PROBABILITY AND STATISTICS

CHAPTER 1: INTRODUCTION TO PROBABILITY

Dr. Phan Thi Huong

HoChiMinh City University of Technology
Faculty of Applied Science, Department of Applied Mathematics
Email: huongphan@hcmut.edu.vn



HCM city - 2021.



Basic concepts of events and probability. Addition rule and Product rule. Total probability formula and Bayes formula. Independence

OUTLINE

1 BASIC CONCEPTS OF EVENTS AND PROBABILITY.

OUTLINE

- BASIC CONCEPTS OF EVENTS AND PROBABILITY.
- 2 Addition rule and Product rule.

OUTLINE

- 1 BASIC CONCEPTS OF EVENTS AND PROBABILITY.
- 2 Addition rule and Product rule.
- 3 TOTAL PROBABILITY FORMULA AND BAYES FORMULA.

OUTLINE

- BASIC CONCEPTS OF EVENTS AND PROBABILITY.
- 2 Addition rule and Product rule.
- 3 TOTAL PROBABILITY FORMULA AND BAYES FORMULA.
- 4 INDEPENDENCE

LEARNING OUTCOMES

After careful study of this chapter, you should be able to do the following:

- Understand and describe sample spaces and events for random experiments
- 2 Interpret and calculate probabilities of events.
- Interpret and calculate conditional probabilities of events
- Determine the independence of events and use independence to calculate probabilities
- Use Bayes' theorem to calculate conditional probabilities
- Use Total probability formula to calculate a probability of an event.



Sample Spaces and Events Definitions and Interpretations Probability

• An experiment is a procedure that is

- An experiment is a procedure that is
 - carried out under controlled conditions, and

- An experiment is a procedure that is
 - carried out under controlled conditions, and
 - · executed to discover an unknown result.

- An experiment is a procedure that is
 - carried out under controlled conditions, and
 - executed to discover an unknown result.
- A trial is a single performance of an experiment.

- An experiment is a procedure that is
 - carried out under controlled conditions, and
 - executed to discover an unknown result.
- A trial is a single performance of an experiment.
- An outcome is the result of a trial.

- An experiment is a procedure that is
 - carried out under controlled conditions, and
 - executed to discover an unknown result.
- A trial is a single performance of an experiment.
- An outcome is the result of a trial.
- Deterministic or predictable experiments: only one possible result or outcome.

- An experiment is a procedure that is
 - carried out under controlled conditions, and
 - executed to discover an unknown result.
- A trial is a single performance of an experiment.
- An outcome is the result of a trial.
- Deterministic or predictable experiments: only one possible result or outcome.
- Random experiment: different outcomes even when repeated in the same manner every time, ex: Tossing a coin or a dice, measuring a person wight, measuring the current of a thin copper wire, ...

• The set of all possible outcomes of a random experiment is called the sample space of the experiment. The sample space is denoted as Ω .

- The set of all possible outcomes of a random experiment is called the sample space of the experiment. The sample space is denoted as Ω .
- Points ω in Ω are called sample outcomes, realizations, or elements

- The set of all possible outcomes of a random experiment is called the sample space of the experiment. The sample space is denoted as Ω .
- Points ω in Ω are called sample outcomes, realizations, or elements
- A sample space is discrete if it consists of a finite or countable infinite set of outcomes

- The set of all possible outcomes of a random experiment is called the sample space of the experiment. The sample space is denoted as Ω .
- Points ω in Ω are called sample outcomes, realizations, or elements
- A sample space is discrete if it consists of a finite or countable infinite set of outcomes
- A sample space is **continuous** if it contains an interval (either finite or infinite) of real numbers.

EXAMPLE 1

EXAMPLE 1

•
$$\Omega = \mathbb{R}^+ = \{x | x > 0\}$$

EXAMPLE 1

- $\Omega = \mathbb{R}^+ = \{x | x > 0\}$
- If it is known that all recycle times are between 1.5 and 5 seconds, $\Rightarrow \Omega = \{x \mid 1.5 < x < 5\}$

EXAMPLE 1

- $\Omega = \mathbb{R}^+ = \{x | x > 0\}$
- If it is known that all recycle times are between 1.5 and 5 seconds, $\Rightarrow \Omega = \{x \mid 1.5 < x < 5\}$
- If the objective of the analysis is to consider only whether the recycle time is low, medium, or high, ⇒
 Ω = {low, minium, high}

EXAMPLE 1

- $\Omega = \mathbb{R}^+ = \{x | x > 0\}$
- If it is known that all recycle times are between 1.5 and 5 seconds, $\Rightarrow \Omega = \{x \mid 1.5 < x < 5\}$
- If the objective of the analysis is to consider only whether the recycle time is low, medium, or high, ⇒
 Ω = {low, minium, high}
- If the objective is only to evaluate whether or not a particular camera conforms to a minimum recycle time specification, ⇒ Ω = {yes, no}.

