rolls of a die, you are told that the sum of the two rolls is 9. How likely is it that the first roll was a 6?

Conditional Probability. Example

- Roll a fair die. If we know the outcome is even, then there are 3 outcomes left for consideration.

$$P(\text{outcome is } 6 | \text{outcome is even}) = \frac{1}{3}$$

Conditional Probability

- We know that some given event B occurs. We wish to know the probability that some other given event A also occurs.
- In other words, we know that the outcome is in B. We wish to quantify the likelihood that the outcome also belongs to A.
- This is a new probability law: the *conditional probability* of A given B, denoted by $\mathrm{P}(A|B)$.

$$\mathrm{P}(A|B) \triangleq \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} \quad \text{if } \mathrm{P}(B) > 0$$

- Note. "Given B" means that B must occur. That is, $\mathrm{P}(B) > 0$.

Conditional Probability. Axioms

- We have to prove that the above definition of conditional probabilities satisfies the probability axioms.
  - Nonnegativity:

$$\mathrm{P}(A|B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} > 0$$

    since $\mathrm{P}(A \cap B) > 0$ and $\mathrm{P}(B) > 0$

  - Normalization:

$$\mathrm{P}(\Omega|B) = \frac{\mathrm{P}(\Omega \cap B)}{\mathrm{P}(B)} = \frac{\mathrm{P}(B)}{\mathrm{P}(B)} = 1$$

  - Additivity.

Example 1.6

- Toss a fair coin three successive times.
- Let A and B are the events

$$A = \{\text{more heads than tails come up}\},$$

and

$$B = \{\text{1st toss is a head}\}.$$

- Find the conditional probability $\mathrm{P}(A|B)$.

Example 1.7

- A fair 4-sided die is rolled twice and we assume that all sixteen possible outcomes are equally likely.
- Let X and Y be the result of the 1st and the 2nd roll, respectively.
- Events A and B are

$$A = \{\max(X, Y) = m\}, B = \{\min(X, Y) = 2\}$$

  and $m = 1, 2, 3, 4$.
- Determine the conditional probability $P(A|B)$,

Example 1.8

- We have two design teams “C” and “N”. They are asked to separately design a new product within a month. From past experience we know that:
  1. The probability that team C is successful is 2/3.
  2. The probability that team N is successful is 1/2.
  3. The probability that at least one team is successful is 3/4.
- Assuming that exactly one successful design is produced, what is the probability that it was designed by team N?

Example 1.9 Radar Detection.

- If an aircraft is present in a certain area, a radar detects it and generates an alarm signal with probability 0.99.
- If an aircraft is not present, the radar generates a (false) alarm, with probability 0.10.
- We assume that an aircraft is present with probability 0.05.
- What is the probability of no aircraft presence and a false alarm? What is the probability of aircraft presence and no detection?

Example 1.9 Radar Detection. Fig. 1.9

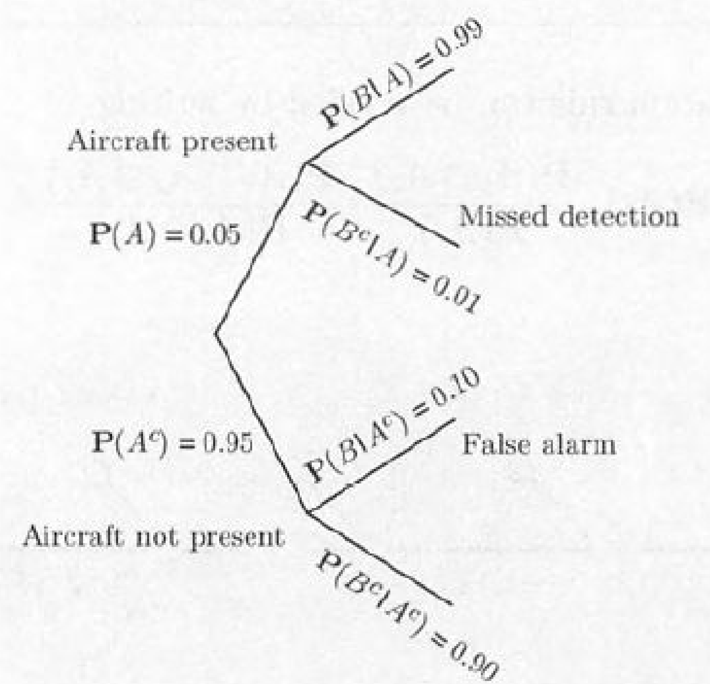


Figure 1.9: Sequential description of the experiment for the radar detection problem in Example 1.9.

