

Ministry of Higher Education and Scientific Research

University Mustapha Stambouli of Mascara

Faculty of Exact Sciences

Department of Mathematics



INTRODUCTION TO PROBABILITY AND DESCRIPTIVE STATISTICS



This course is designed for first-year Bachelor's students in Mathematics.

Dr. Ahmed Bouchenak

Email: a.bouchenak@univ-mascara.dz / bouchenakmath@gmail.com

April, 2026

LIST OF CONTENTS

List of Contents	ii
List of Figures	iii
Introduction	1
1 BASIC CONCEPTS AND STATISTICAL VOCABULARY	3
1.1 Vocabulary	4
1.1.1 Statistical Test	4
1.1.2 Population	5
1.1.3 Individual (Statistical Unit)	6
1.1.4 Character (Statistical Variable)	6
1.1.5 Modalities (Terms)	7
1.2 Types of Characters	7
1.2.1 Qualitative Character	7
1.2.2 Quantitative Character	9
2 STUDY OF A QUALITATIVE AND DISCREET QUANTITATIVE	

CHARACTER	10
2.1 Partial Number - Cumulative Number	11
2.2 Partial Frequency - Cumulative Frequency	13
2.3 Graphical Representation	15
2.3.1 Graphical Representation of a Qualitative Character	15
2.3.2 Graphical Representation of a Discrete Quantitative Character . . .	17
2.4 Representation as a Curve and Distribution Function	18
2.5 Position Parameters (Central Tendency Characteristic)	20
2.5.1 Mode - Median	20
2.5.2 Averages	22
2.5.3 Quartiles	24
2.6 Dispersion Parameters (Variability Characteristic)	26
2.6.1 Range - Variance - Standard Deviation - Coefficient of Variation . .	27
2.6.2 Average Absolute Deviation	30
3 STUDY OF A CONTINUOUS QUANTITATIVE CHARACTER	32
3.1 Continuous Character	32
3.2 Graphic Representation of a Continuous Character	37
3.3 Position Parameters (Central Tendency Characteristic)	40
3.4 Dispersion Parameters (Variability Characteristic)	46
4 Probability Calculus	50
4.1 Combinatorial Analysis	51
4.1.1 Fundamental Principle of Combinatorial Analysis	51
4.1.2 Arrangements	53
4.1.3 Permutations	55
4.1.4 Combinations	56

4.2	Basic Concepts	56
4.2.1	Random experiment, Outcome and Sample space	56
4.2.2	Events	57
4.2.3	Operations on Events	59
4.2.4	Complete System of Events	63
4.3	Probability Measure	64
4.3.1	Probabilizable Space and Probability Function	64
4.3.2	Probability Space and Probability of an Event	65
4.4	Conditional Probability	66
4.4.1	Compound probability	67
4.4.2	Independence and Conditional Probability	67
4.4.3	Bayes' Theorem	68
	References	71

LIST OF FIGURES

1.1	Vocabulary	4
1.2	Types of characters	8
2.1	The number of individuals who take the value x_i	12
2.2	Some characteristics of the graph	16
2.3	Bar diagram	16
2.4	Pie diagram	17
2.5	Bar chart	18
2.6	Representation of a discrete quantitative variable by the cumulative curve.	20
3.1	Characterization for a class values.	33
3.2	A representation of the distribution of values within a class.	34
3.3	Histogram of frequencies (or repetitions).	38
3.4	Histogram of repetitions	38
4.1	Illustration of the 9 arrangements of order 2 from the set $\{a, b, c\}$	54
4.2	Operations on events.	60

INTRODUCTION

Statistics and probability constitute fundamental pillars in the mathematical sciences, providing powerful tools for understanding variability, interpreting real-world data, and modeling uncertain phenomena. Whether in economics, engineering, biology, social sciences, medicine, or artificial intelligence, statistical reasoning has become an essential skill for analyzing information and making informed decisions.

This course offers a progressive and structured introduction to the main concepts of descriptive statistics and probability calculus. It is designed for students encountering statistics for the first time, as well as those seeking to consolidate their foundational knowledge before advancing to more theoretical or applied modules.

The first part is devoted to the basic vocabulary of statistics, where we introduce key notions such as population, sample, statistical variables, and their modalities. We also develop a precise classification of variable types, which is essential before any analysis.

The second part focuses on the study of qualitative and discrete quantitative characters. Here, the student learns how to build and interpret statistical tables, frequencies, cumulative distributions, and how to represent data using appropriate graphs. Measures of central tendency and dispersion are introduced in detail to help describe and summarize

a dataset effectively.

The third part extends these ideas to continuous quantitative characters. Through class intervals, histograms, cumulative curves, and distribution functions, the student gains experience in handling continuous data and extracting meaningful characteristics from them.

The final part of the course provides a rigorous introduction to probability calculus. We explore combinatorial analysis, sample spaces, events, and operations on events, followed by the axiomatic definition of probability. The course culminates with conditional probability, independence of events, and Bayes' theorem — fundamental concepts for studying random phenomena and preparing for more advanced topics such as random variables, distributions, and inferential statistics.

Throughout the course, numerous corrected examples and additional problems are provided to reinforce learning, develop analytical skills, and offer practical situations where statistical and probabilistic reasoning must be applied.

This material aims not only to present the theoretical foundations of statistics and probability but also to foster intuition, critical thinking, and correct interpretation of data — indispensable competencies in the modern world.

CHAPTER 1

BASIC CONCEPTS AND STATISTICAL VOCABULARY

Statistics is the study of the collection of data, their analysis, their processing, the interpretation of the results and their presentation in order to make the data understandable by everyone. It is at the same time a science, a method and a set of techniques.

Data analysis is used to write down the phenomena studied, make predictions and make decisions about them. In this way, statistics is an essential tool for understanding and managing complex phenomena.

The data studied can be of any nature, which makes statistics useful in all disciplinary fields and explains why it is taught in all university courses, from mathematics to economics and to biology through psychology and of course engineering sciences. The statistics consist of:

- Collect data.
- Present and summarize this data.

- Draw conclusions about the population studied and assist in decision making.
- In the presence of time dependent data, we try to make predictions.

1.1 Vocabulary

Statistics consists of various methods of classifying data such as tables, histograms and graphs, to organize a large amount of data. Statistics developed in the second half of the 19th century in the field of human sciences (sociology, economics, anthropology, etc.). They have adopted a particular vocabulary.

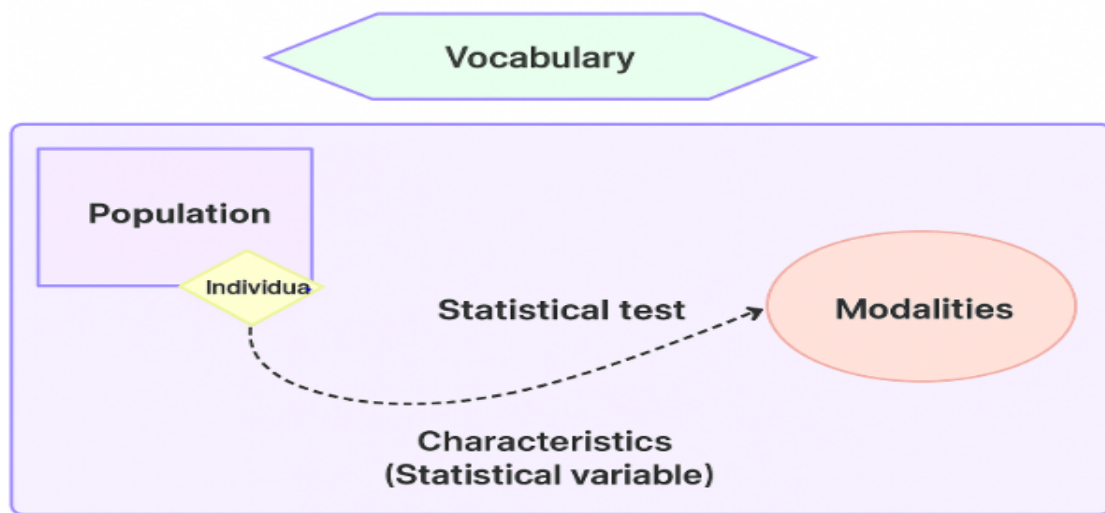


Figure 1.1: Vocabulary

1.1.1 Statistical Test

Descriptive statistics aim to study the characteristics of a set of observations such as the measurements obtained during an experiment. The experiment is the preliminary step to any statistical study. It's about making "contact" with the observations. Generally speaking, the statistical method is based on the following concept.

Definition 1.1.1. *The statistical test is an experiment that we provoke.*

Example 1.1.2. (*The lifespan of the lamps*)

Let's imagine the following case: a manufacturer of electric lamps having the choice between 4 types of filaments intends to study the influence of the nature of the filament on the lifespan of the lamps manufactured. To do this, he will have 4 samples of identical lamps made, except for the filament, burn the lamps until they go out, then compare the results obtained.

1.1.2 Population

In statistics, we work on populations. This term comes from the fact that demography, the study of human populations, occupied a central place in the beginnings of statistics, particularly through population censuses. But, in statistics, the term population applies to any statistical object studied, whether students (of a university or a country), households or any other group on which we make statistical observations. We define the notion of population.

Definition 1.1.3. *We call population the group on which our statistical study relates. This set is denoted Ω .*

Example 1.1.4.

– *We consider all the students in section A. We are interested in the number of brothers and sisters of each student. In this case*

$$\Omega = \text{all students.}$$

– *If we now look at automobile traffic in a city, the population is then made up of all the vehicles likely to be circulating in this city on a given date. In this case*

$$\Omega = \text{all vehicles.}$$

1.1.3 Individual (Statistical Unit)

A population is made up of individuals. The individuals who make up a statistical population are called statistical units.

Definition 1.1.5. *We call individual any element of the population Ω , it is denoted w (w in Ω).*

Remark 1.1.6. *The set Ω can be a set of people, things or animals...,etc.*

The statistical unit is an object for which we are interested in collecting information.

Example 1.1.7.

- *In the example shown above, an individual is any student in the section.*
- *If we study the annual production of a factory of metal beverage cans (cans). The population is all the boxes produced during the year and one box constitutes an individual.*

1.1.4 Character (Statistical Variable)

“Descriptive” statistics, as its name suggests, seeks to describe a given population. We are interested in the characteristic of units which can take different values.

Definition 1.1.8. *We call character (or statistical variable, denoted $S.V$) any application*

$$X : \Omega \rightarrow C$$

The set C is said to be set of values of the character X (this is what is measured or observed on individuals)

Example 1.1.9.

Height, temperature, nationality, eye color, socio-professional category...,etc.

Remark 1.1.10. Let Ω be a set. We call and note $Card(\Omega)$, the number of elements of Ω .

$$Card(\Omega) := \text{number of elements of } \Omega = N.$$

1.1.5 Modalities (Terms)

Definition 1.1.11. The modalities (terms) of a statistical variable are the different values that it can take.

Example 1.1.12.

– Variable is “family situation”

Terms are “single, married, divorced”

– Variable is “switch status”

Modalities are “0 and 1”.

– Variable is “socio-professional categories”

Terms are “Employees, workers, retirees,...,etc.”

The modalities are the different situations in which individuals can find themselves with regard to the character considered.

1.2 Types of Characters

We distinguish two categories of characteristics: qualitative characteristics and quantitative characteristics.