Sample Spaces and Events Definitions and Interpretations Probability

• An event *A* is a subset of the sample space of a random experiment. (= statement, sentence, clause)

- An event *A* is a subset of the sample space of a random experiment. (= statement, sentence, clause)
- A null event is the one never happened, denoted by \emptyset .

- An event *A* is a subset of the sample space of a random experiment. (= statement, sentence, clause)
- A null event is the one never happened, denoted by \emptyset .
- A true event is the one always happened, denoted by Ω .

EXAMPLE 2

- An event *A* is a subset of the sample space of a random experiment. (= statement, sentence, clause)
- A null event is the one never happened, denoted by \emptyset .
- A true event is the one always happened, denoted by Ω .

EXAMPLE 2

- An event *A* is a subset of the sample space of a random experiment. (= statement, sentence, clause)
- A null event is the one never happened, denoted by \emptyset .
- A true event is the one always happened, denoted by Ω .

EXAMPLE 2

②
$$\Omega = \{x \mid 1.5 < x < 5\} \Rightarrow A_1 = (1,3), A_2 = [2,3.5], ...$$

- An event *A* is a subset of the sample space of a random experiment. (= statement, sentence, clause)
- A null event is the one never happened, denoted by \emptyset .
- A true event is the one always happened, denoted by Ω .

EXAMPLE 2

- ② $\Omega = \{x \mid 1.5 < x < 5\} \Rightarrow A_1 = (1,3), A_2 = [2,3.5], ...$
- $Ω = {low, minium, high} ⇒ A₁ = {low, minium},$ $A₂ = {low}, ...$

- An event *A* is a subset of the sample space of a random experiment. (= statement, sentence, clause)
- A null event is the one never happened, denoted by \emptyset .
- A true event is the one always happened, denoted by Ω .

EXAMPLE 2

- ② $\Omega = \{x | 1.5 < x < 5\} \Rightarrow A_1 = (1,3), A_2 = [2,3.5], ...$
- $Ω = {low, minium, high} ⇒ A₁ = {low, minium},$ $A₂ = {low}, ...$
- **1** $\Omega = \{yes, no\}, \Rightarrow A_1 = \{yes\}, A_2 = \{no\}.$





OPERATIONS

• The complement of event A is an event which consists of all outcomes in the sample space that are not in A: $\overline{A} = \Omega \setminus A$. Also denoted by A^c or A'.

OPERATIONS

- **1** The complement of event *A* is an event which consists of all outcomes in the sample space that are not in *A*: $\overline{A} = \Omega \setminus A$. Also denoted by A^c or A'.
- ② The product of two events A and B is the intersection of A and B which consist of all outcomes that are both in A and in B: $AB = A \cap B = \{x : x \in A \text{ and } x \in B\}.$

OPERATIONS

- **1** The complement of event *A* is an event which consists of all outcomes in the sample space that are not in *A*: $\overline{A} = \Omega \setminus A$. Also denoted by A^c or A'.
- ② The product of two events A and B is the intersection of A and B which consist of all outcomes that are both in A and in B: $AB = A \cap B = \{x : x \in A \text{ and } x \in B\}.$
- The sum of two events *A* and *B* is the union of *A* and *B* which consist of all outcomes that are in *A* or in *B*:

$$A+B=A\cup B=\{x:x\in A\ or\ x\in B\}.$$

OPERATIONS

- **1** The complement of event *A* is an event which consists of all outcomes in the sample space that are not in *A*: $\overline{A} = \Omega \setminus A$. Also denoted by A^c or A'.
- ② The product of two events A and B is the intersection of A and B which consist of all outcomes that are both in A and in B: $AB = A \cap B = \{x : x \in A \text{ and } x \in B\}.$
- **③** The sum of two events *A* and *B* is the union of *A* and *B* which consist of all outcomes that are in *A* or in *B*: $A + B = A \cup B = \{x : x \in A \text{ or } x \in B\}.$
- The different between *A* and *B* is the event which consist of all outcomes that are in *A* but not in *B*:

$$A - B = A \setminus B = \{x : x \in A \text{ and } x \notin B\}.$$

MUTUALLY EXCLUSIVE EVENTS

If $AB = \emptyset$, we say that A and B are disjoint, or mutually exclusive, e.i, A and B cannot simultaneously occur.

