الجـمـهـوريـة الـجـزائريـة الديمـقـراطية الشـعـبيـة
PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA

MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH
MOHAMED BOUDIAF UNIVERSITY OF M'SILA
FACULTY OF TECHNOLOGY
DEPARTMENT OF ELECTRICAL ENGINEERING

وزارة التـعـلـيم الـعـالي و البحث العلمي
جامــعــة محمد بوضياف بالمسيلــة
كــلـيــة التكنولوجيا
قسم الهندسة الكهربائية

رقم:.....103/ ق.هـ.ك /2025 — المسيلة في : 29 أفريل 2025

شهادة إدارية

بخصوص مطبوعة الدروس الخاصة بالأستاذ

بوقرة خير الدين

بناءً على محضر اللجنة العلمية لقسم الهندسة الكهربائية تحت رقم: 41/ق.هـ.ك/2025 المنعقد بتاريخ 17 فيفري 2025 والمتضمن تعين الخبراء: الأستاذة لوكال كلثوم أستاذة محاضرة-أ- بجامعة المسيلة الأستاذ حريزي عبد الغفور أستاذ محاضر –أ- بجامعة المسيلة، والأستاذ لعماري يحي أستاذ محاضر –أ- بجامعة باتنة 02 وذلك لتقييم مطبوعة الدروس الخاصة بالأستاذ بوقرة خير الدين أستاذ محاضر "أ" بقسم الهندسة المدنية لجامعة المسيلة المدرسة بقسم الهندسة الكهربائية تحت عنوان :

" **Probabilités et statistiques** "

مطبوعة دروس مكتوبة باللغة الإنجليزية تحت عنوان: "**Lecture notes on probability and statistics** "

وبعد إطلاع رئيس اللجنة العلمية ورئيس القسم على التقارير الواردة و التي كانت كلها ايجابية، وعليه فإن اللجنة لا ترى مانعا أن تتخذه سندا في تدريس طلبة السنة الثانية ليسانس في الكهروتقني ، شعبة الكهرتقني ، ميدان علوم و تكنولوجيا و أن تعتمد في أي تقييم للمسار العلمي للأستاذ المعني.

رئيس اللجنة العلمية
أ .د بوقرة عبد الرحمان

رئيس القسم
دفداف المبروك

ملاحظة: سلمت هذه الشهادة للمعني(ة) لاستعمالها في حدود ما يسمح به القانون.

☐ Department : Electrical Engineering

☐ Course : Probability and Statistics

LECTURE NOTES

ON

PROBABILITY AND STATISTICS

BY

KHEIREDDINE BOUGUERRA

جامعة محمد بوضياف - المسيلة
Université Mohamed Boudiaf - M'sila

Overview

This document serves as a valuable resource for engineering students seeking to gain a deep understanding of fundamental concepts in probability and statistics. It covers statistical series for both one and two variables, enabling rigorous data analysis and interpretation.

The first part focuses on one-variable statistical series, emphasizing central tendency indicators (mean, median, mode) and dispersion measures (variance, standard deviation). The second part addresses two-variable statistical series, introducing key concepts such as covariance and correlation, which are essential for establishing relationships between different quantities.

Furthermore, the document explores probability theory in a finite framework, providing students with the necessary tools to quantify uncertainty and model random phenomena. Concepts such as conditional probabilities and event independence are also developed.

Finally, special attention is given to random variables, whether discrete or continuous, as well as their associated probability distributions, including the binomial, Poisson, and normal distributions. Through this pedagogical approach, engineering students acquire the foundational knowledge needed to model and analyze complex phenomena while developing a strong mathematical intuition.

TABLE OF CONTENTS

Chapter 1

Basic Definitions - One-Variable Statistical Series

1.1 Generalities :

The Purpose of Descriptive Statistics

Descriptive statistics aims to organize, summarize, and present the information contained in a dataset. It provides numerical and graphical summaries that help us understand the characteristics, patterns, and trends of the data.

1.1.1 Individuals or Statistical Units

Individuals and statistical units are fundamental concepts in statistics that refer to the basic elements from which data is collected and analyzed.

Each of the "individuals" studied, e.g.: human person, country, grades.

1.1.2 Population :

Set of observed individuals; Ex: students aged 12-25.

1.1.3 Population Size:

Population size refers to the total number of individuals that make up a population.

1.1.4 Sample:

A sample is a subset of individuals selected from a larger population. It is a group of individuals chosen to represent the characteristics of the entire population. Samples are used to collect data and make inferences about the population.

Example 1.1.1 :

Population : students in their second year of university

Statistical Units : students

Sample: Group of TD

1- **Characteristic (statistical variable):**

It is the property studied on these subjects

Example 1.1.2 :

Statistical variables: sex, height, number of children...

2- **Modalities:**

The different values that a variable can take.

Example 1.1.3 :

Statistical variable: the Exam mark; the modalities ∈ [0; 20].

Statistical variable: Genre; the modalities: male and female.

1.1.5 Types of variables

1- **Qualitative Variables**: A statistical variable is considered qualitative when its modalities are not measurable. It can be of the following types:

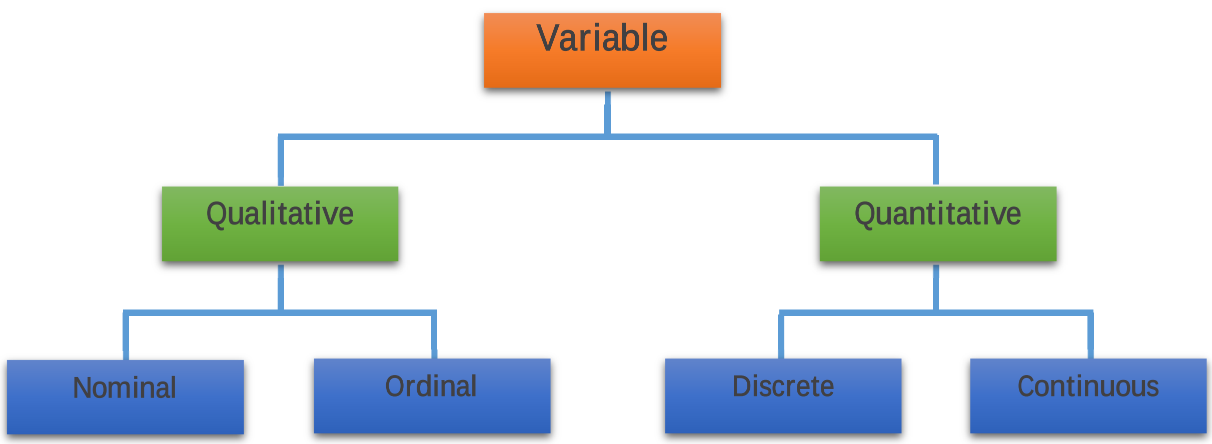
a. ***Nominal***: The groups are not ordered, for instance: eye colors.

b. ***Ordinal***: The groups are ordered, for example: baccalaureate grade (passable, assez-bien, bien)

2- **Quantitative Variables**: A statistical variable is considered quantitative when its modalities are measurable. It can be of the following types:

a. ***Discrete***: The modalities are countable, for example: the number of children per family.

b. ***Continuous:*** The modalities are defined on a continuous interval, for example: height, weight.



Exercise 1.1.4:

What is the type of statistical variation in the following cases?

- The statistical variable (color of houses in a neighborhood) is: qualitative.
- The statistical variable (students' grades) is: quantitative discrete.
- The statistical variable (students' heights) is: quantitative continuous.

1.2 Univariate Statistical Series

1.2.1 Organizing Data:

All data of the statistical series are grouped in a table indicating the distribution of individuals according to the studied characteristic. Grouping is done by classes:

If the characteristic is qualitative or discontinuous, a class contains all individuals having the same modality or the same value of the characteristic.

If the characteristic is continuous, a class is an interval.

- To construct these intervals, the following rules are observed:

The number of classes is between 5 and 20 (preferably between 6 and 12)

Whenever possible, the amplitudes of the classes are equal.

Each class (except the last) contains its lower limit but not its upper limit.

- In calculations, a class will be represented by its center, which is the midpoint of the interval.
- What must be indicated for each class?

1. The class size (effectif): Number of individuals in the class: denoted by n_i (i is the class index).

2. The frequency: Proportion of individuals in the population or sample belonging to the class: denoted by f_i.

f_i and n_i are related by : $f_i = \frac{n_i}{N}$

N : is the total number of individuals or data points in the population or sample

<u>*Note*</u>: f_i can be replaced by f_i x 100, which then represents a percentage. We always have :

$$\sum_{i=1}^{k} n_i = N, \ \sum_{i=1}^{k} f_i = 1, \ 0 \le f_i \le 1$$

Where k represents the number of classes.