1.2.1 Qualitative Character

Qualitative characters are those whose modalities cannot be ordered, that is to say if we consider two characters taken at random, we cannot say of one of the characters that it is inferior or equal to the other. More precisely, we have the following definition:

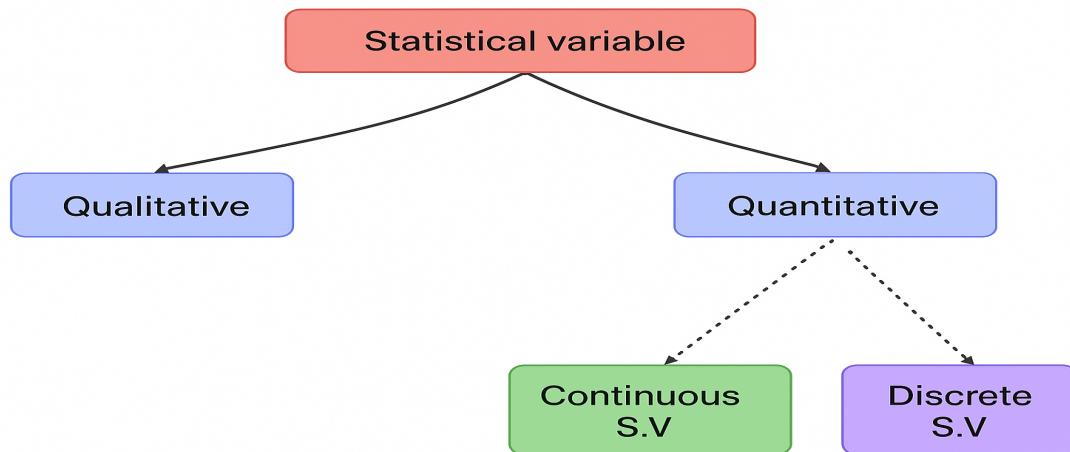


Figure 1.2: Types of characters

Definition 1.2.1. *The elements of C are represented by something other than numbers.*

Example 1.2.2.

The condition of a house: we can consider the following modalities:

- *Old.*
- *Degraded.*
- *New.*
- *Renovated.*

Remark 1.2.3. (nominal - ordinal character)

We distinguish two types of qualitative statistical variables: either ordinal or nominal (non-ordinal).

Example 1.2.4.

- *The eyes color (black, brown, blue, ...) of the students in this section is nominal qualitative character.*
- *The mention of the Baccalaureate (excellent, very good, good, ...) of the students in this section is ordinal qualitative character.*

1.2.2 Quantitative Character

Quantitative characters are characters whose modalities can be ordered. Thus, the age, life size or salary of an individual are quantitative characters. So we have the following definition:

Definition 1.2.5. *All values are represented by numbers. Likewise, it is divided into two kinds of characters, discrete and continuous (see the example).*

Example 1.2.6.

– *The salary of factory employees.*

Terms: 10000da, 20000da..., etc.

Type: Discreet.

– *The rigidity of the springs.*

Modalities: $[10, 20]$ N/m

Type: continuous.

In general, the discrete quantitative variable is a variable that only takes integer values (more rarely decimal). The number of distinct values of such a variable is usually quite small. Let us cite, for example, the number of houses per neighborhood of a city. A quantitative variable is said to be continuous when the observations associated with it are not precise values, but intervals. This is the case when we have a large number of distinct observations.

The aim of descriptive statistics is to synthesize the information contained in data sets using tables, figures or numerical summaries. Statistical variables are analyzed differently depending on their nature (quantitative, qualitative).

CHAPTER 2

STUDY OF A QUALITATIVE AND DISCREET QUANTITATIVE CHARACTER

The statistical character can take a reasonable finite number of values (score, number of children, number of rooms,...etc.). In this case, the statistical character studied is then called a discrete character.

Throughout the rest of the chapter, we consider the following situation:

$$X : \Omega \rightarrow \{x_1, x_2, \dots, x_k\},$$

with $Card(\Omega) := N$ is the number of individuals in our study.

We will often use the example below as a general example to illustrate the statements in this chapter.

Example 2.0.1. *A survey carried out in a village concerns the number of dependent children per family. We note X the number of children, the results are given by this table:*

x_i (term)	0	1	2	3	4	5	6
n_i (number)	18	32	66	41	32	9	2

We have

- Ω all families.
- w a family.
- X number of children per family

$$X : w \rightarrow X(w).$$

We read, to family w , we associate $X(w) =$ the number of children in this family.

2.1 Partial Number - Cumulative Number

We study here a numerical statistical character represented by a sequence x_i describing the value of the character with i varies from 1 to k .

Definition 2.1.1. (*Partial number (Partial repetition)*)

For each value x_i , we set by definition

$$n_i = \text{Card}\{w \in \Omega : X(w) = x_i\}.$$

n_i : *the number of individuals who have the same x_i , this is called number or (partial number) of x_i .*

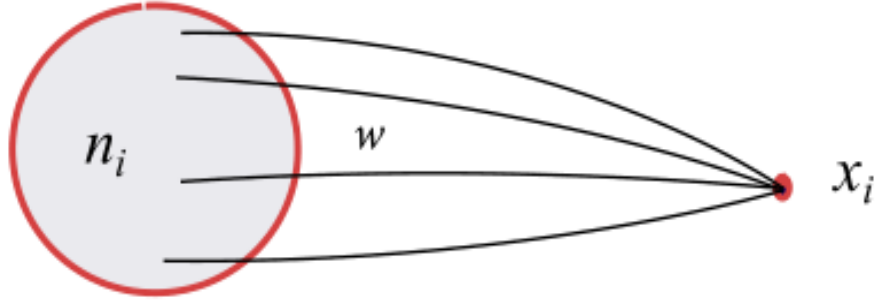


Figure 2.1: The number of individuals who take the value x_i

Example 2.1.2.

In the previous example, 66 is the number of families (repetition) that have 2 children.

x_i (term)	...	2	...
n_i (number)	...	66	...

Definition 2.1.3. (Cumulative number)

For each value x_i , we set by definition

$$N_i = n_1 + n_2 + \dots + n_i = \sum_{j=1}^i n_j.$$

The cumulative number N_i of a value is the sum of the repetition of this value and all the repetitions of the previous values.

Example 2.1.4.

In the previous example: 50 is the number of families which have a number of children less than 1. We look at it in the following table:

x_i (term)	0	1	2	3	4	5	6
N_i (cumulative number)	18	50	116	157	189	198	200

Interpretation:

N_i is the number of individuals whose character value is less than or equal to x_i . Therefore,

the total number is given by

$$N = \text{Card}(\Omega) = \sum_{i=1}^k n_i.$$

In our previous example, we have $N = 200$.

2.2 Partial Frequency - Cumulative Frequency

Typically the numbers are large and it is interesting to calculate quantities allowing the series to be summarized.

Definition 2.2.1. (*Partial frequency*)

For each value x_i , we set by definition

$$f_i = \frac{n_i}{N}.$$

f_i is called the partial frequency of x_i . The frequency of a value is the ratio (division) of the number of this value to the total number.

Remark 2.2.2. We can replace f_i by $f_i \times 100$ which then represents a percentage.

Interpretation:

f_i is the percentage of w such that $X(w) = x_i$.

Example 2.2.3.

In the previous example, $0.33 :=$ there are 33% of families whose number of children equals

2. This percentage is calculated as follows ($N = 200$):

x_i (term)	...	2	...
n_i (number)	...	66	...
f_i (frequency)	...	$\frac{66}{200} = 0.33$	...

We can conclude the following property:

Proposition 2.2.4. *Let f_i be defined as previously. So,*

$$\sum_{i=1}^k f_i = 1.$$

Proof : Let's remember that

$$\sum_{i=1}^k n_i = N.$$

Which implies

$$\sum_{i=1}^k f_i = \sum_{i=1}^k \frac{n_i}{N} = \frac{1}{N} \sum_{i=1}^k n_i = \frac{1}{N} \cdot N = 1.$$

□

Definition 2.2.5. (Cumulative frequency)

For each value x_i , we set by definition

$$F_i = f_1 + f_2 + \dots + f_i = \sum_{j=1}^i f_j.$$

The quantity F_i is called the cumulative frequency of x_i .

Interpretation:

F_i = is the percentage of w such that the value $X(w)$ is less than or equal to x_i .

Example 2.2.6.

In the previous example, 0.785 represents 78.5% of families whose number of children is less than or equal to 3.

x_i (term)	0	1	2	3	4	5	6
n_i (number)	18	32	66	41	32	9	2
F_i (cumulative number)	$\frac{18}{200}$	$\frac{50}{200}$	$\frac{116}{200}$	$\frac{157}{200} = 0.785$	$\frac{189}{200}$	$\frac{198}{200}$	$\frac{200}{200} = 1$

We have seen that tables are an often essential, in any case very useful, means of classifying and presenting the units of a statistical population. In the following paragraph, we will see how we translate our tables into graphs that also allow us to summarize the data in a visual way.

2.3 Graphical Representation

We distinguish the methods of representing a statistical variable depending on the nature of this variable (qualitative or quantitative). The recommended and most frequent representations are diagrams (graphs).

The graphic is a visual support that allows:

The summary: visualize the main characteristics at a glance (but we lose a quantity of information), see Figure 2.2.

Discovery: highlights trends.

Control: we see anomalies better on a graph than in a table.

The search for regularities: regularity in the movement, repetition of the phenomenon.

2.3.1 Graphical Representation of a Qualitative Character

From the observation of a qualitative variable, two diagrams make it possible to represent this variable: the bar diagram (called organ pipes) and the angular sector diagram (called pie diagram).

Bar diagram:

We plot the modalities on the abscissa, arbitrarily. We carry rectangles on the ordinate whose length is proportional to the numbers, or frequencies, of each modality.

Pie diagram:

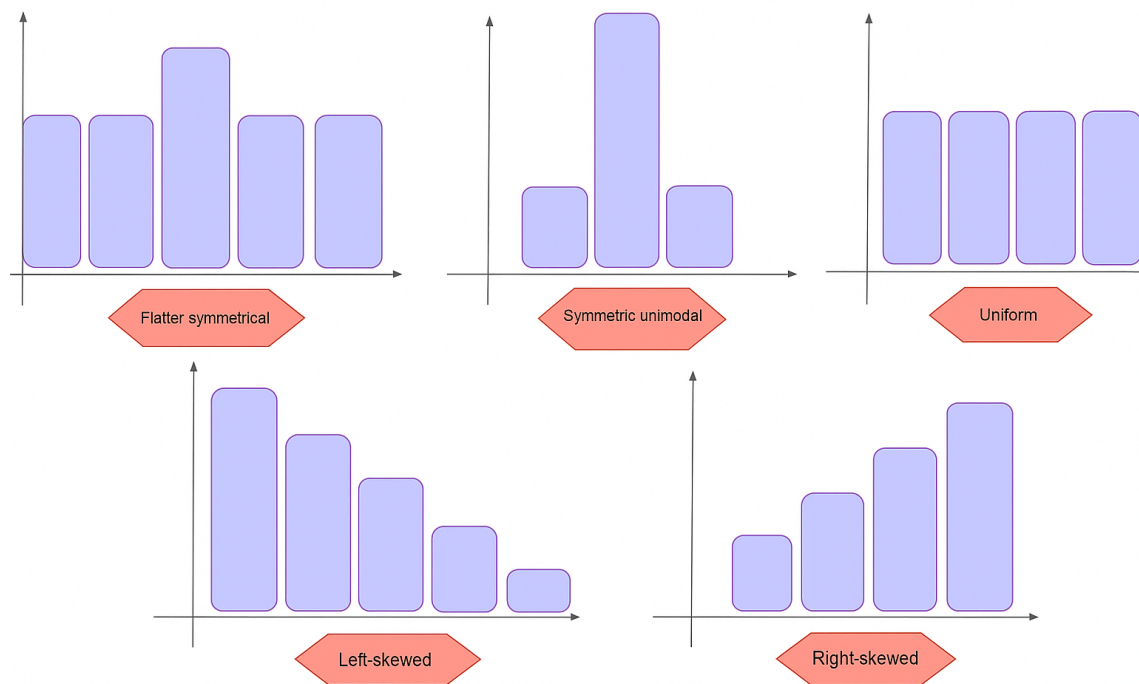


Figure 2.2: Some characteristics of the graph

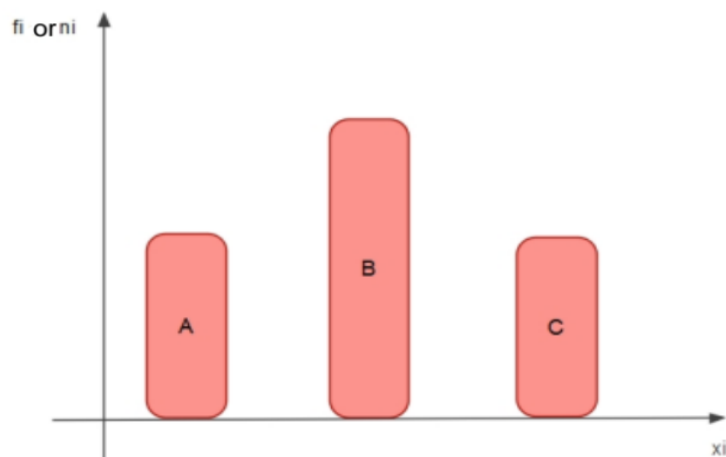


Figure 2.3: Bar diagram

Circular, or semi-circular, diagrams consist of dividing a disk or half-disk, into slices, or sectors, corresponding to the modalities observed and whose surface area is proportional to the number, or frequency, of the modality.