Example 1.10

- Three cards are drawn from an ordinary 52-card deck without replacement (drawn cards are not placed back in the deck).
- Find the probability that none of the three cards is a heart.

Multiplication Rule

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

In general, we have

$$P(A_1 \cap A_2 \cap \cdots A_n)$$
$$= P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P(A_n|A_1 \cap A_2 \cap \cdots A_{n-1})$$

Proof.

$$P(A_1 \cap A_2 \cap \cdots A_n)$$
$$= P(A_1)\frac{P(A_1 \cap A_2)}{P(A_1)}\frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cap A_2)} \cdots \frac{P(A_1 \cap A_2 \cdots \cap A_n)}{P(A_1 \cap A_2 \cdots \cap A_{n-1})}$$

1.4 Total Probability Theorem And Bayes’ Rule

Total Probability Theorem

- Let $A_1, A_2..., A_n$ be disjoint events that form a partition of the sample space and assume that $P(A_i) > 0$, for all i. Then, for any event B, we have

$$
\begin{aligned}
P(B) &= P(A_1 \cap B) + P(A_2 \cap B) + \cdots + P(A_1 \cap B) \\
&= P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + \cdots \\
&\quad + P(A_n)P(B|A_n).
\end{aligned}
$$

Fig. 1.13

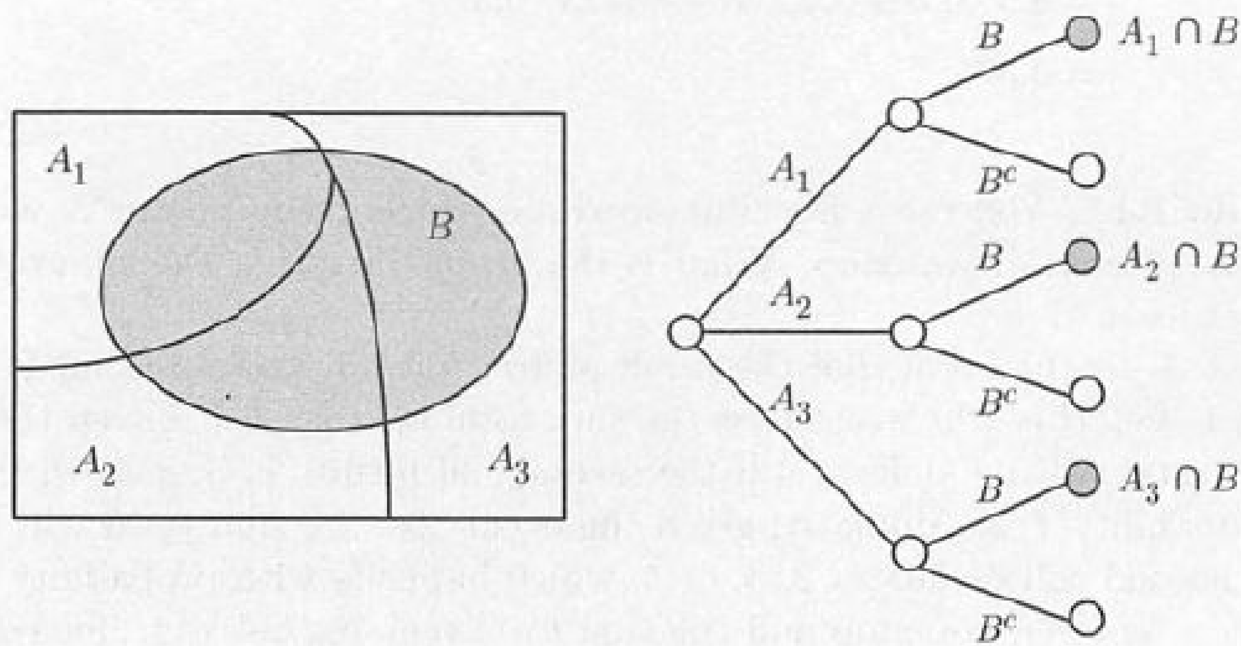


Figure 1.13: Visualization and verification of the total probability theorem. The events $A_1, \ldots, A_n$ form a partition of the sample space, so the event B can be decomposed into the disjoint union of its intersections $A_i \cap B$ with the sets A_i,

Example 1.13

- In a chess game your probability of winning a game is 0.3 against 50% of the players (call them type 1), 0.4 against 25% of the players (call them type 2), and 0.5 against the remaining 25% of the players (call them type 3).
- You play game against a randomly chosen opponent.
- What is the probability of winning?