3. Cumulative Frequency

Given a statistical series

Cumulative increasing frequencies: For a given value, the cumulative increasing frequency is the sum of the frequencies of all values less than or equal to the given value.

Cumulative decreasing frequencies: For a given value, the cumulative decreasing frequency is the sum of the frequencies of all values greater than or equal to the given value.

Example 1.2.1

Value	0	1	2	3	4	5
The class size (effectif)	3	3	12	6	3	3
Frequency	0.1	0.1	0.4	0.2	0.1	0.1
Cumulative deceasing effect	3	6	18	24	27	30
Incresing cumulative effect	30	27	24	18	6	3
Cumulative increasing frequencies	0.1	0.2	0.6	0.8	0.9	1
Cumulative decreasing frequencies	1	0.9	0.8	0.4	0.2	0.1

1.2.2 ***Data Visualizations:***

A. For a qualitative statistical variable:

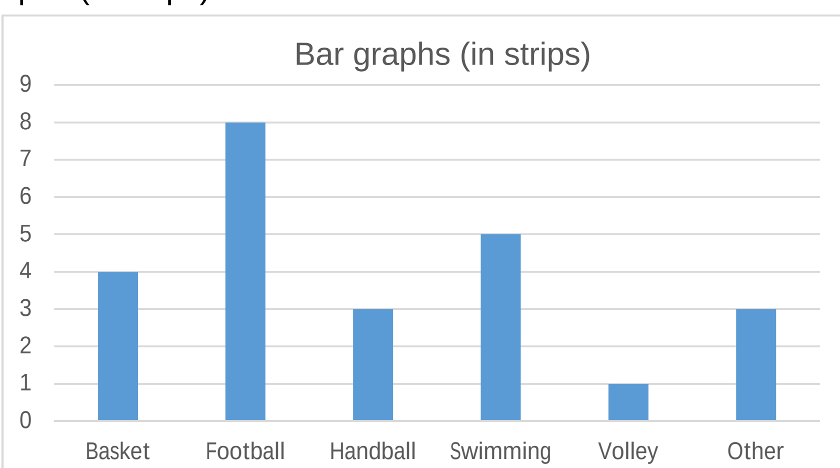
Example 1.2.2 :

24 middle school students were asked about their favorite subject. The results are presented in the table below:

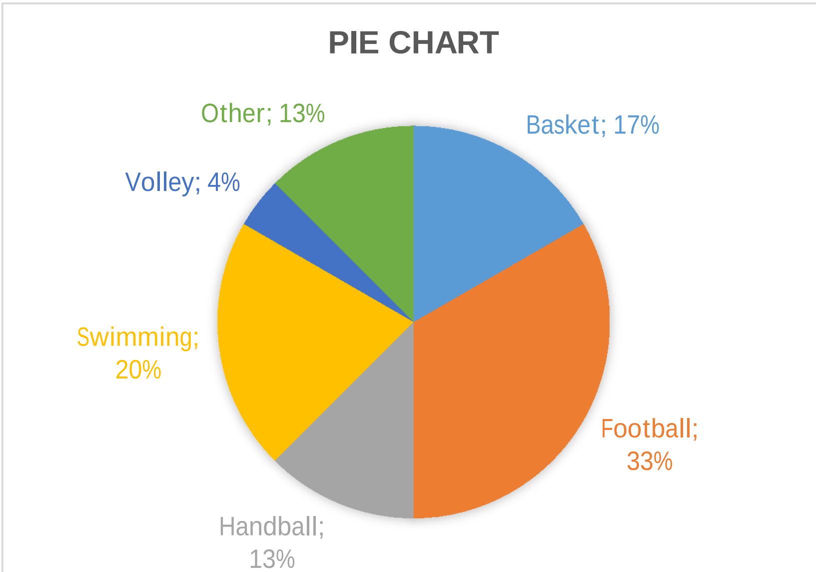
N = 24

Sport	Basket	Football	Handball	Swimming	Volley	Other
Frequency	4	8	3	5	1	3
Angle(°)	60	120	45	75	15	45

1- Bar graphs (in strips)



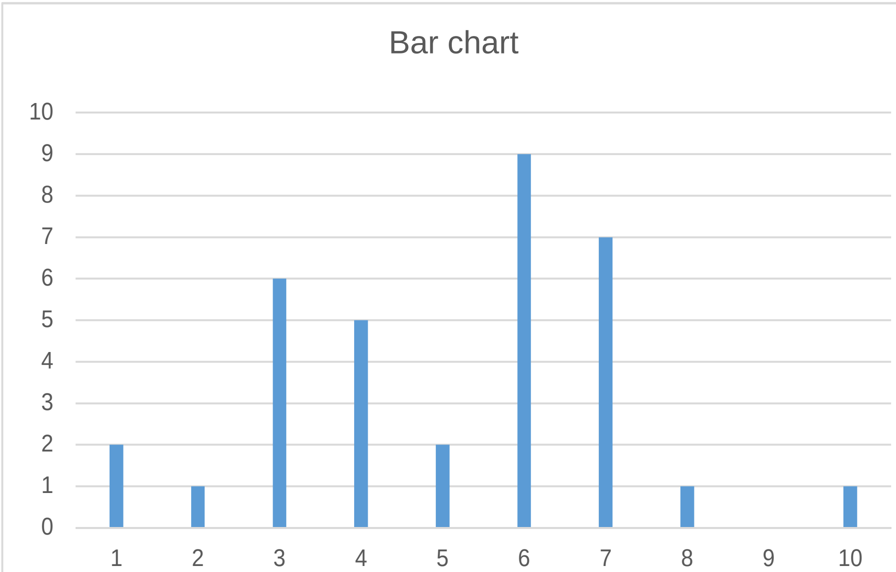
2- Pie or chart or Semi-circular chart:



B. For a discrete statistical variable:

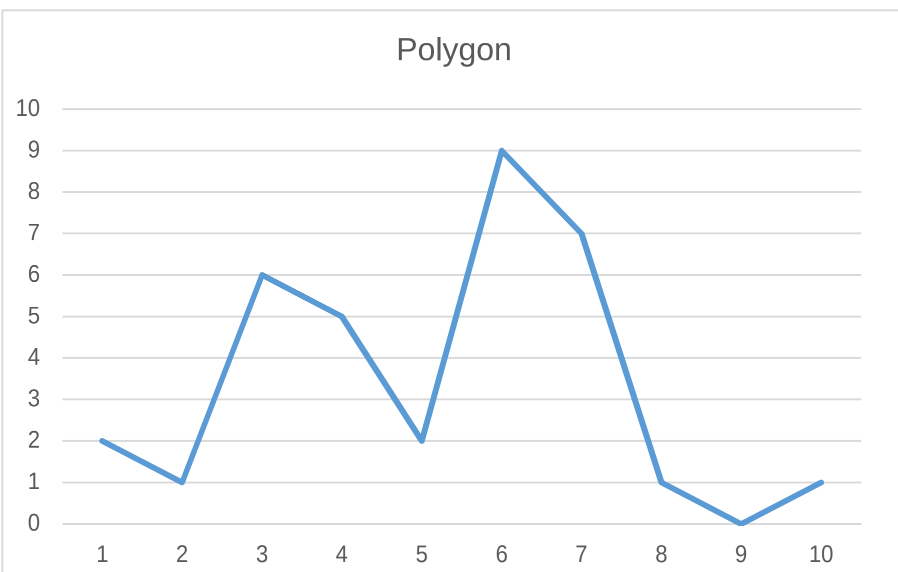
1. Bar charts

Grades	1	2	3	4	5	6	7	8	9	10
n_i	2	1	6	5	2	9	7	1	0	1



Sometimes the frequency is used instead of the count (effectif), which of course gives the diagram the same appearance.