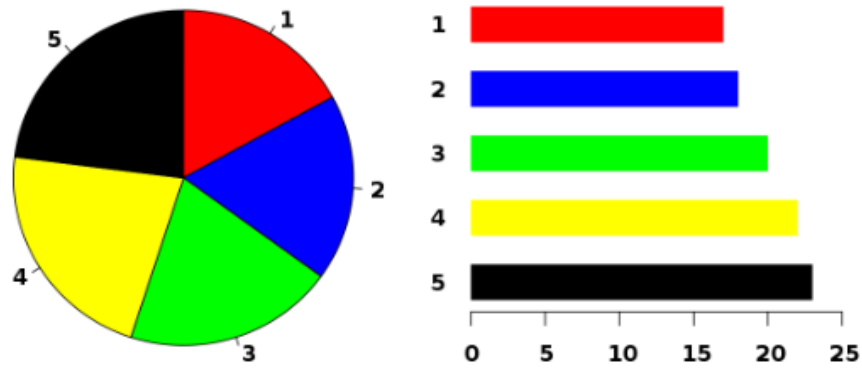


Figure 2.4: Pie diagram

The degree of a sector is determined using the rule of three as follows:

$$N \longrightarrow 360$$

$$n_i \longrightarrow d_i \text{ (degree of modality } i\text{)}.$$

Hence,

$$d_i = \frac{n_i \times 360}{N}.$$

2.3.2 Graphical Representation of a Discrete Quantitative Character

From the observation of a discrete quantitative variable, two diagrams make it possible to represent this variable: the Bar chart and the cumulative curve which will be presented in the next section.

For the illustration, we take the previous example of departure (number of children per family). We recall the associated statistical table.

x_i (<i>term</i>)	0	1	2	3	4	5	6
n_i (<i>number</i>)	18	32	66	41	32	9	2

Bar Chart

We want to represent this distribution in the form of a bar chart. Each mark corresponds to a stick. The heights of the sticks are proportional to the numbers represented (see Figure 2.5).

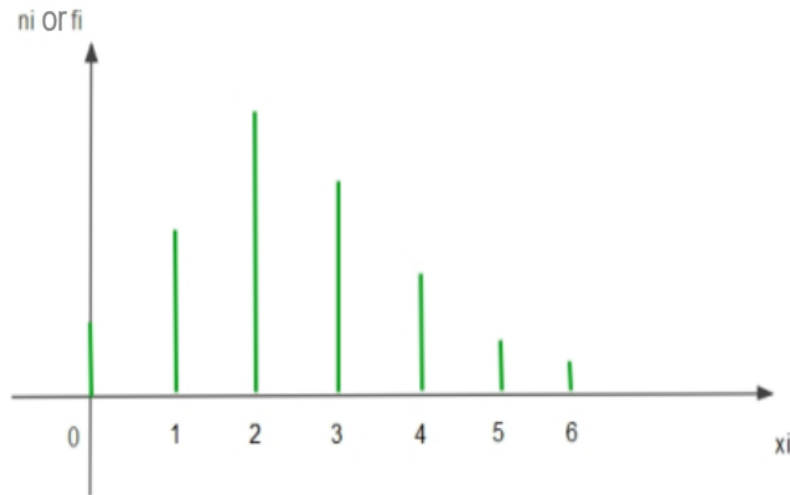


Figure 2.5: Bar chart

2.4 Representation as a Curve and Distribution Function

We have already discussed the cumulative distributions of a statistical variable. In this part, we will exploit its cumulative values to introduce the notion of the distribution function. This notion only concerns quantitative variables.

Definition 2.4.1. *Let the function $F_x : \mathbb{R} \rightarrow [0, 1]$ be defined by*

$$F_x := \text{percentage of individuals whose character value is less than or equal } x.$$

This function is called the distribution function of character X .

Remark 2.4.2. For all $i \in \{1, \dots, k\}$, we have

$$F_x(x_i) = F_i.$$

The F_x curve passes through the points $(x_1, F_1), (x_2, F_2), \dots$ and (x_k, F_k) .

Proposition 2.4.3. The distribution function satisfies, for $i \in \{1, \dots, k\}$,

$$\begin{aligned}
 & - \text{Equality, } F_x(x_i) = F_i, \\
 & - \text{The expression, } F_x(x) = \begin{cases} 0 & \text{if } x < x_1, \\ F_1 & \text{if } x_1 \leq x < x_2, \\ F_i & \text{if } x_i \leq x < x_{i+1}, \\ \dots & \text{if } \dots \leq x < \dots, \\ 1 & \text{if } x \geq x_k. \end{cases}
 \end{aligned}$$

Example 2.4.4.

Based on our example, the F_x curve is shown below (Figure 2.6) on

$\mathbb{R} =] - \infty, 0[\cup [0, 1[\cup \dots \cup [6, +\infty[$. In this case we have

$$F_x(x) = \begin{cases} 0 & \text{if } x < 0, \\ 0.09 & \text{if } 0 \leq x < 1, \\ 0.25 & \text{if } 1 \leq x < 2, \\ 0.58 & \text{if } 2 \leq x < 3, \\ 0.785 & \text{if } 3 \leq x < 4, \\ 0.945 & \text{if } 4 \leq x < 5, \\ 1 & \text{if } x \geq x_k. \end{cases}$$

This curve is called "the cumulative frequency curve". The cumulative curve is a step

curve representing relative cumulative frequencies.

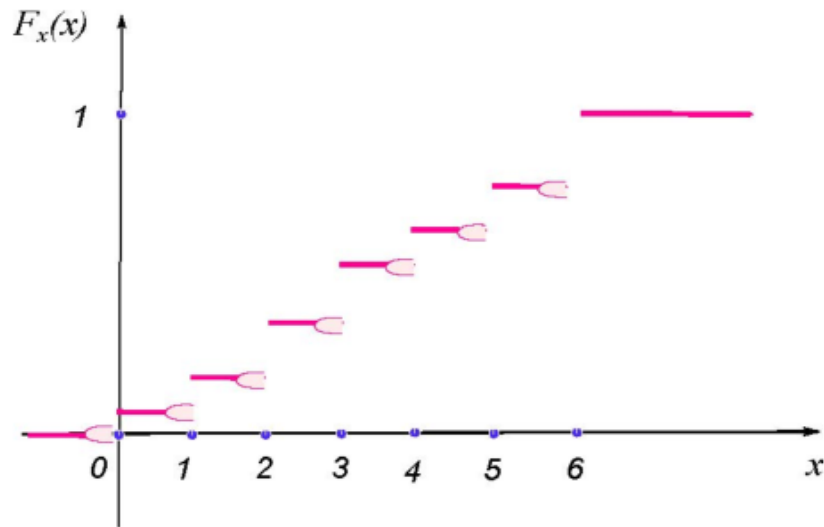


Figure 2.6: Representation of a discrete quantitative variable by the cumulative curve.

2.5 Position Parameters (Central Tendency Characteristic)

The statistical indicators of central tendency (also known as position parameters) frequently considered are the mode and the median, the averages (arithmetic, geometric, harmonic, quadratic) and the quartiles (the first and the third).

2.5.1 Mode - Median

Definition 2.5.1. (*Mode*)

The mode of a statistical variable is the value which has the greatest partial repetition (or the greatest partial frequency) and it is denoted by Mo .

Example 2.5.2.

In the previous example, the mode is equal to 2 which corresponds to the greatest number

(repetition).

x_i (term)	0	1	2	3	4	5	6
n_i (number)	18	32	66	41	32	9	2

Remark 2.5.3. The mode is not unique. ie, (We can have more than one mode).

Definition 2.5.4. (Median)

We call median the value Me of the statistical variable X which verifies the following relationship:

$$F_x(Me^-) < 0.5 \leq F_x(Me^+) = F_x(Me).$$

We can use F_i (cumulative frequencies) instead of F_x (distribution function).

The median divides the statistical series into two groups of the same size (repetition).

Example 2.5.5.

In the previous example, the relationship

$$F_x(0) = 0 < 0.5 \leq F_x(0^+) = 0.09.$$

is not satisfied. Therefore, the median is different from 0.

On the other hand, we have

$$F_x(2^-) = 0.25 < 0.5 \leq F_x(2^+) = F(2) = 0.58$$

Hence $Me = 2$.

2.5.2 Averages

Definition 2.5.6. (*Arithmetic Mean*)

We call the average (the arithmetic mean) of x , the amount

$$\bar{x} = \frac{1}{N} \sum_{i=1}^k n_i x_i = \sum_{i=1}^k f_i x_i,$$

with $N = \text{Card}(\Omega) = \sum_{i=1}^k n_i$. We can therefore express and calculate the arithmetic mean with numbers or with frequencies.

Example 2.5.7.

In our general example. We have,

$$\bar{x} = \frac{1}{N} \sum_{i=1}^k n_i x_i = \frac{1}{200} (0 \times 18 + 1 \times 32 + 2 \times 66 + \dots + 2 \times 6) = 2.36$$

If $\bar{x} = 2.36$, then we have on average a neighborhood family with 2.36 children.

Remark 2.5.8. The value of the average is abstract. As in the previous example, $\bar{x} = 2.36$ is a number that does not correspond to a concrete fact.

The arithmetic mean whose formula has just been indicated is called the weighted average; this means that each value of the variable is multiplied (weighted) by a coefficient, here by the number corresponding to it. In this case, each value x_i of the variable intervenes in the calculation of the average as many times as it has been observed. We speak of a simple arithmetic average when we do not carry out weighting. For example, if 5 students have respective ages of 18, 19, 20, 21 and 22, their average age is given by $(18 + 19 + 20 + 21 + 22)/5 = 20$ years.

Remark 2.5.9. We should mention here that there are other means than the arithmetic mean.

Definition 2.5.10. (*Geometric mean*)

We call the geometric mean of x , the amount

$$\bar{x}_G = 10^{\left(\frac{1}{N} \sum_{i=1}^k n_i \log(x_i) \right)}$$

Definition 2.5.11. (*Harmonic mean*)

We call the harmonic mean of x , the amount

$$\bar{x}_H = \frac{N}{\sum_{i=1}^k \frac{n_i}{x_i}}$$

Definition 2.5.12. (*Quadratic mean*)

We call the quadratic mean of x , the amount

$$\bar{x}_Q = \sqrt{\frac{\sum_{i=1}^k n_i x_i^2}{N}}$$

Example 2.5.13. If we consider the following statistical series, we get:

x_i	1	2	3	4	5	6	Σ
n_i (repetition)	32	66	41	32	9	2	182
$n_i \log(x_i)$	0	19.86	19.56	29.26	6.29	1.55	66.52
$\frac{n_i}{x_i}$	32	33	13.66	8	1.8	0.33	88.79
$n_i x_i^2$	32	264	369	512	225	72	1474

- The geometric mean is giving as:

$$\bar{x}_G = 10^{\left(\frac{1}{N} \sum_{i=1}^k n_i \log(x_i) \right)} = 10^{\left(\frac{66.52}{182} \right)} = 10^{(0.36)} = 2.29$$

- The harmonic mean is giving as

$$\bar{x}_H = \frac{N}{\sum_{i=1}^k \frac{n_i}{x_i}} = \frac{182}{88.79} = 2.04$$

- The quadratic mean is giving as

$$\bar{x}_Q = \sqrt{\frac{\sum_{i=1}^k n_i x_i^2}{N}} = \sqrt{\frac{1474}{182}} = \sqrt{8.09} = 2.84$$

2.5.3 Quartiles

In statistics, Quartiles are the set of values which has three points dividing the data set into four identical parts. We ordinarily deal with a large amount of numerical data, in stats. There are several concepts and formulas, which are extensively applicable in various researches and surveys.