EXHAUSTIVE EVENTS

A set of n events $A_1, A_2, ..., A_n$ is called an exhaustive set if and only if

$$\begin{cases} A_i.A_j = \emptyset, \ i, j \in \overline{1, n} \\ A_1 + A_2 + \dots + A_n = \Omega \end{cases}$$





ALGEBRA OF EVENTS' OPERATIONS

- De Morgan's Rules
 - $\bullet \ \overline{A+B} = \overline{A} \, \overline{B}.$
 - $\overline{AB} = \overline{A} + \overline{B}$.
- 2 Distributive law
 - $\bullet \ \ A(B+C) = AB + AC.$
 - A + (BC) = (A + B)(A + C).
- Opening in the property of the property of
 - A (B + C) = (A B)(A C)
 - A (BC) = (A B) + (A C)

EXAMPLE 3

Suppose that the recycle times of two cameras are recorded. If the objective of the analysis is to consider only whether or not the cameras conform to the manufacturing specifications, either camera may or may not conform. We abbreviate yes and no as y and n. If the ordered pair (yn) indicates that the first camera conforms and the second does not, write below events by notations.

- \blacksquare At least one camera conforms E_1 .
- **1** Both cameras do not conform E_2 .



EXAMPLE 4

Disks of polycarbonate plastic from a supplier are analyzed for scratch and shock resistance. The results from 100 disks are summarized here:

Shock Resistance

		High	Low
Scratch	High	70	9
Resistance	Low	16	5

Let *A* denote the event that a disk has high shock resistance, and let *B* denote the event that a disk has high scratch resistance. Determine the number of disks in A.B, \overline{A} , and A + B.

DEFINITIONS AND INTERPRETATIONS PROBABILITY

INTERPRETATIONS OF PROBABILITY

Probability is used to quantify the likelihood or chance that a particular outcome or event from a random experiment will occur. The likelihood of an outcome is quantified by assigning a number from the interval [0, 1] to the outcome.

EQUALLY LIKELY OUTCOMES

If all outcomes in a finite set Ω are equally likely, the probability of A is the number of outcomes in A divided by the total number of outcomes:

$$P(A) = \frac{\#A}{\#\Omega}$$



EXAMPLE 5

If the last digit of a weight measurement is equally likely to be any of the digits 0 through 9,

- What is the probability that the last digit is 0?
- What is the probability that the last digit is greater than or equal to 5?

FREQUENCIES

A relative frequency is a proportion measuring how often, or how frequently, something or other occurs in a sequence of observations. Think of some experiment which can be repeated n trial. Let A be a possible result of such a trial. If A happens m times in n trials, then $f_n(A) = \frac{m}{n}$ is the relative frequency of A in the n trials. Then,

$$\mathbb{P}(A) = \lim_{n \to \infty} f_n(A) = \frac{m}{n} \tag{1}$$

or

$$f_n(A) \approx P(A)$$

for large n.

FREQUENCIES

A relative frequency is a proportion measuring how often, or how frequently, something or other occurs in a sequence of observations. Think of some experiment which can be repeated n trial. Let A be a possible result of such a trial. If A happens m times in n trials, then $f_n(A) = \frac{m}{n}$ is the relative frequency of A in the n trials. Then,

$$\mathbb{P}(A) = \lim_{n \to \infty} f_n(A) = \frac{m}{n} \tag{2}$$

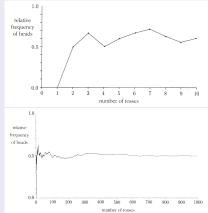
or

$$f_n(A) \approx P(A)$$

for large n.

EXAMPLE 6

Suppose a coin is tossed and recorded the number of head.



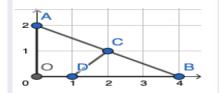
GEOMETRIC PROBABILITY

Consider an experiment with equally likely outcomes, infinite sample space Ω . Let A be an event of this experiment. If both of Ω and A can be illustrated by geometric regions,

$$P(A) = \frac{\text{the measure of the region of } A}{\text{the measure of the region of } \Omega} = \frac{|A|}{|\Omega|}.$$
 (3)

EXAMPLE 7

If a point M is randomly chose inside the triangle OAB. Find the probability that M is in the quadrangle OACD.