2. The polygon of frequencies: The polygon of frequencies is obtained by connecting the ends of the bars of the previous histogram



C. For a continuous statistical variable:

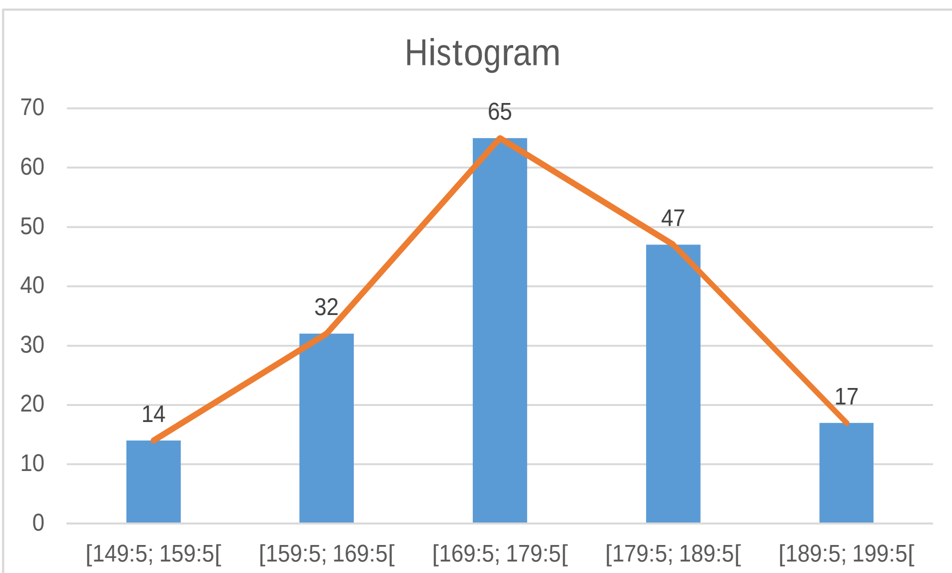
1- Histogram:

Example 1.2.3:

Let's resume the grouped distribution of the heights of the 175 students:

Class size	[149:5; 159:5[	[159:5; 169:5[	[169:5; 179:5[	[179:5; 189:5[	[189:5; 199:5[
Count n_i	14	32	65	47	17
Cumulative count	14	46	111	158	175
Frequency f_i	0.08	0.18	0.37	0.27	0.10
Cumulative frequency F_i	8%	26%	63%	90%	100%

The histogram corresponding to this data is:



2- Cumulative curves (cumulative frequency polygon)

The cumulative curve is obtained by plotting the points whose abscissas represent the upper bound of each class and ordinates the corresponding cumulative frequencies, then connecting these points by straight line segments. Its equivalent in probability theory is the distribution function.

1.2.3 Measures of central tendency

A. For a discrete statistical variable:

Example 1.2.4:

Statistics Exam Grades: N = 14

Grades x_i	2	4	6	8	9	10	11	12	13	14	15	16
Freq n_i	1	1	1	1	1	1	2	2	1	1	1	1

1- **The mode**: Mo: the most frequent value, Mo = 11; 12.
2- **The median**: Med: which divides the statistical series into two groups of equal frequency.

- If N =2n : Even Then; $Med = \frac{a_n + a_{n+1}}{2}$
- If N = 2n +1 Odd Then; $Med = a_{n+1}$

So $\frac{N}{2}$ = 7, and Med = 11

3- **The mean** (average) :

$$\overline{x} = \frac{\sum_{i=1}^{k} n_i x_i}{N} = \sum_{i=1}^{k} f_i x_i, \overline{x} = 10.21$$

B. For a continuous statistical variable:

Example 1.2.5: Class Grades

Class	[0; 4[	[4; 8[	[8; 12[	[12; 16[	[16; 20]
Count n_i	1	2	5	5	1
Center c_i	2	6	10	14	18
Freq f_i	0.07	0.14	0.36	0.36	0.07
Cum F_i	0.07	0.21	0.57	0.93	1

1- The Mode: is [8; 12[and [12; 16[
2- The median class: We have $\frac{N}{2}$ = 7, So The med class = [8; 12[

Med: The value corresponding to the 50% cumulative frequency on the cumulative frequency polygon

3- **The mean** (average) :

$$\overline{x} = \frac{\sum_{i=1}^{k} n_i c_i}{N} = \sum_{i=1}^{k} f_i c_i, \overline{x} = 10.88$$

1.2.4 Dispersion statistics (Measures of dispersion)

1- **The range :** The range of a statistical series is equal to the difference between the largest and smallest value in the series.

$$e = X_{max} - X_{min}$$

Example 1.2.6 :

A group of people of the following age 24, 16, 18, 22,16, 26, 35, 25, 15, 76.

Then we have the range e = 76 - 16 = 50 years

2- **The range :** V(x) is the average of the squared differences between each data point

$$V(x) = \frac{1}{N}\sum_{i=1}^{p} n_i (x_i - \overline{x})^2 = \sum_{i=1}^{p} f_i (x_i - \overline{x})^2$$

(x_i) and the mean (x); in other words:

3- **Standard deviation**: In statistics, the standard deviation is a measure of how spread out the values of a set of data are from their mean. It is calculated as the square root of the variance. The variance is the average of the squared deviations from the mean.

$$\sigma_x = \sqrt{V(x)}$$

Remark 1.2.1: The standard deviation is a finer parameter than the range, because it considers the distribution of the values. This means that it can be used to compare the dispersion of different series, even if they have different ranges.

- The standard deviation has the same unit as the values of the series being studied.
- The standard deviation measures the dispersion of the values of the series around the mean. The more concentrated the values of the series are around the mean, the smaller the standard deviation.
- The larger σ_x is, the more dispersed the observed series is.

Other (much more practical)

$$V(x) = \frac{1}{N}\sum_{i=1}^{p} n_i x_i^2 - \overline{x}^2 = \overline{x^2} - \overline{x}^2$$

Remark 1.2.2: Just as for the mean, to calculate a variance (or σ_x) for a continuous variable, we replace the x_i by c_i, the class centers:

$$V(x) = \frac{1}{N}\sum_{i=1}^{p} n_i c_i^2 - \overline{x}^2 = \overline{x^2} - \overline{x}^2$$

Example 1.2.1:

Calculate the variance and standard deviation of the following series:

$x_i\,(en\ €)$	$[0, 1600[$	$[1600, 2400[$	$[2400, 3200[$
c_i	*800*	*2000*	*2800*
n_i	*9*	*7*	*4*

It should be remembered $\overline{x} = 1620€, N = 20$ that :

Method 1:

$$V(X) = \frac{1}{20}\sum n_i (x_i - 1620)^2 = 631600\ €^2$$

Method 2:

$$\begin{aligned} \overline{x^2} &= \frac{1}{20}\left(9 \times 800^2 + 7 \times 2000^2 + ...\right) = 3256000\ €^2 \\ V(X) &= \overline{x^2} - \overline{x}^2 = 3256000 - 1620^2 = 631600\ €^2 \end{aligned}$$

Standard deviation:

$$\sigma_x = \sqrt{631600} = 794.7€$$

1.2.5 Coefficient of Variation

To compare the dispersion of two series that are not expressed in the same units, the coefficient of variation is used.

$$Cv = \frac{\sigma_x}{\overline{x}} =$$

It is the percentage variation relative to the mean, without units.

- The larger the coefficient of variation, the greater the dispersion.

<u>*Example 1.2.7*</u> (Series Comparison)

1- Let *x* be the statistical series of 4 products in Francs: 100F, 200F, 300F, 400F.

2- Let *y* be the statistical series of 4 products in euros: 15€, 30€, 45€, 60€.

Série	N	$\overline{x}$	$V(X)$	σ_x	Cv
1	4	$250F$	$12500F^2$	111.80F	0.45
2	4	$37,5$€	218.25€	16.77€	0.45

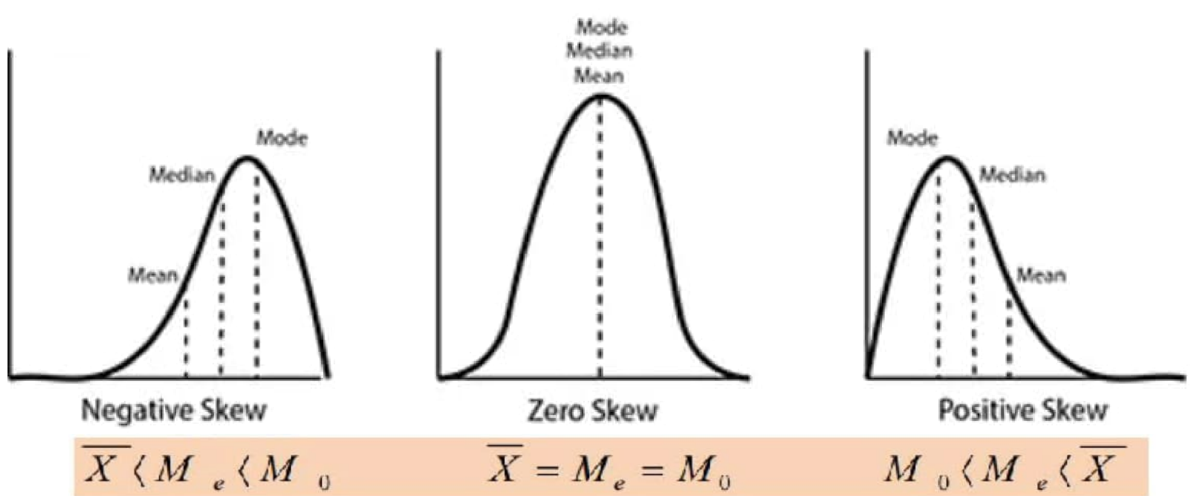
We have $Cv_1 = Cv_2$; the two series have the same dispersion.