Definition 2.5.14. (*First quartile*)

The first quartile denoted Q_1 is the median of the lower half of the data set, and it is determined by its position in the cumulative repetition ($N/4$).

Or just it is the value Q_1 of the statistical variable X which verifies the following relationship:

$$F_x(Q_1^-) < 0.25 \leq F_x(Q_1^+) = F_x(Q_1).$$

We can use F_i (cumulative frequencies) instead of F_x (distribution function).

Remark 2.5.15. (*Second Quartile*)

The median is considered as the second quartile (Q_2)

Definition 2.5.16. (Third Quartile)

The third quartile denoted (Q_3) is the median of the upper half of the data, and it is determined by its position in the cumulative repetition ($3N/4$).

Or just it is the value Q_3 of the statistical variable X which verifies the following relationship:

$$F_x(Q_3^-) < 0.75 \leq F_x(Q_3^+) = F_x(Q_3).$$

We can use F_i (cumulative frequencies) instead of F_x (distribution function).

Definition 2.5.17. (Interquartile Range)

Interquartile Range denoted (IQR) is the difference between the third quartile and first quartile is known as the interquartile range. (i.e.) $IQR = Q_3 - Q_1$

Example 2.5.18. In the general example, we get the statistical table:

x_i (terms)	0	1	2	3	4	5	6
n_i (repetition)	18	32	66	41	32	9	2
N_i (cumulative repetition)	18	50	116	157	189	198	200
F_i (cumulative number)	0.09	0.25	0.58	0.785	0.945	0.99	1

- The first quartile Q_1 :

We have $\frac{N}{4} = \frac{200}{4} = 50$ which is the position of Q_1 in N_i .

Then, $Q_1 = 1$.

Or, we have the relationship:

$$F_i(Q_1^-) < 0.25 \leq F_i(Q_1^+) = F_i(Q_1).$$

From the table we find

$$F_i(0) = 0.09 < 0.25 \leq F_i(1) = 0.25$$

Hence, $Q_1 = 1$.

- The third quartile Q_3 :

We have $\frac{3N}{4} = \frac{600}{4} = 150$ which is the position of Q_3 in N_i .

Then, $Q_3 = 3$.

Or, we have the relationship:

$$F_i(Q_3^-) < 0.75 \leq F_i(Q_3^+) = F_i(Q_3).$$

From the table we find

$$F_i(2) = 0.58 < 0.75 \leq F_i(3) = 0.785$$

Hence, $Q_3 = 3$.

The interquartile range is given as

$$IQR = Q_3 - Q_1 = 3 - 1 = 2.$$

2.6 Dispersion Parameters (Variability Characteristic)

The common statistical indicators of dispersion are the range, the variance, the standard deviation, the coefficient of variation, the average absolute deviation.

2.6.1 Range - Variance - Standard Deviation - Coefficient of Variation

Definition 2.6.1. (*Extent or Range*)

The difference between the largest value and the smallest value of the character, given by the quantity

$$e = x_{\max} - x_{\min}.$$

is called the extent of the statistical variable X . Calculating the range is very simple. It gives a first idea of the dispersion of the observations. This is a very rudimentary indicator and there are more sophisticated dispersion indicators (see below).

Example 2.6.2. *In the general example the range is given as:*

$$e = x_{\max} - x_{\min} = 6 - 0 = 6.$$

Definition 2.6.3. (*Variance*)

We call the variance of a statistical series X , the number

$$\text{Var}(X) = \frac{1}{N} \sum_{i=1}^k n_i (x_i - \bar{x})^2$$

We say that the variance is the average of the squares of the deviations from the mean $\bar{x}$

The following theorem (König-Huygens Theorem) gives a remarkable identity linking the variance and the mean, sometimes more practical in calculating the variance.

Theorem 2.6.4. (König-Huygens Theorem)

Let (x_i, n_i) be a statistical series with mean $\bar{x}$ and variance $Var(X)$. SO

$$Var(X) = \left(\frac{1}{N} \sum_{i=1}^k n_i x_i^2 \right) - \bar{x}^2$$

Proof : By definition, we have

$$Var(X) = \frac{1}{N} \sum_{i=1}^k n_i (x_i^2 - \bar{x})^2 = \frac{\sum_{i=1}^k n_i (x_i - \bar{x})^2}{\sum_{i=1}^k n_i}.$$

Then

$$Var(X) = \frac{\sum_{i=1}^k n_i (x_i - \bar{x})^2}{\sum_{i=1}^k n_i} = \frac{\sum_{i=1}^k n_i (x_i^2 + \bar{x}^2 - 2x_i \bar{x})}{\sum_{i=1}^k n_i}.$$

By equality, we have

$$Var(X) = \frac{\sum_{i=1}^k n_i x_i^2}{\sum_{i=1}^k n_i} + \frac{\sum_{i=1}^k n_i \bar{x}^2}{\sum_{i=1}^k n_i} - \frac{\sum_{i=1}^k 2n_i x_i \bar{x}}{\sum_{i=1}^k n_i}.$$

Which implies

$$Var(X) = \frac{\sum_{i=1}^k n_i x_i^2}{\sum_{i=1}^k n_i} + \bar{x}^2 - 2\bar{x}\bar{x} = \left(\frac{1}{N} \sum_{i=1}^k n_i x_i^2 \right) - \bar{x}^2.$$

□

Definition 2.6.5. (Standard deviation)

The standard deviation of the statistical variable X is the amount

$$\sigma_X = \sqrt{\text{Var}(X)}.$$

Definition 2.6.6. (Coefficient of variation)

The coefficient of variation of the statistical variable X is giving as

$$Cv = \frac{\sigma_X}{\bar{x}}.$$

Remark 2.6.7. It should be mentioned here that if we have two populations subject to the same statistical study, we say that the population with the smallest variance, standard deviation, or coefficient of variation is the most homogeneous.

Example 2.6.8. In the general example, we recall the associated statistical table.

x_i (term)	0	1	2	3	4	5	6	Σ
n_i (number)	18	32	66	41	32	9	2	200
$n_i x_i$	0	32	132	123	128	45	12	472
$n_i x_i^2$	0	32	264	369	512	225	72	1474

- The arithmetic mean is calculated as:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^k n_i x_i = \frac{472}{200} = 2.36$$

- The variance of this statistical series is determined as:

$$\text{Var}(X) = \left(\frac{1}{N} \sum_{i=1}^k n_i x_i^2 \right) - \bar{x}^2 = \frac{1474}{200} - (2.36)^2 = 1.8004$$

- The standard deviation of the given statistical series is founded on this way:

$$\sigma_X = \sqrt{Var(X)} = \sqrt{1.8004} = 1.34$$

- We conclude this example by calculating the coefficient of variation:

$$Cv = \frac{\sigma_X}{\bar{x}} = \frac{1.34}{2.36} = 0.567$$

2.6.2 Average Absolute Deviation

The average absolute deviation of a data set is the average of the absolute deviations from a central point. It is a summary statistic of statistical dispersion or variability. In the general form, the central point can be a mean, median, mode, or the result of any other measure of central tendency or any reference value related to the given data set.

Definition 2.6.9. (*Average absolute deviation from the arithmetic mean*)

The average absolute deviation from the arithmetic mean is the amount

$$A.A.D_{\bar{x}} = \frac{1}{N} \sum_{i=1}^k n_i |x_i - \bar{x}|$$

Definition 2.6.10. (*Average absolute deviation from the mode*)

The average absolute deviation from the mode is giving by

$$A.A.D_{Mo} = \frac{1}{N} \sum_{i=1}^k n_i |x_i - Mo|$$

Definition 2.6.11. (*Average absolute deviation from the median*)

The average absolute deviation from the median is defined as

$$A.A.D_{Me} = \frac{1}{N} \sum_{i=1}^k n_i |x_i - Me|$$

Example 2.6.12. To illustrate the average absolute deviation, we back to the general example of this chapter, and we consider the following statistical table:

x_i (term)	0	1	2	3	4	5	6	Σ
n_i (number)	18	32	66	41	32	9	2	200
$n_i x_i - \bar{x} $	42.48	43.52	23.76	26.24	52.48	23.76	7.28	219.52
$n_i x_i - Me $	36	32	0	41	64	27	8	208
$n_i x_i - Mo $	36	32	0	41	64	27	8	208

- The average absolute deviation from the arithmetic mean is the amount:

$$A.A.D_{\bar{x}} = \frac{1}{N} \sum_{i=1}^k n_i |x_i - \bar{x}| = \frac{219.52}{200} = 1.0976$$

- The average absolute deviation from the mode and the median is given as:

Since the mode and the median are equals ($Mo = Me = 2$) then we get:

$$A.A.D_{Mo} = A.A.D_{Me} = \frac{1}{N} \sum_{i=1}^k n_i |x_i - Mo| = \frac{208}{200} = 1.04$$

CHAPTER 3

STUDY OF A CONTINUOUS QUANTITATIVE CHARACTER

We recall that a quantitative statistical variable (S.V) concerns a measurable quantity. Its values are numbers expressing a quantity and on which arithmetic operations (addition, multiplication, etc.) have meaning. In this chapter we will focus on a continuous quantitative statistical variable.

3.1 Continuous Character

Definition 3.1.1. *We call continuous statistical variable (or continuous character) any application of Ω and with real values and which takes a "large" number of values (Continuous characters are those which have an infinity of modalities).*

Example 3.1.2. *Let Ω be all of the newborns born at the university hospital of a city during the first 3 months of 2017. We designate by X the weight of the newborns. We*

suppose that

$$x_{\min} = 2.701 \quad \text{and} \quad x_{\max} = 5.001.$$

Question: How to study this character?

Answer: Divide the values taken by X into classes of values.

Definition 3.1.3. (Values of class)

We call a class of values of X an interval of type $[a, b[$ such that $X \in [a, b[$ if and only if $a \leq X(w) < b$, that is, the character values are in the class $[a, b[$.

As soon as a character is identified as continuous, these modalities $C_k = [L_k, L_{k+1}[$ are intervals (classes) with

- L_k : lower bound.
- L_{k+1} : upper bound.
- $a_k = L_{k+1} - L_k$: its amplitude, its step or its length.
- $c_k = x_k = (L_{k+1} + L_k)/2$: its center.

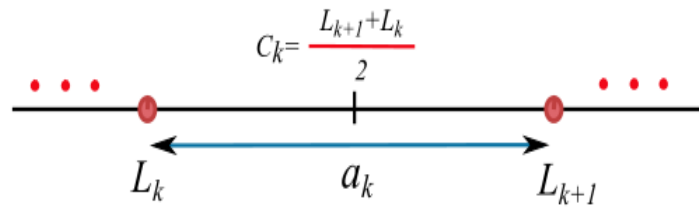


Figure 3.1: Characterization for a class values.

Remark 3.1.4. We will assume in all the cases studied that the distribution within the classes is uniform (see Figure 3.2). This hypothesis makes it possible to justify the fact that we choose the center of the classes as representative.