Example 1.14

- Roll a fair 4-sided die.
- If the result is 1 or 2, you roll once more but otherwise, you stop.
- What is the probability that the sum total of your rolls is at least 4?

Example 1.15

- You are taking a probability class and at the end of each week you can be either up-to-date or you may have fallen behind.
- If you is up-to-date in a given week, the probability that she will be up-to-date (or behind) in the next week is 0.8 (or 0.2, respectively).
- If you is behind in a given week. the probability that you will be up-to-date (or behind) in the next week is 0.4 (or 0.6, respectively).
- You are (by default) up-to-date when you starts the class.
- What is the probability that you are up-to-date after three weeks?

Bayes' Rule

Let $A_1, A_2, \ldots, A_n$ be disjoint events that form a partition of the sample space. Then, for any event B such that $\mathrm{P}(B) > 0$, we have

$$\mathrm{P}(A_i|B) = \frac{\mathrm{P}(A_i \cap B)}{\mathrm{P}(B)}$$

$$= \frac{\mathrm{P}(A_i)\mathrm{P}(B|A_i)}{\mathrm{P}(A_1)\mathrm{P}(B|A_1) + \mathrm{P}(A_2)\mathrm{P}(B|A_2) + \cdots \mathrm{P}(A_n)\mathrm{P}(B|A_n)}$$

- Calculate $\mathrm{P}(A|B)$ from $\mathrm{P}(B|A)$

Inference

- We may regard the events $A_1, \ldots, A_n$ as the causes and the event B represents the associated effect.
- By observing the effect, we want to infer the cause.
- Bayes' rule is a method used for inference.
- Given that the effect B has been observed, we want to calculate the probability $\mathrm{P}(A_i|B)$ that the cause A_i is present.
- $\mathrm{P}(A_i|B)$ is called the **posterior probability** of event A_i given the information,
- $\mathrm{P}(A_i)$ is called the **prior probability**.

Example 1.16/1.9 Radar Detection.

- If an aircraft is present in a certain area, a radar detects it and generates an alarm signal with probability 0.99.
- If an aircraft is not present, the radar generates a (false) alarm, with probability 0.10.
- We assume that an aircraft is present with probability 0.05.
- Let $A = \{\text{an aircraft is present}\}$ and $B = \{\text{the radar generates an alarm}\}$
- Calculate $\mathrm{P}(\text{aircraft present}|\text{alarm}) = \mathrm{P}(A|B)$.

Example 1.17/1.13

- In a chess game your probability of winning a game is 0.3 against 50% of the players (call them type 1), 0.4 against 25% of the players (call them type 2), and 0.5 against the remaining 25% of the players (call them type 3).
- You play game against a randomly chosen opponent.
- Suppose that you win. What is the probability that you had an opponent of type 1?

Example 1.18. The False-Positive Puzzle.

- A test for a certain rare disease is assumed to be correct 95% of the time, that is,
  - if a person has the disease, the test results are positive with probability 0.95, and
  - if the person does not have the disease, the test results are negative with probability 0.95.
- A random person drawn from a certain population has probability 0.001 of having the disease.
- Given that the person just tested positive, what is the probability of having the disease?

1.5 Independence

Independence

- $\mathrm{P}(A|B)$ is the partial information that event B provides about event A.
- When B provides no such information and does not alter the probability that A has occurred, that is

$$\mathrm{P}(A|B) = \mathrm{P}(A),$$

we say that A is **independent of** B.

- By above definition, if A is independent of B, then

$$\mathrm{P}(A|B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} = \mathrm{P}(A)$$
$$\Rightarrow \mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B)$$

Independence

$$\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B)$$

- Above identity is the definition of independence.
- This definition can be used even when $\mathrm{P}(B) = 0$, in which case $P(A|B)$ is undefined.
- Independence is a symmetric property: if A is independent of B, then B is independent of A.
- Property. If A and B are independent, then A^c and B are independent.