AXIOMS OF PROBABILITY

Denote by ${\mathscr A}$ the collection of all events. Probability is a function

$$\mathbb{P}: \mathcal{A} \to [0,1]$$

AXIOMS OF PROBABILITY

Denote by ${\mathscr A}$ the collection of all events. Probability is a function

$$\mathbb{P}: \mathscr{A} \to [0,1]$$

AXIOMS OF PROBABILITY

Denote by ${\mathscr A}$ the collection of all events. Probability is a function

$$\mathbb{P}: \mathscr{A} \to [0,1]$$

AXIOMS OF PROBABILITY

Denote by ${\mathscr A}$ the collection of all events. Probability is a function

$$\mathbb{P}: \mathscr{A} \to [0,1]$$

AXIOMS OF PROBABILITY

Denote by A the collection of all events. Probability is a function

$$\mathbb{P}: \mathscr{A} \to [0,1]$$

such that

- For two disjoint events *A* and *B*

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

For any event $A \in \mathcal{A}$, $\mathbb{P}(A)$ is called the probability of A.

PROPERTIES OF PROBABILITY

- $0 \le \mathbb{P}(A) \le 1.$
- **3** For a sequence of disjoint events $A_1, ..., A_n$

$$\mathbb{P}(A_1 + \dots A_n) = \mathbb{P}(A_1) + \dots + \mathbb{P}(A_n)$$

ADDITION RULE

For two events

$$\mathbb{P}(A+B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(AB)$$

2 For three events

$$\mathbb{P}(A+B+C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(AB) - \mathbb{P}(AC) - \mathbb{P}(BC) + \mathbb{P}(ABC)$$

Note the alternating signs.

In general,

$$\mathbb{P}\left(\sum_{i} A_{i}\right) = \sum_{i} \mathbb{P}(A_{i}) - \sum_{i < j} P(A_{i} A_{j}) + \sum_{i,j,k} \mathbb{P}(A_{i} A_{j} A_{k}) - \dots$$

ADDITION RULE

Example 8

Consider the semiconductor wafer data in the table below.

Location in Sputtering Tool			
Contamination	Center	Edge	Total
Low	514	68	582
High	112	246	358
Total	626	314	

A wafer is randomly selected from the batch. Let H denote the event that the chosen wafer contains high levels of contamination. Let C denote the event that the wafer is in the center of a sputtering tool. Compute $\mathbb{P}(H)$, $\mathbb{P}(C)$, $\mathbb{P}(HC)$ and $\mathbb{P}(H+C)$.

ADDITION RULE

EXAMPLE 9

In a big company, they announced a new requirement that staffs have to able to speak at least one of three foreign languages including English, Chinese, or Japanese. Then they have reported as below:

There are respectively 46.45, 40, and 42.58 percentage of staffs can speak English, Chinese and Japanese. Furthermore, 12.9% of them can speak both English and Chinese, 15.48% of them can speak both English and Japanese, 12.9% of them can speak both Japanese and Chinese. Also, 5.16% of staffs able to speak all three languages. Compute the percentage of staffs confirm this new requirement.

A communications channel has an error rate of 1 per 1000 bits transmitted. Errors are rare, but do tend to occur in bursts. If a bit is in error, the probability that the next bit is also in error is greater that. \Rightarrow Sometimes probabilities need to be reevaluated as additional information becomes available.

Definition 2.1

The probability of an event B under the knowledge that the outcome will be in event A is denoted as P(B|A) and this is called the conditional probability of B given A.

FORMULA OF CONDITIONAL PROBABILITY

Let two events *A* and *B* where $\mathbb{P}(B) > 0$. The conditional probability of event *A* given event *B* is

$$P(A|B) = \frac{P(AB)}{P(B)}, \quad P(B) > 0$$
 (4)

Similarly, where P(A) > 0, the conditional probability of event B given event A is

$$P(B|A) = \frac{P(AB)}{P(A)}, \quad P(A) > 0$$
 (5)

EXAMPLE 10

An example of 400 parts classified by surface flaws and as (functionally) defective.

		Surface Flaws		
		Yes (event F)	No	Total
Defective	Yes (event D)	10	18	28
	No	30	342	372
	Total	40	360	400

Suppose that a part is chosen randomly, compute the following probability.

- The chosen part is defective.
- The chosen part is defective given that it has surface flaws.
- The chosen part is defective given that it has no surface flaws.