1.2.6 Shape Characteristics

The statisticians have explained two criteria to describe the shape of such a curve: its symmetry and its flatness (Kurtosis).

1- Symmetry of a distribution

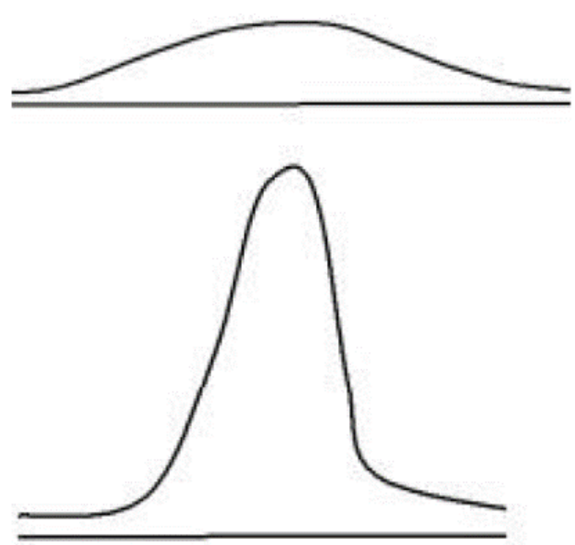
If the curve of a frequency distribution is symmetrical, then the values are distributed in the same proportions around the mode; and therefore $\bar{x}$ = Mo = Me:



2- Flattening (Kurtosis) of a distribution

In the case of a unimodal variable:

- The more dispersed the values, the flatter the distribution curve appears.
- The more the values are concentrated around the mode, the less flat the distribution curve appears.



Chapter 2

Bivariate statistical series

Definition 2.0.1:

A bivariate statistical series X and Y is defined as any list of pairs (x_1, y_1); (x_2, y_2); ...; (x_n, y_n):

x_i	x_1	$\dots$	x_n
y_i	y_1	$\dots$	y_n

n being the number of these pairs.

Example 2.0.1:

We consider 10 employees who are observed using two variables (age) and (salary). The data is given in the following table:

Salary	6000	7400	7500	8200	8207	8900	9100	9900	9950	10750
Age	21	26	20	43	47	37	52	34	50	44

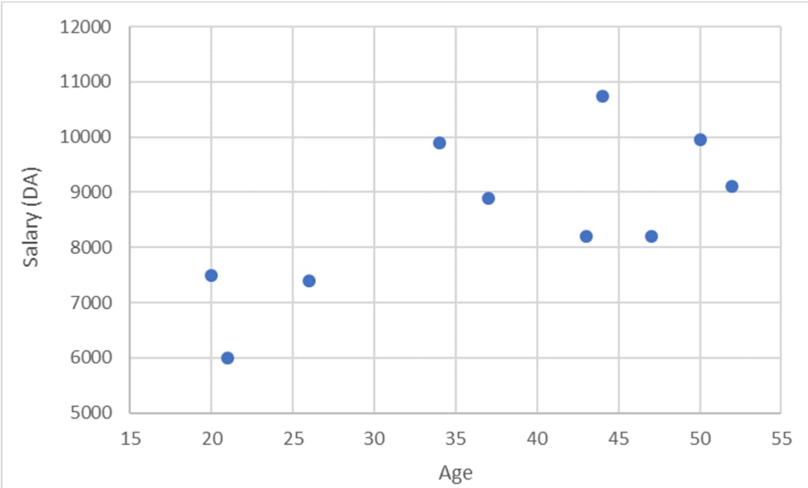
2.1. Point Cloud

Definition 2.1.1

Given a statistical series, we call the associated scatter plot the set of n points M_1, M_2, ..., M_n in the plane with coordinates (x_1, y_1), (x_2, y_2), ..., (x_n, y_n).

Example 2.0.1 (continued):

The scatter plot is plotted, from the raw data, in the following figure:



Definition 2.1.2:

The mean point of a scatter plot is the point G with coordinates ($\bar{x}$; $\bar{y}$) where $\bar{x}$ is the mean of $x_1, x_2, ..., x_n$ and $\bar{y}$ is the mean of $y_1, y_2, ..., y_n$: G($\bar{x}$; $\bar{y}$)

Example 2.0.1 (continued) :

$\bar{x} = 7595.7$; $\bar{y} = 36.8$

2.2 Statistical distribution of a pair of variables

We call the statistical distribution of the pair (X; Y) the grouping ((x_i; y_i) ; n_{ij}) where:

- $x_1, x_2, ..., x_k$ the *k* categories or levels of X
- $x_1, x_2, ..., x_l$ the *l* categories or levels of Y
- n_{ij} represents the number of observations that fall into both categories x_i and y_i

2.2.1 Contingency table

Representation of the distribution of the pair (X; Y); we use a double-entry table called a contingency table.

x/y	y_1	.....	y_j	....	y_l	Totals
x_1	n_{11}		n_{1j}		n_{1l}	n_{1*}
....	....		...		...	....
x_i	n_{i1}		n_{ij}		n_{il}	n_{i*}
....	....		....		....	....
x_k	n_{k1}		n_{kj}		n_{kl}	n_{k*}
Totals	n_{*1}		n_{*j}		n_{*l}	N

Marginal distribution of variable X $(x_i ; n_{i*})$

Marginal distribution of variable Y $(y_j ; n_{*j})$

2.2.2 Marginal distributions

We add row and column totals to the contingency table.

Definition 2.2.1 :

The k pairs $(x_i ; n_{i*})$ define the marginal distribution of variable X.

The k pairs $(y_j ; n_{*j})$ define the marginal distribution of variable Y.

Remark 2.2.1:

$$\begin{cases} n_{i\bullet} = \sum_{j=1}^{l} n_{ij} = \text{The total of the row } i \\ \sum_{i=1}^{k} n_{i\bullet} = N \\ n_{\bullet j} = \sum_{i=1}^{k} n_{ij} = \text{The total of the column } j \\ \sum_{j=1}^{l} n_{\bullet j} = N \end{cases}$$

Example 2.2.1 (continued)

1. Construct a contingency table with age (X) and salary (Y) as variables. For the age variable, create classes with a 10-year interval. For the salary variable, create classes with a 1000 Algerian dinars (DA) interval.
2. Calculate the marginal distributions of X and Y.

Ans.

Number of classes for the salary $= \frac{e}{a_{sal}} = \frac{10750-6000}{1000} = 4.75 \simeq 5\ classes$

Number of classes for age $= \frac{e}{a_{\hat{a}ge}} = \frac{52-15}{10} = 3.7 \simeq 4\ classes$

Therefore:

Age x/Salaire y	$[6,7[$	$[7,8[$	$[8,9[$	$[9,10[$	$[10,11[$	$n_{i\bullet}$	$f_{i\bullet}$
$[15,25[$	1	1	0	0	0	2	0.2
$[25,35[$	0	1	0	1	0	2	0.2
$[35,45[$	0	0	2	0	1	3	0.3
$[45,55[$	0	0	1	2	0	3	0.3
$n_{\bullet j}$	1	2	3	3	1	10	1
$f_{\bullet j}$	0.1	0.2	0.3	0.3	0.1	1	

Remark 2.2.2 :

- the frequency of the pair $(x_i, y_i)\ :\ f_{ij} = \frac{n_{ij}}{N}$,
- The marginal frequency (distribution) of y_j: $f_{\bullet j} = \frac{n_{\bullet j}}{N} = \sum_{i=1}^{k} f_{ij}$,
- The marginal frequency (distribution) of x_i: $f_{i\bullet} = \frac{n_{i\bullet}}{N} = \sum_{j=1}^{l} f_{ij}$,

Example 2.2.1 (continued)

3. Compute $f_{21}, f_{12}, f_{45}, f_{33}$.

$$f_{21} = \frac{n_{21}}{10} = 0 \qquad f_{45} = \frac{n_{45}}{10} = 0$$

$$f_{12} = \frac{n_{12}}{19} = 0.1 \qquad f_{33} = \frac{n_{33}}{10} = 0.2$$

4. Construct the frequency distribution table for the marginal series X and Y.

X	$[15, 25[$	$[25, 35[$	$[35, 45[$	$[45, 55[$
$n_{i\bullet}$	2	2	3	3
$f_{i\bullet}$	0.2	0.2	0.3	0.3
x_i le centre	20	30	40	50

Y	$[6, 7[$	$[7, 8[$	$[8, 9[$	$[9, 10[$	$[10, 11[$
$n_{i\bullet}$	1	2	3	3	1
$f_{i\bullet}$	0.1	0.2	0.3	0.3	0.1
y_i le centre	6.5	7.5	8.5	9.5	10.5

2.2.3 Conditional Distributions

Definition 2.2.2:

The conditional distribution of Y given X = x_i is defined as the distribution $((y_1; n_{i1}), \ldots, (y_l; n_{il}))$.