Example 1.19

- Two successive rolls of a fair 4-sided die.
- Let $A_i = \{$ lst roll results in $i\}$, and $B_k = \{$ 2nd roll results in $k\}$. Are A_i and B_k independent?
- Let $A = \{$ lst roll results in $1\}$, and $B = \{$ sum of the two rolls is a $5\}$. Are A and B independent?
- Let $A = \{$maximum of the two rolls is $2\}$, and $B = \{$minimum of the two rolls is $2\}$. Are A and B independent?

Conditional Independence

- Given an event C, the events A and B are called **conditionally independent** if

$$\mathrm{P}(A \cap B|C) = \mathrm{P}(A|C)\mathrm{P}(B|C)$$

- If A and B are conditionally independent given C, then

$$\mathrm{P}(A|B \cap C) = \mathrm{P}(A|C).$$

- $\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B) \not\Rightarrow \mathrm{P}(A \cap B|C) = \mathrm{P}(A|C)\mathrm{P}(B|C)$
- $\mathrm{P}(A \cap B|C) = \mathrm{P}(A|C)\mathrm{P}(B|C) \not\Rightarrow \mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B)$

Example 1.20

- Toss two independent fair coins. Let

$$\begin{aligned} H_1 &= \{\text{1st toss is a head}\}, \\ H_2 &= \{\text{2nd toss is a head}\}, \\ D &= \{\text{the two tosses have different results}\}. \end{aligned}$$

- Are H_1 and H_2 independent?
- Given D, are H_1 and H_2 independent?

Example 1.21

- There are two biased coins C_1 and C_2. We know that $\mathrm{P}(C_1 = H) = 0.99$ and $\mathrm{P}(C_2 = H) = 0.01$,
- We choose one of the two equally likely, and proceed with two independent tosses.
- Let

$$\begin{aligned} H_1 &= \{\text{1st toss is a head}\}, \\ H_2 &= \{\text{2nd toss is a head}\}, \\ B &= \{C_1 \text{ is selected}\}. \end{aligned}$$

- Given B, are H_1 and H_2 independent?
- Are H_1 and H_2 independent?

Independence of 3 or More Events

- Definition. The events $A_1, A_2, \ldots, A_n$ are independent if

$$\mathrm{P}(A_i \cap A_j) = \mathrm{P}(A_i)\mathrm{P}(A_j), \quad \text{for any } i \text{ and } j.$$

$$\mathrm{P}(A_i \cap A_j \cap A_k) = \mathrm{P}(A_i)\mathrm{P}(A_j)\mathrm{P}(A_k), \quad \text{for any } i, j \text{ and } k.$$

$$\vdots$$

$$\mathrm{P}(A_i \cap A_j \cap \cdots \cap A_s) = \mathrm{P}(A_i)\mathrm{P}(A_j)\cdots\mathrm{P}(A_s), \quad \text{for any } i, j \ldots s.$$

- **NOT ONLY** piecewise independent.

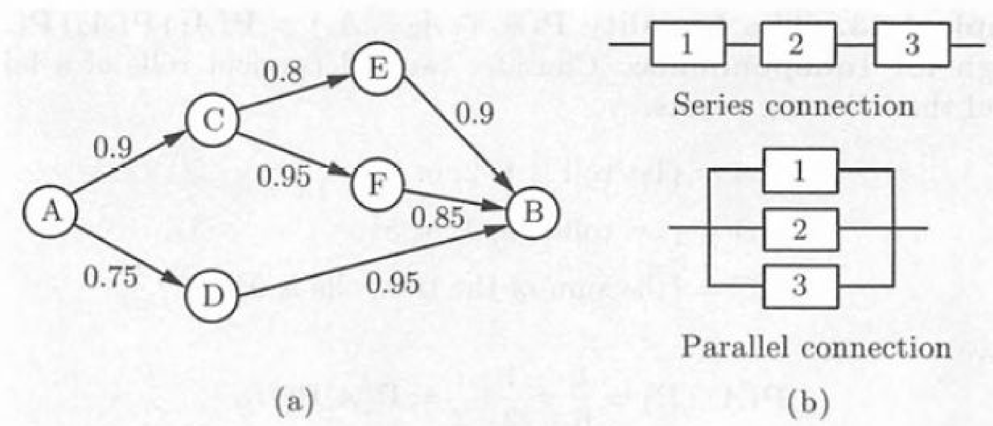
Example 1.23

- Two independent rolls of a fair six-sided die.Let :

$$A = \{\text{ 1st roll is 1, 2, or 3}\},$$
$$B = \{\text{ 1st roll is 3, 4, or 5}\},$$
$$C = \{\text{ the sum of the two rolls is 9}\}.$$

- Find $P(A), P(B), P(C), P(A \cap B), P(A \cap C), P(B \cap C)$ and $P(A \cap B \cap C)$.
- $P(A \cap B \cap C) = P(A)P(B)P(C)$ is not enough for independence.