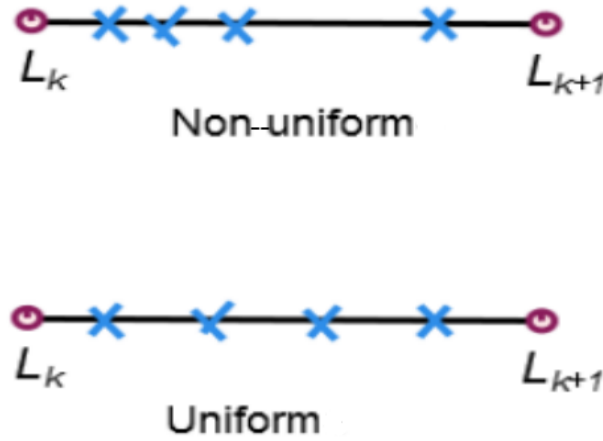


Figure 3.2: A representation of the distribution of values within a class.

Number of classes:

In how many classes do we share values? The answer is not unique, but the following theory explains the sufficient answer.

Theorem 3.1.5. *Let N be the total number in some statistical series contain a continuous character. Then the number of classes k can be found as:*

1) **Classical formula:**

$$k \simeq \sqrt{N}$$

2) **Sturge's formula:**

$$k = 1 + 3.3 \log_{10}(N)$$

3) **Yule formula:**

$$k = 2.5 \sqrt[4]{N}$$

Since k is the number of classes, so it must be integer, and if the result is not integer number we should take the integer part plus one.

Example 3.1.6. *Consider 30 values between 56.5 cm and 97.8 cm. In this case, the number of classes is given as:*

- The classical formula give the result: $k = \sqrt{30}$ and we take $k = 6$.

- The Sturge formula give the result: $k = 1 + 3.3 \log_{10}(30) \simeq 6$,
- The Yule formula give the result: $k = 2.5 \sqrt[4]{30} \simeq 6$.

We mention that the two formulas (Sturge and Yule) are almost the same if $N \leq 200$.

Remark 3.1.7. As a result, we can have several statistical tables depending on the number of classes.

We now recall the definition of extent. Furthermore, in the continuous case we also talk about the step or the length of the class.

Definition 3.1.8. The number $e = x_{\max} - x_{\min}$ is called the range of X . In this case, we can define the amplitude by

$$a_i := \frac{\text{the range}}{\text{the number of classes}} = \frac{x_{\max} - x_{\min}}{k}$$

In the rest of this chapter we will use the bellow example as a general example to illustrate the information provided about this type.

Example 3.1.9. At a plastic tube manufacturer, we took a sample of 100 tubes whose diameter we measured in decimeters.

1.94	2.20	2.33	2.39	2.45	2.50	2.54	2.61	2.66	2.85
1.96	2.21	2.33	2.40	2.46	2.51	2.54	2.62	2.68	2.87
2.07	2.26	2.34	2.40	2.47	2.52	2.55	2.62	2.68	2.90
2.09	2.26	2.34	2.40	2.47	2.52	2.55	2.62	2.68	2.91
2.09	2.28	2.35	2.40	2.48	2.52	2.56	2.62	2.71	2.94
2.12	2.29	2.36	2.41	2.49	2.52	2.56	2.63	2.73	2.95
2.13	2.30	2.37	2.42	2.49	2.53	2.57	2.63	2.75	2.99
2.14	2.31	2.38	2.42	2.49	2.53	2.57	2.65	2.76	2.99
2.19	2.31	2.38	2.42	2.49	2.53	2.59	2.66	2.77	3.09
2.19	2.31	2.38	2.42	2.50	2.54	2.59	2.66	2.78	3.12

1. Identification of this statistical test,

- Population: the tubes.
- Individual: the tube.
- Character: diameter.
- Type: continuous quantitative.
- Terms(modalities): 1.94,..., 3.12.

2. - By the Yule method, we find

$$(\text{number of classes}) = k = 2.5\sqrt[4]{100} = 7.9 \simeq 8.$$

- By Sturge's method, we get

$$(\text{number of classes}) = k = 1 + 3.3 \log_{10}(N) = 1 + 3.3 \log_{10}(100) = 7.6 \simeq 8.$$

We therefore have the amplitude which equals

$$a_i = \frac{x_{\max} - x_{\min}}{k} \simeq 0.15.$$

We obtain the following statistical table, and we need added terms for the next calculation:

C_i	[1.94, 2.09[	[2.09, 2.24[	[2.24, 2.39[	[2.39, 2.54[	[2.54, 2.69[
n_i	3	9	18	29	25

C_i	[2.69, 2.84[	[2.84, 2.99[	[2.99, 3.14[	Σ
n_i	6	6	4	100

3.2 Graphic Representation of a Continuous Character

Histogram of frequencies (or repetitions)

We can represent the statistical table with a histogram. We plot the classes on the abscissa axis and, above each of them, we draw a rectangle whose area is proportional to the associated frequency f_i (or effective n_i). This graph is called the frequency histogram (see Figure 3.3).

Example 3.2.1. The graphical representation of the general example (see FIGURE 3.4 below),

Definition 3.2.2. For each class C_i , we set by definition

$$n_i = \text{Card}\{w \in \Omega : X(w) = C_i\}.$$

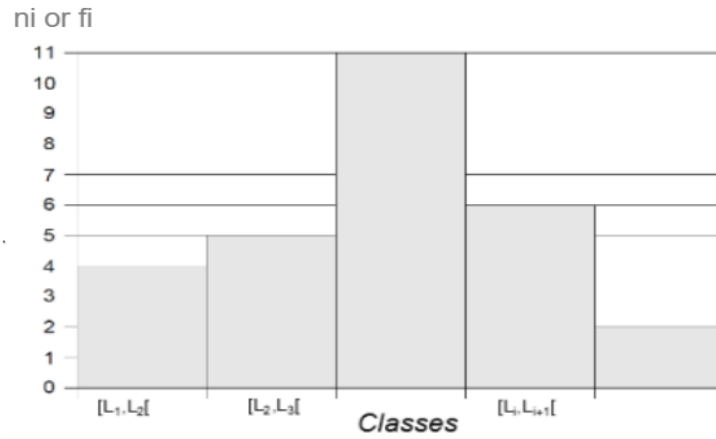


Figure 3.3: Histogram of frequencies (or repetitions).

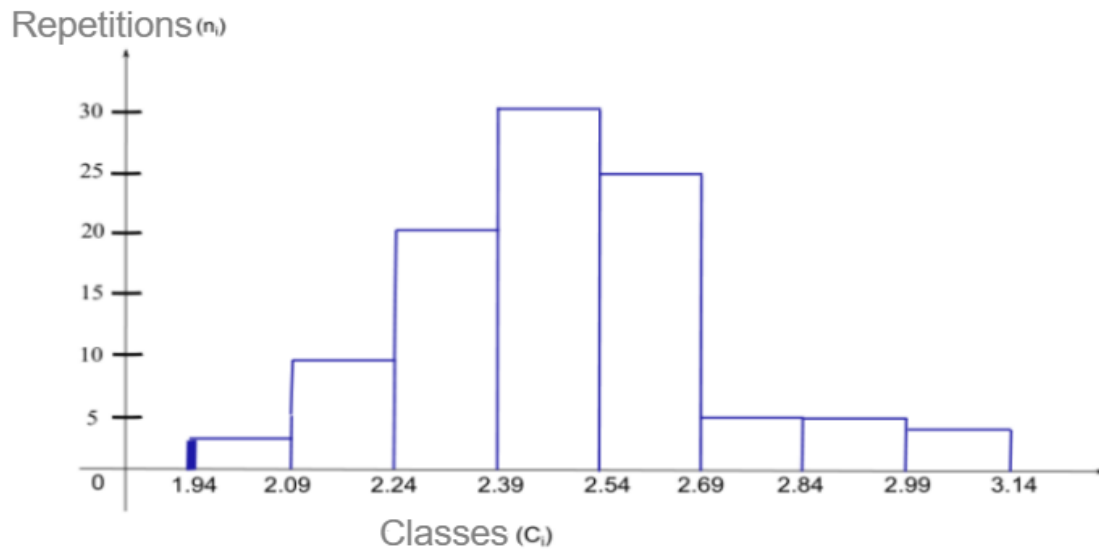


Figure 3.4: Histogram of repetitions

n_i : the number of individuals who have the same C_i , this is called repetition or (partial repetition) of C_i .

Definition 3.2.3. For each class C_i , we set by definition

$$N_i = n_1 + n_2 + \dots + n_i.$$

The cumulative number N_i of a value is the sum of the repetition of this class and all the repetitions of the previous values.

Interpretation:

N_i is the number of individuals whose character value is less than or equal to $x_i \in C_i$.

Therefore, the total number is given by

$$N = \text{Card}(\Omega) = \sum_{i=1}^k n_i.$$

In the previous example, we have $N = 100$.

Definition 3.2.4. For each class C_i , we set by definition

$$f_i = \frac{n_i}{N}.$$

f_i is called the partial frequency of C_i . The frequency of a value is the ratio (division) of the number of this value to the total number.

Remark 3.2.5. We can replace f_i by $f_i \times 100$ which then represents a percentage.

Interpretation:

f_i is the percentage of w such that $X(w) \in C_i$.

Definition 3.2.6. For each class C_i , we set by definition

$$F_i = f_1 + f_2 + \dots + f_i.$$

The quantity F_i is called the cumulative frequency of C_i .

Interpretation:

F_i is the percentage of w such that the value $X(w)$ is less than or equal to $x_i \in C_i$.

Example 3.2.7. *The following table represent classes, partial numbers (repetitions), cumulative number, partial frequencies and cumulative frequencies of the general example.*

C_i	[1.94, 2.09[	[2.09, 2.24[	[2.24, 2.39[	[2.39, 2.54[	[2.54, 2.69[
n_i	3	9	18	29	25
N_i	3	12	30	59	84
f_i	0.03	0.09	0.18	0.29	0.25
F_i	0.03	0.12	0.30	0.59	0.84
c_i	2.015	2.165	2.315	2.465	2.615

C_i	[2.69, 2.84[	[2.84, 2.99[	[2.99, 3.14[	Σ
n_i	6	6	4	100
N_i	90	96	100	-
f_i	0.06	0.06	0.04	1
F_i	0.9	0.96	1	-
c_i	2.765	2.915	3.065	-

3.3 Position Parameters (Central Tendency Characteristic)

The statistical indicators of central tendency (also known as position) frequently considered are the means (arithmetic, geometric, harmonic, quadratic), the quartiles (the first and the third), the median and the mode.

Definition 3.3.1. *The mode of a statistical variable is the value which has the greatest partial repetition (or the greatest partial frequency) and it is denoted by Mo .*

To determine the mode of a continuous character we have two cases:

1) *case of equal amplitude:*

$$Mo = \inf(C_{Mo}) + a_i \frac{n_i - n_{i-1}}{(n_i - n_{i-1}) + (n_i - n_{i+1})},$$

such that the modal class (C_{Mo}) is the class which correspond to the greatest repetition.

Also, $\inf(C_{Mo})$ is the lower bound of the modal class (C_{Mo}) and n_i its repetition.

2) *case of unequal amplitude:*

$$Mo = \inf(C_{Mo}) + a_i \frac{d_i - d_{i-1}}{(d_i - d_{i-1}) + (d_i - d_{i+1})},$$

such that the modal class (C_{Mo}) is the class which correspond to the greatest density.

Also, $\inf(C_{Mo})$ is the lower bound of the modal class (C_{Mo}) and d_i its density.

Example 3.3.2. In the general example, the mode is calculated using the given method:

First we must determine the modal class C_{Mo} :

$$29 \hookrightarrow [2.39, 2.54[\quad \text{so} \quad C_{Mo} = [2.39, 2.54[$$

The amplitude is constant and equal $a_i = 0.15$ then

$$Mo = \inf(C_{Mo}) + a_i \frac{n_i - n_{i-1}}{(n_i - n_{i-1}) + (n_i - n_{i+1})}.$$

$$\text{Hence } Mo = 2.39 + 0.15 \frac{29 - 18}{(29 - 18) + (29 - 25)} = 2.5$$

Remark 3.3.3. The mode need not be unique, so we can have more than one mode.