- This is the distribution of a single variable given by the i^{th} row of the contingency table (denoted $Y/_{X=x_i}$).

  Similarly, we can define the conditional distribution of X given Y = y_j as $((x_1; n_{1j}), \ldots, (x_k; n_{kj}))$

- This is the j^{th} column of the table (denoted $X/_{Y=y_j}$).

Example 2.2.1 (continued)

5. $X/_{Y=[7,8[}; Y/_{X=[45,55[}$

$X/_{Y=[7,8[}$	$[15,25[$	$[25,35[$	$[35,45[$	$[45,55[$
n_i	1	1	0	0

$Y/_{X=[45,55[}$	$[6,7[$	$[7,8[$	$[8,9[$	$[9,10[$	$[10,11[$
n_i	0	0	1	2	0

2.3 Covariance

Definition 2.3.1

The covariance of the bivariate statistical series of variables x and y is defined as the real number

$$COV(X,Y) = \sigma_{xy} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})(y_i - \overline{y})$$

We can also use for calculations:

$$\sigma_{xy} = \frac{1}{N}\sum_{i=1}^{N} x_i y_i - \overline{x}\overline{y}$$

Remark 2.3.1

1. $COV(X,X) = Var(X)$.
2. $COV(Y,Y) = Var(Y)$.

Proof:

1. $COV(X,X) = \frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})(x_i - \overline{x}) = \frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})^2 = Var(X)$.
2. $COV(Y,Y) = \frac{1}{N}\sum_{i=1}^{N}(y_i - \overline{y})(y_i - \overline{y}) = \frac{1}{N}\sum_{i=1}^{N}(y_i - \overline{y})^2 = Var(Y)$.

Example 2.3.1 *(continued)*

6. Compute $COV(X,Y), COV(X,X)$.

$$COV(X,Y) = \frac{1}{N}\sum_{i=1}^{N} x_i y_i - \overline{x}\overline{y} = \frac{1}{10}(600 \times 15 + ... + 10750 \times 44) - (7595.7 \times 36.8) = 10763.64$$

$$COV(X,X) = \frac{1}{N}\sum_{i=1}^{N} x_i^2 - \overline{x}^2 = \frac{1}{10}\left(15^2 + ... + 44^2\right) - (36.8)^2 = 148.16$$

2.4 Linear regression

The scatter plot associated with a bivariate statistical series thus immediately provides qualitative information.

To extract more quantitative information, we need to pose the problem of fitting.

The problem of establishing a functional relationship between the two series is the problem of fitting.

In the case of an elongated scatter plot, it is possible to replace this plot with a line called a linear regression line.

To draw this line, we use the following methods:

Mayer's method

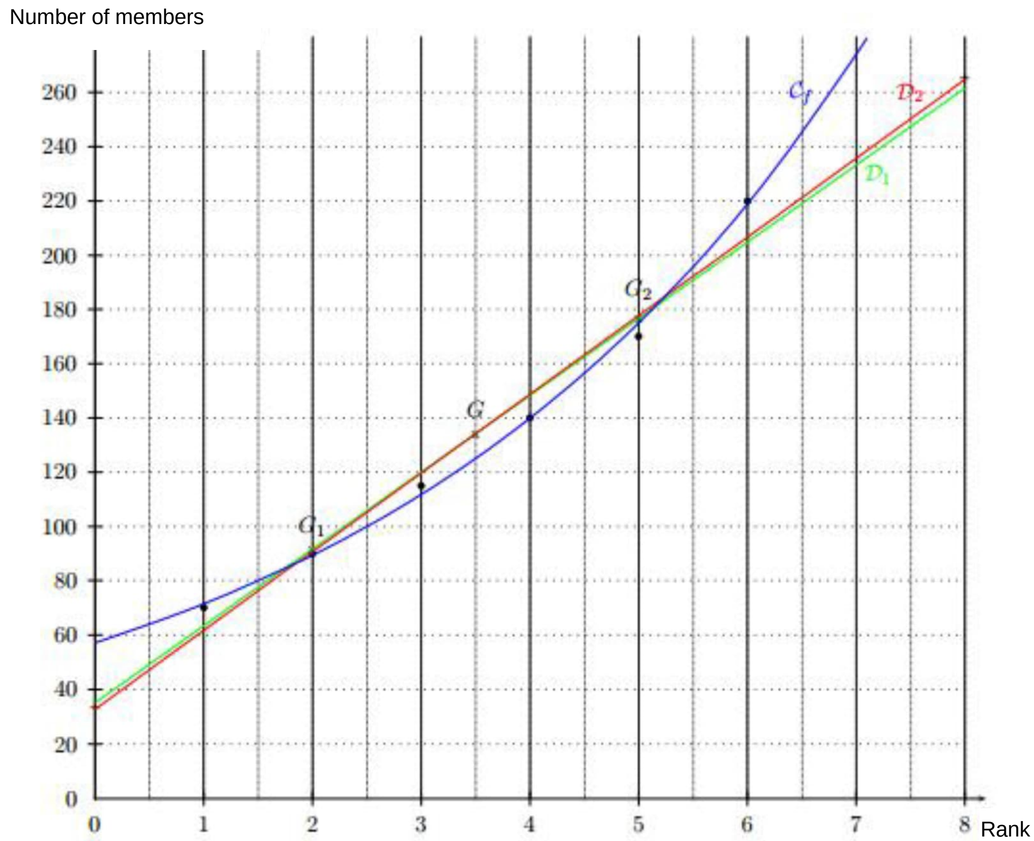
Example 2.4.1

The following table shows the evolution of the number of members of a football club from 2001 to 2006.

Year	2001	2002	2003	2004	2005	2006
Rank x_i	1	2	3	4	5	6
Number of members y_i	70	90	115	140	170	220

The goal is to study this bivariate statistical series (rank and number of members) in order to forecast the evolution of the number of members for the following years.

a- On a graph with orthogonal axes, using a scale of 2 cm per year on the x-axis and 1 cm for 20 members on the y-axis, plot the scatter plot associated with the series (x_i , y_i).



b- Calculer des coordonnées des points moyens G_1 et G_2 :
We divide the data into two equally sized groups according to the increasing values of x_i , and we calculate the coordinates G_1 and G_2 of each group of points:
G_1 for the years from 2001 to 2003,
G_2 for the years from 2004 to 2006.

$$G_1\left(\frac{1+2+3}{3};\frac{70+90+115}{3}\right), G_1\,(2;91.7)$$

$$G_2\left(\frac{4+5+6}{3};\frac{140+170+220}{3}\right), G_2\,(5;176.7)$$

c- The line (D1), represented by the equation y = ax + b, is the Mayer affine regression line that passes through the two points G_1 and G_2.

$$a=\frac{176.7-91.7}{5-2}=28.3$$

$$y_{G_1}=ax_{G_1}+b \Longleftarrow b=91.7-28.3\times 2=35.1$$

$$(D_1):y=28.3x+35.1$$

To construct (D_1), simply plot points G_1 and G_2, then draw the line segment joining them.

Method of least squares (regression line):

The goal is to obtain a line that is equidistant from the points located on either side of it.

To achieve this, we seek to minimize the sum of the squared distances of the points to the line.

We consider a bivariate statistical series represented by a scatter plot that justifies a linear fit.

This line is called the regression line of x on y.

Definition 2.4.1

The regression line D of y on x has the equation y = a*x* + b where:

$$a = \frac{COV(X,Y)}{V(X)} = \frac{\sigma_{xy}}{\sigma_x^2}$$
$$b = \overline{y} - a\overline{x}$$

Example 2.4.2:

Find the equation of the least squares regression line (D2) for y on x and add it to the previous graph.

===> (D_2) : $y = ax + b$ with $a = 29$ and $b = 32.7$.

(D_2) : $y = 29x + 32.7$

Draw

x	0	8
y	32.7	264.7

Comparison

Example 2.4.3:

Assuming the adjustments are still valid, predict the number of members in 2007

based on both methods.

- ⇨ we need to find the value of y when x is 2007 (rank 7).
- Mayer's method: y = 28; 3 x 7 + 35.1 = 233.2, or approximately 233 members.
- Adjusted estimate: y = 29 x 7 + 32.7 = 235.7, or approximately 236 members.