PROPERTIES OF CONDITIONAL PROBABILITY

- $0 \le P(A|B) \le 1$
- **2** P(B|B) = 1
- **3** If $AC = \emptyset$, P[(A + C)|B] = P(A|B) + P(C|B)
- **1** $P(\bar{A}|B) = 1 P(A|B)$

• For two events A and B

$$\mathbb{P}(AB) = \mathbb{P}(A|B)\mathbb{P}(B) = \mathbb{P}(B|A)P(A)$$

For two events A and B

$$\mathbb{P}(AB) = \mathbb{P}(A|B)\mathbb{P}(B) = \mathbb{P}(B|A)P(A)$$

• For three events *A*, *B* and *C*

$$\mathbb{P}(ABC) = \mathbb{P}(A|BC)\mathbb{P}(B|C)\mathbb{P}(C)$$

For two events A and B

$$\mathbb{P}(AB) = \mathbb{P}(A|B)\mathbb{P}(B) = \mathbb{P}(B|A)P(A)$$

• For three events *A*, *B* and *C*

$$\mathbb{P}(ABC) = \mathbb{P}(A|BC)\mathbb{P}(B|C)\mathbb{P}(C)$$

In general,

$$\mathbb{P}(A_1 A_2 \dots A_n) = \mathbb{P}(A_n | A_1 \dots A_{n-1}) \mathbb{P}(A_{n-1} | A_1 \dots A_{n-2}) \dots \mathbb{P}(A_2 | A_1) \mathbb{P}(A_1)$$

EXAMPLE 11

The bin contains 3 defective parts and 47 nondefective parts.

- Two parts are randomly chosen from the bin. What is the probability that the second part is defective given that the first part is defective?
- Three parts are randomly chosen from the bin. What is the probability that the first two parts selected are defective and the third is not defective?

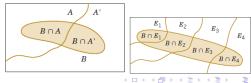
TOTAL PROBABILITY FORMULA

• For two events *A* and *B*,

$$P(A) = P(A|B)P(B) + P(A|\overline{B})P(\overline{B})$$

• In general, assume $A_1, A_2, ..., A_n$ are n mutually exclusive and exhaustive sets, e.i $A_i \cap A_j = \emptyset$ and $\bigcup_i A_i = \Omega$. Then,

$$P(B) = P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + \dots + P(B|A_n)P(A_n)$$
$$= \sum_{i=1}^{n} P(A_i)P(B|A_i)$$



TOTAL PROBABILITY FORMULA

EXAMPLE 12

Information about product failure based on chip manufacturing process contamination is given below. Find the probability of failure.

Probability of Failure	Level of Contamination	Probability of Level
0.1	High	0.2
0.005	Not high	0.8

BAYSE FORMULA

• For two events *A* and *B*.Then,

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

• In general, assume $A_1, A_2, ..., A_n$ are n mutually exclusive and exhaustive sets, e.i $A_i \cap A_j = \emptyset$ and $\bigcup_i A_i = \Omega$. Then,

$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(A_i)\mathbb{P}(B|A_i)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A_i)\mathbb{P}(B|A_i)}{\sum_{i=1}^n \mathbb{P}(A_i)\mathbb{P}(B|A_i)}$$

BAYSE FORMULA

EXAMPLE 13

Because a new medical procedure has been shown to be effective in the early detection of an illness, a medical screening of the population is proposed. The probability that the test correctly identifies someone with the illness as positive is 0.99, and the probability that the test correctly identifies someone without the illness as negative is 0.95. The incidence of the illness in the general population is 0.0001. If a person takes the test, compute the following probabilities.

- Mis result is positive.
- He has the illness in case the test shows a positive result.
- The test is performing correctly.

BAYSE FORMULA

EXAMPLE 14

The conditional probability that a high level of contamination was present when a failure occurred is to be determined. The information is summarized in the table below.

Probability of Failure	Level of Contamination	Probability of Level
0.1	High	0.2
0.005	Not high	0.8

Suppose that a failed product has been taken, compute the chance that it has high levels of contamination.

INDEPENDENCE

- Two events are independent if any one of the following equivalent statements is true:
- The event A_1 , A_2 , ..., A_n are independent if and only if for any subset of these events

$$\mathbb{P}(A_{i_1}A_{i_2}\dots A_{i_k}) = \mathbb{P}(A_{i_1})\mathbb{P}(A_{i_2})\dots\mathbb{P}(A_{i_k})$$



INDEPENDENCE

EXAMPLE 15

Disks of polycarbonate plastic from a supplier are analyzed for scratch and shock resistance. The results from 100 disks are summarized as follows:

		Shock Resistance	
		High	Low
Scratch	High	70	9
Resistance	Low	16	5

Let *A* denote the event that a disk has high shock resistance, and let *B* denote the event that a disk has high scratch resistance. Are events *A* and *B* independent?

INDEPENDENCE

EXAMPLE 16

The following circuit operates if and only if there is a path of functional devices from left to right. The probabilities that each device functions are as shown where p = 0.9. Assume that the probability that a device is functional does not depend on whether or not other devices are functional. What is the probability that the circuit operates?