Definition 3.3.4. We call median the value Me of the statistical variable X which verifies

the following formula

$$Me = \inf(C_{Me}) + a_i \frac{\frac{N}{2} - N_{i-1}}{n_i},$$

such that the median class (C_{Me}) verify the following formula:

$$F_i(C_{Me}^-) < 0.5 \leq F_i(C_{Me}^+) = F_i(C_{Me})$$

Also, $\inf(C_{Me})$: the lower bound of the median class (C_{Me}), n_i its repetition, N the total number and N_{i-1} is the cumulative repetition of the class which is before the median class (C_{Me}).

The median divides the statistical series into two groups of the same size(repetition).

Example 3.3.5. According to the general example. The median is determined as:

First we need the middle class which satisfy:

$$F_i(C_{Me}^-) < 0.5 \leq F_i(C_{Me}^+) = F_i(C_{Me}) \implies 0.3 < 0.5 \leq 0.59$$

So $C_{Me} = [2.39, 2.54[$. Therefore the median is given by

$$Me = \inf(C_{Me}) + a_i \frac{\frac{N}{2} - N_{i-1}}{n_i}.$$

$$\text{Hence } Me = 2.39 + 0.15 \frac{50 - 30}{29} = 2.49$$

Definition 3.3.6. We call the average (the arithmetic mean) of x , the amount

$$\bar{x} = \frac{1}{N} \sum_{i=1}^k n_i c_i = \sum_{i=1}^k f_i c_i$$

with $N = \text{Card}(\Omega)$. We can therefore express and calculate the arithmetic mean with numbers or with frequencies.

Remark 3.3.7. *The value of the average is abstract. It means it is a number that does not correspond to a concrete fact.*

Definition 3.3.8. *We call the geometric mean of x , the amount*

$$\bar{x}_G = 10^{\left(\frac{1}{N} \sum_{i=1}^k n_i \log(c_i) \right)}$$

Definition 3.3.9. *We call the harmonic mean of x , the amount*

$$\bar{x}_H = \frac{N}{\sum_{i=1}^k \frac{n_i}{c_i}}$$

Definition 3.3.10. *We call the quadratic mean of x , the amount*

$$\bar{x}_Q = \sqrt{\frac{\sum_{i=1}^k n_i c_i^2}{N}}$$

Example 3.3.11. *The averages in the general example are founded as:*

- *The arithmetic mean is given by*

$$\bar{x} = \frac{1}{N} \sum_{i=1}^k n_i c_i = \frac{250.4}{100} = 2.5$$

- *The geometric mean is given by*

$$\bar{x}_G = 10^{\left[\frac{1}{N} \sum_{i=1}^k n_i \log(c_i) \right]} = 10^{\left[\frac{39.68}{100} \right]} = 10^{0.3968} = 2.5$$

- The harmonic mean is given by

$$\bar{x}_H = \left[\frac{N}{\sum_{i=1}^k n_i/c_i} \right] = \frac{100}{40.26} = 2.48$$

- The arithmetic mean is given by

$$\bar{x}_Q = \sqrt{\frac{1}{N} \sum_{i=1}^k n_i c_i^2} = \sqrt{\frac{632.44}{100}} = 2.51$$

In statistics, Quartiles are the set of values which has three points dividing the data set into four identical parts. We ordinarily deal with a large amount of numerical data, in stats. There are several concepts and formulas, which are extensively applicable in various researches and surveys.

Definition 3.3.12. (First Quartile)

The first quartile denoted (Q_1) is the median of the lower half of the data set, and it is determined by the following formula

$$Q_1 = \inf(C_{Q_1}) + a_i \frac{\frac{N}{4} - N_{i-1}}{n_i},$$

such that the first quartile class (C_{Q_1}) verify the following formula:

$$F_i(C_{Q_1}^-) < 0.25 \leq F_i(C_{Q_1}^+) = F_i(C_{Q_1}),$$

also, $\inf(C_{Q_1})$ is the lower bound of the first quartile class and N_{i-1} is the cumulative repetition of the class which is before the first quartile class C_{Q_1} .

Remark 3.3.13. The median is considered as the second quartile (Q_2).

Definition 3.3.14. (Third Quartile)

The third quartile denoted (Q_3) is the median of the upper half of the data. and it is determined by the given formula

$$Q_3 = \inf(C_{Q_3}) + a_i \frac{\frac{3N}{4} - N_{i-1}}{n_i},$$

such that the third quartile class (C_{Q_3}) verify the following formula:

$$F_i(C_{Q_3}^-) < 0.75 \leq F_i(C_{Q_3}^+) = F_i(C_{Q_3}),$$

also, $\inf(C_{Q_3})$ is the lower bound of the third quartile class and N_{i-1} is the cumulative repetition of the class which is before the first quartile class C_{Q_1} .

Definition 3.3.15. Interquartile Range (IQR) is the difference between the third quartile and first quartile is known as the interquartile range. (i.e.) $IQR = Q_3 - Q_1$

Also, the interquartile interval is defined as $[Q_1, Q_3]$.

Example 3.3.16. In the general example,

- The first quartile is calculated as:

First we should determine the class of the first quartile C_{Q_1} :

$$F_i(C_{Q_1}^-) < 0.25 \leq F_i(C_{Q_1}^+) = F_i(C_{Q_1}) \Rightarrow 0.12 < 0.25 \leq 0.30$$

So the first quartile class is the class $C_{Q_1} = [2.24, 2.39[$,

Consequently

$$Q_1 = \inf(C_{Q_1}) + a_i \frac{\frac{N}{4} - N_{i-1}}{n_i}$$

$$Q_1 = 2.24 + 0.15 \frac{25 - 12}{18} = 2.35$$

- The third quartile is determined as:

It is first necessary to determine the class of the third quartile C_{Q_3} :

$$F_i(C_{Q_3}^-) < 0.75 \leq F_i(C_{Q_3}^+) = F_i(C_{Q_3}) \Rightarrow 0.59 < 0.75 \leq 0.84$$

So the third quartile class is the class $C_{Q_3} = [2.54, 2.69[$,

Hence,

$$Q_3 = \inf(C_{Q_3}) + a_i \frac{\frac{3N}{4} - N_{i-1}}{n_i}$$

$$Q_3 = 2.54 + 0.15 \frac{75 - 59}{25} = 2.64$$

- The interquartile range is given as

$$IQR = Q_3 - Q_1 = 2.64 - 2.35 = 0.29$$

- The interquartile interval is founded as

$$[Q_1, Q_3] = [2.35, 2.64]$$

3.4 Dispersion Parameters (Variability Characteristic)

The common statistical indicators of dispersion are the range, the variance, the standard deviation, the coefficient of variation, the average absolute deviation.

Definition 3.4.1. The difference between the largest value and the smallest value of the

character, given by the quantity

$$e = x_{\max} - x_{\min}.$$

is called the extent of the statistical variable X . Calculating the range is very simple. It gives a first idea of the dispersion of the observations. This is a very rudimentary indicator and there are more sophisticated dispersion indicators (see below).

Example 3.4.2. In the previous example the range is given as:

$$e = x_{\max} - x_{\min} = 3.12 - 2.94 = 1.18$$

Definition 3.4.3. We call the variance of this statistical series X , the number

$$\text{Var}(X) = \frac{1}{N} \sum_{i=1}^k n_i (c_i - \bar{x})^2$$

We say that the variance is the average of the squares of the deviations from the mean $\bar{x}$

The following theorem (König-Huygens Theorem) gives a remarkable identity linking the variance and the mean, sometimes more practical in calculating the variance.

Theorem 3.4.4. Let (C_i, n_i) be a statistical series with mean $\bar{x}$ and variance $\text{Var}(X)$.

So

$$\text{Var}(X) = \left(\frac{1}{N} \sum_{i=1}^k n_i c_i^2 \right) - \bar{x}^2$$

Definition 3.4.5. The standard deviation of the statistical variable X is the amount

$$\sigma_X = \sqrt{\text{Var}(X)}.$$

Definition 3.4.6. *The coefficient of variation of the statistical variable X is giving as*

$$Cv = \frac{\sigma_X}{\bar{x}}.$$

Example 3.4.7. *In the general example of this chapter, the variance, the standard deviation, the coefficient of variation are founded as follow:*

- *The variance of the variable*

$$Var(X) = \sum_{i=1}^k f_i c_i^2 - \bar{x}^2 = 6.32 - 6.27 = 0.05$$

- *The standard deviation of the variable*

$$\sigma_X = \sqrt{Var(X)} = \sqrt{0.05} = 0.22$$

- *The coefficient of variation*

$$Cv = \frac{\sigma_X}{\bar{x}} = \frac{0.22}{2.5} = 0.088$$

Definition 3.4.8. (Average absolute deviation)

- *The average absolute deviation from the arithmetic mean is the amount*

$$A.A.D_{\bar{x}} = \frac{1}{N} \sum_{i=1}^k n_i |c_i - \bar{x}|$$

- *The average absolute deviation from the mode is giving by*

$$A.A.D_{Mo} = \frac{1}{N} \sum_{i=1}^k n_i |c_i - Mo|$$

- The average absolute deviation from the median is defined as

$$A.A.D_{Me} = \frac{1}{N} \sum_{i=1}^k n_i |c_i - Me|$$

Example 3.4.9. According to the data of the general example in this chapter the averages absolute deviations are calculated as:

- The average absolute deviation from the arithmetic mean

$$A.A.D_{\bar{x}} = \frac{1}{N} \sum_{i=1}^k n_i |c_i - \bar{x}| = \frac{18.03}{100} = 0.18$$

- The average absolute deviation from the mode

$$\text{We have } \bar{x} = Mo \text{ Then } A.A.D_{\bar{x}} = A.A.D_{Mo} = 0.18$$

- The average absolute deviation from the median

$$\text{We have } \bar{x} = Mo \simeq Me \text{ So } A.A.D_{\bar{x}} = A.A.D_{Mo} \simeq A.A.D_{Me} = 0.18$$

CHAPTER 4

PROBABILITY CALCULUS

Uncertainty is an inherent aspect of many phenomena in science, engineering, and everyday life. Whether we analyze the outcome of a game of chance, model the behavior of complex systems, or make inferences based on incomplete information, we need a mathematical framework capable of describing and quantifying randomness. Probability calculus provides this framework.

This chapter presents the fundamental concepts and tools that underlie classical probability theory. We begin with combinatorial analysis, a collection of counting techniques that allow us to determine the number of possible outcomes in complex situations. These methods are essential for solving many probability problems, especially those involving equally likely outcomes.

We then introduce the basic notions of probability, starting from random experiments, outcomes, and sample spaces, followed by a rigorous treatment of events and operations on events. Understanding how events combine—through unions, intersections, and complements—is crucial for developing more advanced results.

Next, we formalize the idea of assigning probabilities by defining probability measures on probabilizable spaces, according to the axioms of Kolmogorov. This framework provides a solid mathematical foundation on which the entire theory rests.

The chapter then focuses on conditional probability, compound probability, and independence, which are essential concepts for modeling dependent events and sequential processes. We conclude with Bayes' theorem, a powerful tool for updating probabilities when new information becomes available, forming the basis of many modern applications in statistics, decision theory, and machine learning.

To reinforce the theoretical concepts, the chapter includes a series of corrected exercises, followed by additional problems for further practice and mastery.

4.1 Combinatorial Analysis

4.1.1 Fundamental Principle of Combinatorial Analysis

Combinatorial analysis provides the fundamental tools for counting structured selections. It is essential for probability theory.

Definition 4.1.1. (*Rule of product*)

If an operation (experiment, process) is performed in k successive steps, and:

- Step 1 can be carried out in n_1 ways.*
- Step 2 can be carried out in n_2 ways.*
- Step 3 in n_3 ways.*
- ...*
- Step k in n_k ways.*

Then the total number of possible outcomes is:

$$N = n_1 \times n_2 \times \dots \times n_k$$

Interpretation :

This principle applies when every step is independent of the others.

We multiply the number of choices at each step to get the total number of possibilities.

Example 4.1.2. (Password of 3 digits)

A password consists of three digits, where each digit can take any value from 0 to 9.

What is the total number of possible passwords that can be formed ?