2.4.2 Linear correlation coefficient

The quantity $r_{xy} = \dfrac{COV(X,Y)}{\sigma_x \sigma_y}$ is the correlation coefficient

Remark 2.4.1

The linear correlation coefficient r_{xy} lies within the interval [-1, 1]; in other words,

$$|r_{xy}| \leq 1; -1 \leq r_{xy} \leq 1.$$

- The closer the absolute value of the linear regression coefficient is to 1, the better the linear fit;
- When r = ±1, the regression line passes through all points in the scatterplot, which are therefore perfectly aligned.
- The linear correlation coefficient r_{xy} justifies the use of linear regression.

$-0.7 \prec r_{xy} \prec 0.7 \Longrightarrow$ is not justified

$0.7 \leq r_{xy} \leq 1 \;\; et \; -1 \leq r_{xy} \leq -0.7 \Longrightarrow$ justified

$r_{xy} = \pm 1 \Longrightarrow$ Perfect

Chapter 3

Combinatorial Analysis

3.1 Arrangements

Consider a set E containing n elements (n is finite) and let's take a sequence (an arrangement) of p elements chosen from the elements of set E.

This sequence can be ordered or unordered, depending on whether we consider the position of the elements or not.

It can be done with or without repetition, depending on whether we can use the same element multiple times or only once.

Definition 3.1.1:

Let E be a set containing n distinct elements (n is finite). An arrangement without repetition of p elements ($0 \leq p \leq n$) is any ordered sequence of p distinct elements chosen from E.

The number of arrangements without repetition is : $A_n^p = \frac{n!}{(n-p)!}$

Note : The order is important

Example 3.1.1 :

From a group of 10 individuals, we aim to select a committee composed of a president, a secretary, and a treasurer. Assuming no individual can hold more than one position, determine the number of possible committees.

Given the distinct roles within the committee, each selection represents a arrangement. As no member can be chosen multiple times, these are arrangements without repetition.

As a result, we have $A_{10}^3 = \frac{10!}{(10-3)!} = 720$ possible committees.

Example 3.1.2 :

Four balls are drawn one at a time without replacement from an urn containing 10 balls numbered 0 to 9.
Each possible outcome is an ordered list of 4 different numbers from the set E = {0, 1, 2, . . . , 9}.
Therefore, the number of possible outcomes is : $A_{10}^4 = 5040$

Definition 3.1.2:

Let E be a set containing n distinct elements (n finite). An arrangement with repetition of p elements (p arbitrary) is any ordered sequence of p elements, each of which may be any element of E.

The number of arrangements with repetition is:

$$\bar{A}_n^p = n^p$$

Example 3.2.3:

We successively draw 4 balls with replacement from an urn containing 10 balls numbered from 0 to 9. Each possible outcome of this experiment is a 4-tuple of the set E = {0, 1, 2, . . . , 9}.

Therefore, the number of possible outcomes is : $\bar{A}_{10}^4 = 10^4$

3.2 Permutations

3.2.1 Basic Permutation

Definition 3.2.1:

A permutation is any ordered sequence of the n elements of a set E.

Note: A permutation is therefore a particular arrangement where p = n (all elements of E).

The number of different permutations is: Pn = n!

Example 3.2.1:

How many words can be formed with the letters of the word DIPLOMES?

n = 8, the number of different words is: N = P_8 = 8! = 40320

Example 3.2.2:

In how many ways can 4 math books, 3 physics books, and 2 chemistry books be arranged on a shelf?

There are N = 9! = 362880 different ways.

3.2.2 Permutations with Repetition

When some elements in the set E are identical:

Given n_1 elements of type 1, n_2 elements of type 2, ..., and n_k elements of type k, the number of distinct permutations is:

$$P_n^{n_1,n_2,\ldots,n_k} = \frac{n!}{n_1!\,n_2!\,\ldots.\,n_k!}$$

Example 3.2.3:

How many words can be formed with the letters of the word STATISTIQUES?

$n = 12$, $n_1 = 3$ (3 S), $n_2 = 3$ (3 T), $n_3 = 2$ (2 I)

The number of different words is :

$$P_{12}^{3,3,2} = \frac{12!}{3!\,3!\,2!} = 6652800$$

3.3 Combinations

3.3.1: Combinations Without Repetition

Definition 3.3.1:

A combination without repetition of p elements (where $0 \le p \le n$) is any unordered sequence of p distinct elements chosen from the elements of set E.

Note: A combination is an arrangement where the order does not matter.

The number of different combinations is:

$$C_n^p = \frac{n!}{p!\,(n-p)!}$$

Properties:

$C_1^0 = 1$, $C_n^n = 1$, $C_n^p = C_n^{n-p}$, $C_n^p = C_{n-1}^p + C_{n-1}^{p-1}$

3.3.2: Combinations with Repetition

Definition 3.3.2:

A combination with repetitions of p elements (where p is any number) is any unordered sequence of p elements, not necessarily distinct, chosen from the elements of E.

$$K_n^p = \frac{(n+p-1)!}{p!\,(n-1)!}$$

Example 3.3.1:

From a group of 5 men and 7 women, how many different committees composed of 2 men and 3 women can be formed?

There is no order, so there are $C_5^2 C_7^3 = 10 * 35 = 350$possible committees.

Example 3.3.2:

Six balls are drawn simultaneously from an urn containing 49 balls numbered from 1 to 49:

Each possible outcome of this experiment is an unordered sequence of 6 distinct elements of the set E = {1, 2, ..., 49}.

Therefore, the number of possible outcomes is: $C_{49}^6 = 13983816$

3.3.3: Binomial Theorem

Let a and b ∈ R, and n ∈ N $\quad (a+b)^n = \sum_{i=0}^{n} C_n^i a^{n-i} b^i$

Example 3.3.3:

1- $(a+b)^2 = C_2^0 a^2 + C_2^1 ab + C_2^2 a^0 b^2 = 1a^2 + 2ab + 1b^2$

2- $(a+b)^3 = C_3^0 a^3 + C_3^1 a^2 b + C_3^2 a^1 b^2 + C_3^3 a^0 b^2 = 1a^3 + 3a^2 b + 3a^1 b^2 + 1b^3$

Conditions	The number of possible draws is the number of	A Common example:
The p elements are not necessarily all distinct but are ordered	p-lists of elements of E, which is: n^p	Successive draws with replacement of p objects from n
The *p* elements are all distinct and ordered	Arrangements of p elements of E, which A_n^p	Successive draws without replacement of p objects from n
The *n* elements are all distinct and ordered	Permutations of n elements of E, which n!	Anagrams of a word with no repeated letters
The *p* elements are all distinct and unordered.	Combinations of p elements of E, which $\binom{p}{n}$	Simultaneous draws of p objects from a set of n

Chapter 4

Introduction to Probability

4.1 Definitions:

4.1.1. Random Experiment:

Definition 4.1.1:

A random experiment is any experiment whose outcome is governed by chance and, if repeated under identical conditions, can yield different results.

Examples of random experiments:

- Rolling a 6-sided die and observing the total.
- Tossing a coin.

4.1.2. Sample Space

Definition 4.1.2:

The sample space of a random experiment is the set of all possible outcomes of that experiment, denoted by Ω.

Examples 3.1.1:

The sample spaces associated with the random experiments presented in the previous example are respectively:

1. Ω = [1, 2, 3, 4, 5, 6]
2. Ω = [F, P]

4.1.3. Event

Definition 4.1.3:

An event related to a random experiment is a subset of the sample space Ω. Events are usually denoted by A, B, C, ...

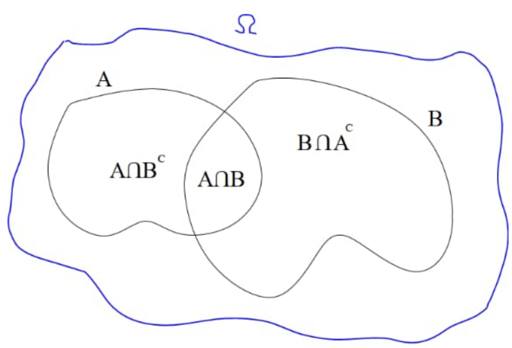
- An event that contains only one outcome is an elementary event.
- The event that contains no outcomes is the impossible event, denoted by ∅.
- The event composed of all outcomes is called the certain event.

Example 3.1.2:

Rolling a six-sided die:

- The sample space: Ω = {1, 2, 3, 4, 5, 6}.
- Obtaining a 2 is an outcome of this random experiment.
- A: "obtaining a 5" is an elementary event, which can be denoted as A = {5}.
- B: "obtaining an even number" is an event, which can be denoted as B = {2, 4, 6}.
- Obtaining a 7 is an impossible event.
- Obtaining a positive number is a certain event.