Example 1.24. Reliability



- A network connects two nodes A and B through intermediate nodes C, D, E, F, as shown in Fig. 1.15(a).
- For every connected pair i and j, there is a given probability P_{ij} that the link from i to j is up.

Example 1.24. Reliability

- Assuming that link failures are independent of each other.
- What is the probability that there is a path connecting A and B in which all links are up?

Bernoulli Trials. Binomial Probability

- If an experiment involves only two results, we say that we have a **Bernoulli trials**.
- If an experiment involves a sequential of Bernoulli trials, a given results form a **binomial probability law**.
- For example. The result of a coin toss forms a Bernoulli trial. The number of heads in n independent tosses form a binomial probability law. Let p be the probability of a head in a coin toss.

Bernoulli Trials. Binomial Probability

- The result in n independent tosses forms a sequence. $\binom{n}{k}$ is the number of sequences that contain k heads.

$$\binom{n}{k} = \frac{n!}{k!(n-k)!},$$
$$n! = 1 \times 2 \times \cdots \times n, \quad 0! \triangleq 1$$

- $n!$ is called n factorial.
- The probability that the number of heads in n independent tosses is k is

$$p(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

1.6 COUNTING

The Counting Principle

- A process consists of r stages. Suppose that
  - There are n_1 possible results at the first stage.
  - For every possible result at the first stage, there are n_2 possible results at the second stage.
  - For every possible result at the $i-1$th stage, there are n_i possible results at the i stage.
  - The total number of possible results of the process is

$$n_1 n_2 \cdots n_r$$

Permutations

- We have n distinct objects.
- We wish to count the number of different ways that we can pick k out of these n objects.
- We can choose any of the n objects for the first one. Having chosen the first, there are only $n-1$ possible choices for the second; given the choice of the first two, there only remain $n-2$ available objects for the 3rd, etc.
- By the counting principle, the number of possible sequences is

$$n \times (n-1) \times \cdots \times (n-k+1) = \frac{n!}{(n-k)!}$$

- This is called the k-permutations.

Combinations

- We have n people.
- We wish to form a team of k people.
- How many different teams are possible?
- For example. 4 people A, B C, and D. a team of 2 persons.
- The first person can be A, B C, and D. The second have 3 choices. We have following possible teams.
  AB, AC, AD, BA, BC, BD, CA, CB, CD, DA, DB, DC
- However, the order is of of matter. Only 6 distinct teams.

Combinations

- In general, for k-team from n people, we have $n!/(n-k)!$ permutations. Each k-team forms $k!$ permutations. Therefore, the number of distinct team is

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

- The is the number of combinations of k objects out of n.

Partitions

- We have n people.
- We wish to form a team of k_1 people and another team of k_2.
- How many different teams are possible?
- First, we choose n_1 people and then choose n_2 from the rest $n - n_1$.

$$\binom{n}{n_1}\binom{n-n_1}{n_2} = \frac{n!}{n_1!n_2!(n-n_1-n_2)!} \triangleq \binom{n}{n_1, n_2, n-n_1-n_2}$$

Partitions

- In general, we have n people.
- We wish to form r teams, the ith team has k_i people, $n_1 + n_2 + \cdots + n_r = n$. The number of distinct teams is

$$\binom{n}{n_1, n_2, \ldots, n_r} = \frac{n!}{n_1! n_2! \ldots n_r!}$$

- This is called the multinomial coefficient.
- $\binom{n}{k}$ is the coefficient of $x^k y^{n-k}$ in the expansion of $(x+y)^n$. (binomial coefficient)
- $\binom{n}{n_1, n_2, \ldots, n_r}$ is the coefficient of $x_1^{n_1} x_2^{n_2} \cdots x_r^{n_r}$ in the expansion of $(x_1 + x_2 + \cdots + x_r)^n$. (multinomial coefficient)