Solution:

Each digit of the password can be chosen in 10 different ways (0, 1, 2, ..., 9).

Since the three digits are chosen independently, the total number of possible passwords is:

$$N = 10 \times 10 \times 10 = 1000.$$

Example 4.1.3. (Clothing choice)

A person has the following clothing options:

- 4 shirts.
- 3 pants.
- 2 pairs of shoes.

How many different outfits can the person create by choosing one shirt, one pair of pants, and one pair of shoes each time ?

Solution:

Each outfit is formed by combining:

one of the 4 shirts, one of the 3 pants, and one of the 2 pairs of shoes.

Since each choice is independent, the total number of possible outfits is:

$$N = 4 \times 3 \times 2 = 24.$$

Example 4.1.4. (Licence plates)

A licence plate consists of:

- 3 letters, each of which can be any of the 26 letters in the alphabet, and
- 3 digits, each of which can be any number from 0 to 9.

How many different licence plates can be formed ?

Solution:

For the letters: $26 \times 26 \times 26 = 26^3$.

For the digits: $10 \times 10 \times 10 = 10^3$.

Since the choices of letters and digits are independent, the total number of possible licence plates is:

$$N = 26^3 \times 10^3 = 17,576,000.$$

4.1.2 Arrangements

To visualize arrangements, consider the selection of 2 ordered elements from the set $\{a, b, c\}$. The diagram below shows all possible ordered pairs:

This figure visually represents that each element may appear in the first or second position, generating $3^2 = 9$ possible arrangements.

Definition 4.1.5. (Arrangements without repetition)

We call arrangement without repetition (denoted by A_n^p) of n elements ($n \geq p$) any ordered set of p of these elements, all distinct. An arrangement is therefore characterized by the

ARRANGEMENTS

An arrangement of a set is an ordered selection of elements by picking them one at a time.

1st	2nd	3rd
a	a	b
b	c	c
a	b	a

Figure 1 Arrangements of the set $\{a, b, c\}$

The number of possible arrangements of n elements chosen from a set of n elements is denoted by A^n and is equal to $n!$

Figure 4.1: Illustration of the 9 arrangements of order 2 from the set $\{a, b, c\}$.

nature of the elements or by their order.

$$A_n^p = \frac{n!}{(n-p)!}, \quad n, p \in \mathbb{N}^* \text{ and } n \geq p.$$

Example 4.1.6. *Third in order in a race of 14 horses.*

$$A_{14}^3 = \frac{14!}{(14-3)!} = \frac{14!}{(11)!} = 2184.$$

Definition 4.1.7. (*Arrangements with repetition (list)*)

An arrangement of n objects taken p at a time with repetition is an arrangement where each object can be repeated up to p times. Hence

$$A_n^p = n^p.$$

Example 4.1.8. *-The total number of arrangements with repetition of order 2 of the letters a, b, c is*

$$n^p = 3^2 = 9, \quad \text{i.e., } \{aa, ab, ac, ba, bb, bc, ca, cb, cc\}.$$

-The total number of arrangements with repetition and order 3 of the letters a, b is

$$n^p = 2^3 = 8, \quad \text{i.e.,} \quad \{aaa, aab, abb, aba, bbb, bab, baa, bba\}.$$

4.1.3 Permutations

Definition 4.1.9. (*Permutation without repetition*)

A permutation denoted by P of n objects is an ordered set of these n objects. The permutations of these n objects constitute a particular case of arrangements, which is the case where $n = p$.

$$P_n = A_n^n = \frac{n!}{(n-n)!} = \frac{n!}{0!} = n!$$

Example 4.1.10. The possible permutations of the three letters x, y, z are:

$$P! = 3! = 6, \quad \text{i.e.,} \quad \{xyz, xzy, yxz, yzx, zxy, zyx\}.$$

Definition 4.1.11. (*Permutation with repetition*)

Let $F = \{a_1, a_2, \dots, a_k\}$ be a finite set of cardinality k . Let $r_1, r_2, \dots, r_k$ be natural numbers and n their sum, i.e., $n = r_1 + r_2 + \dots + r_k$. A permutation of n elements of F with $r_1, r_2, \dots, r_k$ repetitions is an n -tuple of elements of F in which each of the elements $a_1, a_2, \dots, a_k$ of F appears respectively $r_1, r_2, \dots, r_k$ times. We notes by $P_n^{r_1, r_2, \dots, r_k}$ the number of permutations of n elements with $r_1, r_2, \dots, r_k$ repetitions.

$$P_n^{r_1, r_2, \dots, r_k} = \frac{n!}{r_1! r_2! \dots r_k!}$$

Example 4.1.12. The number of possible permutations with the letters of the word "in-

diene” is:

$$\text{we have } n = 8, r_i = 2, r_e = 2, r_n = 3 \text{ then, } P_{n=8}^{r_i, r_e, r_n} = \frac{8!}{2!2!3!} = 1680.$$

4.1.4 Combinations

Definition 4.1.13. (*Combinations (binomial coefficients)*)

We call combinations (denoted C_n^p) of p elements among n ($n \geq p$) any set that can be formed by choosing p of these elements, without consideration of order.

$$C_n^p = \frac{A_n^p}{p!} = \frac{n!}{p!(n-p)!}, \quad n, p \in \mathbb{N}^* \text{ and } n \geq p.$$

Example 4.1.14. The possible combinations of the following 4 letters a, b, c, d 3 to 3 are:

$$C_4^3 = \frac{4!}{3!(4-3)!} = 4 \quad \text{i.e., } \{abc, abd, acd, bcd\}$$

.

4.2 Basic Concepts

4.2.1 Random experiment, Outcome and Sample space

Definition 4.2.1. (*Random Experiment*)

An experiment (or trial) is called random if its outcome cannot be predicted with certainty, and if, when repeated under identical conditions, it may produce different results.

Example 4.2.2. Tossing a coin is a random experiment, because we cannot predict whether it will land on heads or tails.

Definition 4.2.3. (*Elementary Event*)

The outcome of a random experiment is called an elementary event and is usually denoted by ω .

Example 4.2.4. When tossing a coin, the possible elementary events are $\omega_1 = \text{Heads}$ and $\omega_2 = \text{Tails}$.

Definition 4.2.5. (Sample Space)

The set of all possible elementary events of a random experiment is called the sample space, denoted by Ω .

Example 4.2.6. For the coin toss experiment, the sample space is:

$$\Omega = \{\text{Heads}, \text{Tails}\}.$$

4.2.2 Events

Definition 4.2.7. (Event)

- An event is any subset of the sample space Ω .
- It is a specific collection of outcomes that may occur when the experiment is performed.

Remark 4.2.8. The universe Ω can be finite or infinite, countable if we can enumerate the known possibilities, or continuous.

Example 4.2.9.

- **Finite:** The number of outcomes is limited.

Rolling a six-sided die, $\Omega = \{1, 2, 3, 4, 5, 6\}$.

- **Countably infinite:** The outcomes can be listed in a sequence (can be enumerated).

The number of coin tosses needed to get the first head, $\Omega = \{1, 2, 3, \dots\}$.

- **Continuous:** The outcomes form a continuum, often represented by an interval of real numbers.

Measuring the time (in seconds) for a light bulb to burn out, $\Omega = [0, \infty)$.

Definition 4.2.10. (Certain Event)

A certain event is an event that is guaranteed to occur, represented by the sample space Ω .

Example 4.2.11.

D : “rolling a number divisible by 1” = $\{1, 2, 3, 4, 5, 6\}$

$$P(D) = \frac{\text{card}(D)}{\text{card}(\Omega)} = \frac{6}{6} = 1$$

Definition 4.2.12. (Impossible Event)

The impossible event, represented by $\emptyset$, contains no outcomes and cannot occur. Its probability is zero:

$$P(\emptyset) = 0$$

Example 4.2.13.

E : “rolling a number divisible by 7” = $\emptyset$

$$P(E) = \frac{\text{card}(E)}{\text{card}(\Omega)} = 0$$

Definition 4.2.14. Any event A is a set of elementary events and constitutes a subset of the sample space Ω , where we can determine whether it has occurred or not at the outcome of the experiment.

Example 4.2.15. We have a sample space Ω representing the outcomes of rolling a six-sided die. The event B “rolling an even number” can be expressed as:

$$B = \{2, 4, 6\} \quad \text{where} \quad B \subset \Omega$$

Definition 4.2.16. If Ω is finite, each subset A of the universe Ω ($A \subset \Omega$) consists of a finite number of possible outcomes. In this case, the set of events $\varepsilon(\Omega)$ is such that:

$$\varepsilon(\Omega) = P(\Omega)$$

where $P(\Omega)$ is the power set of the sample space, representing the set of all possible events in the universe.

Example 4.2.17. Suppose we roll a standard six-sided die, and the sample space is $\Omega = \{1, 2, 3, 4, 5, 6\}$. In this case:

- $\varepsilon(\Omega)$ is the set of all possible events.
- $\mathcal{P}(\Omega)$ is the power set of Ω , which includes all possible subsets.

The power set $\mathcal{P}(\Omega)$ is:

$$\mathcal{P}(\Omega) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \{1, 2\}, \{1, 3\}, \dots, \{1, 2, 3\}, \dots, \{1, 2, 3, 4, 5, 6\}\}$$

4.2.3 Operations on Events

To simplify the understanding of operations on events, we introduce the following figure. It uses distinct yet soft colours (light blue, light gray, and yellow) to ensure clarity and to highlight the overlapping regions effectively.

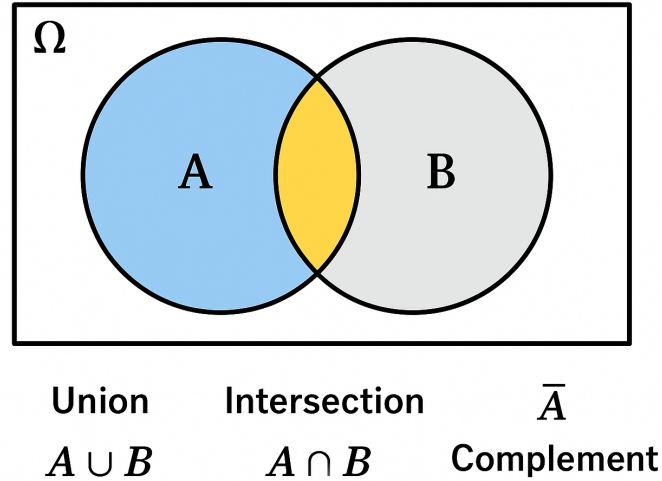


Figure 4.2: Operations on events.

Discussion of the Figure (with colors):

The Venn diagram displays the sample space Ω (the rectangular frame) and two events A and B (the two circles). Colours are used to make each operation visually distinct:

- **Event A :** shaded light blue: this colour marks all outcomes that belong to A .
- **Event B :** shaded light gray: this colour marks all outcomes that belong to B .
- **Intersection $A \cap B$:** shaded yellow (the overlap): the yellow region is the area common to both A and B ; it represents outcomes for which both events occur simultaneously.
- **Union $A \cup B$:** visually the union of the light blue and light gray regions (including the yellow overlap): together these shaded areas show outcomes where at least one of the events occurs. If you want a single colour for the union, one can use a semi-transparent colour overlay so the union appears as a combined coloured region.
- **Complement A^c :** the unshaded part of the rectangle outside the light-blue circle: this represents all outcomes in Ω where A does not occur. (In the figure the background remains white to indicate the complement.)
- **Difference $A \setminus B$:** the portion of A that is light blue but not yellow (i.e., the light-blue area excluding the yellow overlap): these are outcomes where A occurs and B does not.

- **Symmetric difference $A \Delta B$:** the non-overlapping parts of A and B (light blue plus light gray, excluding the yellow intersection): outcomes where exactly one of A or B occurs.