4.2 Algebra of events



4.2.1. Intersection and Union

Definition 4.2.1:

The union of two sets A and B, denoted A ∪ B, is the set consisting of all elements of Ω that belong to either A or B. In other words:

$A \cup B = \{\omega \in \Omega \mid \omega \in A \text{ or } \omega \in B\}$.

Definition 4.2.2:

The intersection of two sets A and B, denoted by A∩B, is the set consisting of the elements of Ω that belong to both A and B. In other words:

$A \cap B = \{\omega \in \Omega / \omega \in A \text{ and } \omega \in B\}$.

Remark 4.2.1:

If $A \cap B = \emptyset$, we say that the events A and B are incompatible (disjoint) or are mutually exclusive events.

Example 4.2.1:

Consider the set of numbers from 1 to 10.

Let A be the event "obtaining an even number" and B be the event "obtaining a number strictly less than six".

• A ∩ B: "obtaining an even number and strictly less than six" A ∩ B = {2, 4}.

• A ∪ B: "obtaining an even number or a number strictly less than six" A ∪ B = {1, 2, 3, 4, 5, 6, 8, 10}.

4.2.2. The Complement

Definition 4.2.3:

The complement of the set A, denoted $\overline{A}$ (or Ac), is the set consisting of all elements of Ω that do not belong to A. In other words:

$\overline{A} = \{\omega \in \Omega \,/\, \omega \notin A\}$.

Remark 4.2.2:

In particular, we have $A \cup \overline{A} = \Omega$ and $A \cap \overline{A} = \emptyset$.

4.2.3. Symmetric Difference

Definition: The symmetric difference of sets A and B, denoted A ▵ B, is the set consisting of the elements of Ω that belong to A ∪ B but do not belong to A ∩ B. In other words:

$A \triangle B = \{\omega \in \Omega \,/\, \omega \in A \cup B \text{ and } \omega \notin A \cap B\} = (A \cup B) - (A \cap B)$.

notations	vocabulaire ensembliste	vocabulaire probabiliste
Ω	ensemble plein	événement certain
$\emptyset$	ensemble vide	événement impossible
ω	élément de Ω	événement élémentaire
A	sous-ensemble de Ω	événement
$\omega \in A$	ω appartient à A	ω réalise A
$A \subset B$	A inclus dans B	A implique B
$A \cup B$	réunion de A et B	A ou B
$A \cap B$	intersection de A et B	A et B
A^c ou $\overline{A}$	complémentaire de A	événement contraire de A
$A \cap B = \emptyset$	A et B disjoints	A et B incompatibles

4.3 Probability Spaces

4.3.1 Probability

Definition 4.3.1:

The probability of an event is a numerical value that represents the proportion of times the event will occur when the experiment is repeated under identical conditions. From this definition, we can deduce that a probability must be between 0 and 1, and that the probability of an event is the sum of the probabilities of each of the elementary events that compose it. Finally, the sum of the probabilities of all elements of Ω is 1.

Remark 4.3.1:

An event is nothing more than a subset of Ω. A probability associates with each event a number between 0 and 1. It is therefore a function from the power set of Ω, denoted $\mathcal{P}(\Omega)$, into [0, 1].

4.3.2. Power Set

Definition 4.3.2:

The power set of Ω, denoted P(Ω), is the set of all subsets of Ω.

Example 4.3.1:

For the universe Ω = {a, b, c}, we have: $\mathcal{P}(\Omega) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \Omega\}$.

Remark 4.3.2:

• $\mathcal{P}(\Omega)$ always contains the empty set ∅ and the set Ω itself.

• The elements of $\mathcal{P}(\Omega)$ are the subsets of Ω, not the elements of Ω. In fact: $A \in \mathcal{P}(\Omega) \Longleftrightarrow A \subset \Omega$.

4.3.3. Probability Space

Definition 4.3.3:

A probability space is a pair $(\Omega, \mathcal{A})$, where $\mathcal{A}$ is a σ-algebra on Ω.

A probability on the sample space Ω is a function P such that:

- P: $\mathcal{P}(\Omega) \rightarrow [0, 1]$.
- P(Ω) = 1.
- The probability of an event in the sample space Ω is the sum of the probabilities of the elementary events that compose it

$$P\left(\bigcup_n A_n\right) = \sum_n P(A_n).$$

Remark 4.3.3:

We say that there is equiprobability when all elementary events have the same probability.

In this case, we have: $P(A) = \frac{Number\ of\ elements\ of\ A}{Number\ of\ elements\ of\ \Omega} = \frac{Card(A)}{Card(\Omega)}$

Example 4.3.2:

An urn contains 1 white ball and 1 black ball. If we draw 2 balls with replacement. What is the probability of getting 2 black balls?

We have Ω = {(B, B); (W, W); (W, B); (B, W)}. Card(Ω) = 4

Let A be the event of getting 2 black balls. Card(A) = {(B, B)}.

$$\Rightarrow P(A) = \frac{1}{4}.$$

Example 4.3.3:

A jury of 3 people is chosen from a group of 4 men and 6 women. What is the probability that the 3 people chosen are 2 men and 1 woman?

Possible outcomes Card(Ω): $C_{10}^{3} = \frac{10!}{3!*7!} = 120$.

Favorable outcomes Card(A) : $C_4^2 * C_6^1 = 36$

$\Rightarrow P(A) = \frac{36}{120} = 0.3$

4.3.2. Probability Space

Definition 4.3.4:

A probability space is a triple (Ω, A, P), where A is a σ-algebra on Ω and P is a probability measure.

Properties

Let A and B be two events, we have the following properties:

- $P(\emptyset) = 0$.
- $P(\Omega) = 1$.
- $0 \leq P(A) \leq 1$.
- $P(\overline{A}) = 1 - P(A)$.
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
- If A and B are mutually exclusive events, then $P(A \cup B) = P(A) + P(B)$.

Example 4.3.4:

Let's say we have a set E containing the integers from 1 to 20. If we randomly pick a number from this set:

Event A is when the chosen number is a multiple of 3: A = {3, 6, 9, 12, 15, 18}.
Event B is when the chosen number is a multiple of 2: B = {2, 4, 6, 8, 10, 12, 14, 16, 18, 20}.

Let's calculate the probabilities:

- P(A) = $\frac{6}{20} = \frac{3}{10}$ = 0.3
- P($\bar{A}$) = 1 − P(A) = 1 − $\frac{3}{10} = \frac{7}{10}$ =0.7
- P(B) = $\frac{10}{20}$ = 0.5
- P(A ∩ B) = $\frac{3}{20}$ = 0.15
- P(A ∪ B) = P(A) + P(B) − P(A ∩ B) = $\frac{6}{20} + \frac{10}{20} - \frac{3}{20} = \frac{13}{20}$ = 0.65

Chapter 5

Conditional Probability and Independence

5.1. Conditional Probability

Definition 5.1.1:

Let A and B be two events such that P(B) ≠ 0.

The conditional probability of A given B, denoted P(A|B), is the probability that event A occurs given that event B has already occurred. This probability is equal to

$$P(A/B) = P_B(A) = \frac{P(A\cap B)}{P(B)}$$

Remark 5.1.1 :

We might also see P(A/B) used instead of $P_B(A)$

Example 5.1.1:

Consider the random experiment of rolling a fair 6-sided die. We assume that all faces are equally likely, and we define the events:

• A: « the face obtained has a number that is a multiple of 3 ».

• B: « the face obtained has an even number ».

Let us determine the probability of obtaining a number that is a multiple of 3, given that we have an even number, in two different ways.

• The event (A/B) corresponds to the event "obtaining a number that is a multiple of 3 among the possibilities of B", in other words among {2, 4, 6}. There is therefore only the outcome "obtaining a 6" that corresponds. Thus, we obtain: $P_B(A) = \frac{1}{3}$

By computation, we find: P(B) = $\frac{3}{6}$ and P(A ∩ B) = 1/6.

So, using the formula:

$$P_B(A) = \frac{P(A\cap B)}{P(B)} = \frac{1/6}{3/6} = 0.33$$

Proposition 1:

For any events A and B with non-zero probabilities, we have:

$P(A \cap B) = P(B) \times P_B(A) = P(A) \times P_A(B)$.

Proposition 2:

If S is an event that can happen (non-zero probability), then

- $0 \leq P_S(A) \leq 1$.
- $P_S(\Omega) = 1$.
- $P_S(\emptyset) = 0$.
- $P_S(\bar{A}) = 1 - P_S(A)$.
- $P_S(A \cup B) = P_S(A) + P_S(B) - P_S(A \cap B)$.
- If events A and B are incompatible (cannot happen at the same time), then : $P_S(A \cup B) = P_S(A) + P_S(B)$.
- $P_S(\bar{A} \cap \bar{B}) = P_S(\overline{A \cup B}) = 1 - P_S(A \cup B)$.