Definition 4.2.18. (Complementary Event)

The complementary event of an event X , denoted $\bar{X}$, is the event that occurs if and only if X does not occur. It consists of all outcomes in Ω that are not in X .

$$P(\bar{X}) = 1 - P(X)$$

Example 4.2.19. From the previous example:

$$P(\bar{A}) = 1 - P(A) = 1 - \frac{1}{2} = \frac{1}{2}, \quad P(\bar{B}) = 1 - P(B) = 1 - \frac{1}{3} = \frac{2}{3}.$$

Definition 4.2.20. (Elementary Event)

An elementary event is an event consisting of a single outcome.

$$P(X) = \frac{1}{\text{card}(\Omega)}$$

Example 4.2.21. Roll a six-sided die once. Consider the event:

$$C : \text{“rolling a number divisible by 5”} = \{5\}.$$

$$P(C) = \frac{\text{card}(C)}{\text{card}(\Omega)} = \frac{1}{6}.$$

Definition 4.2.22. (Intersection of Events)

The intersection of two events A and B , denoted $A \cap B$, is the event that occurs if and only if both A and B occur. It consists of outcomes common to both.

Example 4.2.23.

$$A = \text{"even number"} = \{2, 4, 6\}, \quad B = \text{"divisible by 3"} = \{3, 6\}$$

$$A \cap B = \{6\}, \quad P(A \cap B) = \frac{1}{6}$$

Definition 4.2.24. (Union of Events)

The union of two events A and B , denoted $A \cup B$, occurs if at least one of A or B occurs:

$$A \cup B = \{2, 3, 4, 6\}, \quad P(A \cup B) = \frac{4}{6} = \frac{2}{3}.$$

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

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

Definition 4.2.25. (Inclusion of Events)

An event A is said to be included in event B , denoted $A \subset B$, if whenever A occurs, B also occurs. Then:

$$P(A) \leq P(B)$$

Definition 4.2.26. (Disjoint / Incompatible Events)

Two events A and B are incompatible (or disjoint) if they cannot occur simultaneously:

$$A \cap B = \emptyset \quad \Rightarrow \quad P(A \cap B) = 0$$

Example 4.2.27.

$$A = \{2, 4, 6\}, \quad B = \{1, 3, 5\}, \quad C = \{3, 6\}$$

$$A \cap B = \emptyset \Rightarrow A \text{ and } B \text{ are incompatible.}$$

$A \cap C = \{6\} \neq \emptyset \Rightarrow A$ and C are not incompatible.

Definition 4.2.28. (*Independent Events*)

Two events A and B are independent if:

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

Example 4.2.29.

$$A = \{2, 4, 6\}, \quad B = \{3, 6\}, \quad \Omega = \{1, 2, 3, 4, 5, 6\}$$

$$P(A) = \frac{3}{6} = \frac{1}{2}, \quad P(B) = \frac{2}{6} = \frac{1}{3}$$

$$P(A \cap B) = \frac{1}{6}, \quad P(A) \times P(B) = \frac{1}{6}$$

Hence, A and B are independent.

4.2.4 Complete System of Events

Definition 4.2.30. (*Complete System of Events*)

Events $A_1, A_2, \dots, A_n$ form a complete system if:

$$A_i \neq \emptyset, \quad A_i \cap A_j = \emptyset \quad \forall i \neq j, \quad \bigcup_{i=1}^n A_i = \Omega$$

Example 4.2.31.

$$\Omega = \{1, 2, 3, 4, 5, 6\}, \quad A = \{2, 4, 6\}, \quad C = \{1, 3, 5\}$$

$$A \cap C = \emptyset, \quad A \cup C = \Omega$$

A and C form a complete system.

4.3 Probability Measure

4.3.1 Probabilizable Space and Probability Function

Definition 4.3.1. (*Probabilizable Space*)

A probabilizable space (Ω, C) consists of:

- a sample space Ω , representing all possible outcomes of a random experiment;
- a collection of events $C \subset \mathcal{P}(\Omega)$, where $\mathcal{P}(\Omega)$ is the power set of Ω , representing all possible subsets of outcomes that can be considered events.

The collection C satisfies the following properties:

(P1) If $A \in C$, then its complement $\bar{A} \in C$.

(P2) If $A_i \in C$ for all $i \in \mathbb{N}^*$, then $\bigcup_i A_i \in C$.

(P3) $\Omega \in C$.

Example 4.3.2. Consider rolling a six-sided die:

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

Let C include:

- All single-element outcomes, e.g., $\{1\}, \{2\}, \dots$
- The event of rolling an even number, $E = \{2, 4, 6\}$.
- The complement of E , i.e., rolling an odd number, $\bar{E} = \{1, 3, 5\}$.

Then (Ω, C) is a probabilizable space satisfying (P1), (P2), and (P3).

Definition 4.3.3. (*Probability Function*)

Let (Ω, C) be a probabilizable space. A probability is a function $P : C \rightarrow \mathbb{R}^+$ satisfying the **Kolmogorov axioms**:

1. **Non-negativity:** $P(A) \geq 0$ for all $A \in C$.
2. **Normalization:** $P(\Omega) = 1$.
3. **Additivity:** For any finite or countable sequence of pairwise disjoint events $A_1, A_2, \dots$,

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

4.3.2 Probability Space and Probability of an Event

Definition 4.3.4. (*Probability Space*)

A probability space is a triplet (Ω, C, P) where:

1. (Ω, C) is a probabilizable space;
2. P is a probability function on C .

Definition 4.3.5. (*Classical Probability of an Event*)

If all outcomes in a finite sample space are equally likely, the probability of an event X is given by:

$$P(X) = \frac{\text{card}(X)}{\text{card}(\Omega)} = \frac{\text{number of favorable outcomes}}{\text{total number of possible outcomes}}.$$

Example 4.3.6. Roll a six-sided die once. The sample space is

$$\Omega = \{1, 2, 3, 4, 5, 6\}, \quad \text{card}(\Omega) = 6.$$

Consider the events:

$$A = \text{“rolling an even number”} = \{2, 4, 6\}, \quad \text{card}(A) = 3$$

$$B = \text{“rolling a number divisible by 3”} = \{3, 6\}, \quad \text{card}(B) = 2$$

Then the probabilities are:

$$P(A) = \frac{\text{card}(A)}{\text{card}(\Omega)} = \frac{3}{6} = 0.5$$

$$P(B) = \frac{\text{card}(B)}{\text{card}(\Omega)} = \frac{2}{6} = \frac{1}{3} \approx 0.33$$

4.4 Conditional Probability

Definition 4.4.1. (*Conditional Probability*)

Let A and B be two events in a probability space (Ω, P) with $P(B) \neq 0$. The conditional probability of A given B is defined as:

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

and is sometimes denoted $P_B(A)$.

Theorem 4.4.2. Let B be an event with $P(B) \neq 0$. Then

$$P_B : \mathcal{E}(\Omega) \rightarrow [0, 1], \quad A \mapsto P_B(A) = \frac{P(A \cap B)}{P(B)}$$

defines a probability measure on Ω .

Example 4.4.3. Consider a standard deck of 52 cards. Define events:

A : Drawing a red card, B : Drawing a heart.

Compute $P(A | B)$:

1. $A \cap B = \text{red hearts} = 13 \text{ cards} \Rightarrow P(A \cap B) = \frac{13}{52} = \frac{1}{4}.$

2. $B = \text{hearts} = 13 \text{ cards} \Rightarrow P(B) = \frac{13}{52} = \frac{1}{4}.$

3. Apply the formula:

$$P(A | B) = \frac{P(A \cap B)}{P(B)} = \frac{1/4}{1/4} = 1.$$

4.4.1 Compound probability

Theorem 4.4.4. (Compound Probability Formula)

For any two events A and B :

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

Example 4.4.5. Using the same deck of cards:

A : red card, B : heart.

$$P(A) = \frac{26}{52} = \frac{1}{2}, \quad P(B | A) = \frac{13}{26} = \frac{1}{2}.$$

$$P(A \cap B) = P(A)P(B | A) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}.$$

4.4.2 Independence and Conditional Probability

Properties 4.4.6. (Independence and Conditional Probability)

1. If A and B are independent with $P(B) \neq 0$, then:

$$P(A | B) = P(A).$$

2. Conversely, if $P(A | B) = P(A)$, events A and B are independent.

Example 4.4.7. Consider a fair six-sided die. Let:

$$A : \text{even number } (2, 4, 6), \quad B : \text{number} \leq 4 \ (1, 2, 3, 4).$$

Compute probabilities:

$$P(A) = \frac{3}{6} = \frac{1}{2}, \quad P(B) = \frac{4}{6} = \frac{2}{3}.$$

$$A \cap B = \{2, 4\} \Rightarrow P(A \cap B) = \frac{2}{6} = \frac{1}{3}.$$

Check independence:

$$P(A) \cdot P(B) = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3} = P(A \cap B),$$

so A and B are independent.

Verify conditional probability:

$$P(A | B) = \frac{P(A \cap B)}{P(B)} = \frac{1/3}{2/3} = \frac{1}{2} = P(A).$$

4.4.3 Bayes' Theorem

Theorem 4.4.8. (Bayes' Theorem)

Let $\{A_1, A_2, \dots, A_n\}$ be a complete system of events with $P(A_i) \neq 0$, and let B be an

event with $P(B) \neq 0$. Then, for any i :

$$P(A_i | B) = \frac{P(B | A_i)P(A_i)}{\sum_{j=1}^n P(B | A_j)P(A_j)}.$$

Proof : By definition of conditional probability:

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

From the compound probability formula:

$$P(A_i \cap B) = P(A_i)P(B | A_i).$$

Also, using the law of total probability:

$$P(B) = \sum_{j=1}^n P(B | A_j)P(A_j).$$

Substitute into the conditional probability formula:

$$P(A_i | B) = \frac{P(A_i)P(B | A_i)}{\sum_{j=1}^n P(A_j)P(B | A_j)}. \quad \square$$

□

Example 4.4.9. Two factories produce a product. Factory A produces 60% of the total, Factory B 40%. Defective rates: $P(B | A) = 0.05$, $P(B | \bar{A}) = 0.03$. Let:

A = product from Factory A, B = product is defective.

Use Bayes' theorem to compute $P(A | B)$:

$$P(A | B) = \frac{P(B | A)P(A)}{P(B | A)P(A) + P(B | \bar{A})P(\bar{A})}.$$

Substitute values:

$$P(A | B) = \frac{0.05 \cdot 0.6}{0.05 \cdot 0.6 + 0.03 \cdot 0.4} = \frac{0.03}{0.03 + 0.012} = \frac{0.03}{0.042} \approx 0.714.$$

Thus, the probability that a defective product comes from Factory A is approximately 71.4%.

REFERENCES

- [1] *G. Calot*, Cours de statistique descriptive, *Dunod, Paris, 1973*.
- [2] *P. Bailly*, Exercices corrigés de statistique descriptive, *OPU Alger, 1993*.
- [3] *H. Hamdani*, Statistique descriptive avec initiation aux méthodes d'analyse de l'information économique: exercices et corrigés, *OPU Alger, 2006*.
- [4] *K. Redjdal*, Probabilités, *OPU Alger, 2004*.
- [5] Descriptive Statistics and Probability Theory, *Robert A. Barks Hutchinson, 1972*.
- [6] Statistics: Descriptive statistics and probability, *Elliot A. Tanis Harcourt Brace Jovanovich, 1987*.
- [7] Probability and Statistics: *The Science of Uncertainty* Michael J. Evans and Jeffrey S. Rosenthal University of Toronto.
- [8] *G. Casella and R. L. Berger*, Statistical Inference, *2nd Edition, Duxbury, 2002*.
- [9] *S. M. Ross*, Introduction to Probability and Statistics for Engineers and Scientists, *5th Edition, Academic Press, 2014*.

- [10] *M. H. DeGroot and M. J. Schervish, Probability and Statistics, 4th Edition, Pearson, 2012.*
- [11] *D. S. Moore, G. P. McCabe, and B. Craig, Introduction to the Practice of Statistics, 8th Edition, W. H. Freeman, 2013.*
- [12] *W. Feller, An Introduction to Probability Theory and Its Applications, Vol. 1, 3rd Edition, Wiley, 1968.*