Theorem 1: (Law of Total Probability):

For any two events A and B with non-zero probabilities: $P(A) = P_B(A) \times P(B) + P_{\bar{B}}(A) \times P(\bar{B})$.

Theorem 2: (Bayes' Theorem)

For any events A and B with non-zero probabilities:

$$P_A(B) = \frac{P_B(A) * P(B)}{P(A)}$$

This formula actually gives us $P_A(B)$ directly, without any extra calculations.

Example 5.1.2:

In a workshop, two machines, M_1 and M_2, cut identical metal pieces. M_1 provides 60% of the production (of which 6.3% are defective), with the rest being provided by M_2 (of which 4% of the production is defective).

The daily production consists of the pieces produced by both machines, and at the end of the evening, one piece is randomly selected (all selections are assumed to be equally likely).

1. The probability of selecting a defective piece, given that it is produced by machine M_1, is: $P_{M_1}(D) = 0.063$.

2. The probability of selecting a defective piece, given that it is produced by machine M_2, is: $P_{M_2}(D) = 0.04$.

3. The probability of selecting a defective piece:

Using the law of total probability, we have

$P(D) = P(M_1 \cap D) + P(M_2 \cap D)$

$= P(M_1) \times P_{M_1}(D) + P(M_2) \times P_{M_2}(D)$

$= 0.6 \times 0.063 + 0.4 \times 0.04$

$= 0.0538.$

4. If we select a defective piece, let's calculate the probability that it was produced by machine M_1:

Using Bayes' theorem, we have:

$$P_D(M_1) = \frac{P_{M1}(D) * P(M_1)}{P(D)} = \frac{0.063 * 0.6}{0.0538} = 0.703$$

5.2 Independent Events

Definition 5.2.1:

We say that A and B are independent events if and only if : $P(A \cap B) = P(A) \times P(B)$.

Example 5.2.1:

A fair coin is tossed twice.

Let the events be:

A: "Getting heads on the first toss".

B: "Getting heads on the second toss".

C: "Getting heads-tails or tails-heads".

We will show that events A, B, and C are pairwise independent but not mutually independent.

We have: $\Omega = \{PP, PF, FP, FF\}$.

$P(A) = P(\{PP, PF\}) = 1/2$

$P(B) = P(\{PP, FP\}) = 1/2$

$P(C) = P(\{PF, FP\}) = 1/2$

$P(A \cap B) = P(\{PP\}) = 1/4 = P(A) \times P(B) \Rightarrow$ A and B are independent.

$P(A \cap C) = P(\{PF\}) = 1/4 = P(A) \times P(C) \Rightarrow$ A and C are independent.

$P(B \cap C) = P(\{FP\}) = 1/4 = P(B) \times P(C) \Rightarrow$ B and C are independent.

$P(A \cap B \cap C) = P(\emptyset) = 0 \neq P(A) \times P(B) \times P(C)$

⇒ A, B, and C are not mutually independent.

Thus, events A, B, and C are pairwise independent but not mutually independent.

Proposition:

If A and B are independent events, then: A and not B; not A and B ($\bar{B}$); not A ($\bar{A}$) and not B ($\bar{B}$) are also independent events.

Chapter 6

Random Variables

6.1. Definitions and Properties

Let (Ω, P) be a probability space.

Definition 6.1.1:

A variable is a mapping.

Let (Ω,𝔸,P) be a probability space. A discrete random variable X associated with this probability space is a mapping from the sample space Ω into R ($X: \Omega \to \mathbb{R}$) that assigns a numerical value to each outcome of the random experiment under study.

$$\sum_{x \in X} P(X = x) = 1$$

Let X be a discrete random variable with elements denoted by $x_1, x_2, \ldots, x_i, \ldots, x_n$, where n is a natural number ($n \in \mathbb{N}$). For each i,($i \in \mathbb{N}$) we define:

$$P_X(X = x_i) = P_i = P(X^{-1}(x_i)) \quad \text{and} \quad \sum_{i=1}^{n} P_i = 1$$

Example 6.1.1:

We flip 3 coins and observe the number of heads obtained. It is clear that X takes the values: 0, 1, 2, 3.

Then the application: $X : \Omega \to \mathbb{R}$ is a random variable such that

Ω = *{(hhh) ; (hht) ; (hth) ; (thh) ; (htt) ; (tht) ; (tth) ; (ttt)}* and X(Ω) = {0; 1; 2; 3}

moreover, we have:

Value of X (event)	[X = 0]	[X = 1]	[X = 2]	[X = 3]
Structure of the event	*{hhh}*	*{hht , hth , thh}*	*{htt , tht , tth}*	*{ttt}*
Probability	1/8	3/8	3/8	1/8

$$P(X = 0) = P[(hhh)] = \left(\frac{1}{2}\right)^3$$

$$P(X = 1) = P[\,(hht)\,,(hth)\,,(thh)] = 3\left(\frac{1}{2}\right)^3$$

$$P(X = 2) = P[\,(htt)\,,(tht)\,,(tth)] = 3\left(\frac{1}{2}\right)^3$$

$$P(X = 3) = P[(ttt)] = \left(\frac{1}{2}\right)^3$$

These results are summarized in the following table:

X	0	1	2	3	$\sum_{x=0}^{3} P(X=x)$
P_X	$\left(\frac{1}{2}\right)^3$	$3\left(\frac{1}{2}\right)^3$	$3\left(\frac{1}{2}\right)^3$	$\left(\frac{1}{2}\right)^3$	1

Remark 6.1.1:

If X and Y are two discrete random variables defined on the same probability space, then X + Y is also a discrete random variable defined on the same space.

6.2. Cumulative Distribution Function

Definition 6.2.1:

The cumulative distribution function of a random variable X is the function associated with its distribution, that is, the function F_X: ℝ → [0,1] defined by:

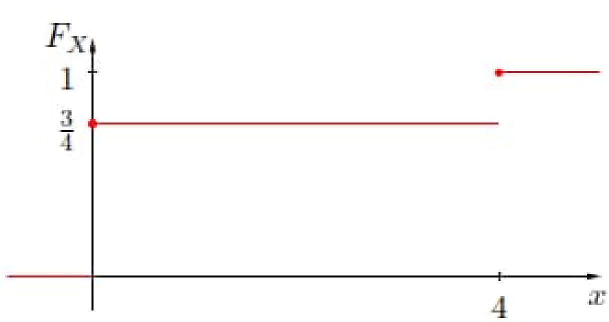
$$F(x) = P_X(\,]-\infty, x]) = P(X \leq x)$$

Example 6.2.1:

Two coins are tossed successively; Ω = {HH, HT, TH, TT}. Suppose a player bets his fortune of 1 DA on the following game based on this random experiment: each time a tail comes up, his fortune doubles, but if a head comes up, he loses everything. The random variable X giving his fortune at the end of the game is given by:

X(HH) = X(HT) = X(TH) = 0, X(TT) = 4.

The cumulative distribution function of this random variable is given by the following figure:



$$F_X(x) = \begin{cases} 0 & si\ x < 0, \\ \frac{3}{4} & si\ 0 \le x < 4, \\ 1 & si\ x \ge 4. \end{cases}$$

Example 6.2.3:

Continuing from Example 6.2.1, focusing on the number of faces.

x	*hhh*	*hht*	*hth*	*thh*	*tth*	*tht*	*htt*	*ttt*
Value of X	3	2	2	2	1	1	1	0

The distribution of X is: P [X = 0] = 1/8, P[X = 1] = P[X = 2] = 3/8, P[X = 3] = 1/8

Where:

$$F(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1/8 & \text{si } 0 \le x < 1, \\ 4/8 & \text{si } 1 \le x < 2, \\ 7/8 & \text{si } 2 \le x < 3, \\ 1 & \text{si } x \ge 3 \end{cases}$$

Proposal 1:

1. $P(X > x) = 1 - F_X(x)$,

2. $P(x < X \le y) = F_X(y) - F_X(x)$

6.3 Expected Value

6.3.1 Expectation:

Definition 6.3.1:

The expected value of the discrete random variable X is the number:

$$E(X) = \sum_{i=1}^{n} P_i\, x_i$$

Remark 6.3.1:

The expected value of a discrete random variable X is the probabilistic version of the arithmetic mean.

Proposal 2: Let X and Y be two discrete random variables defined on the same probability space. Then:

1/ $E(X + Y) = E(X) + E(Y):$

2/ $E(aX + b) = aE(X) + b; \forall (a; b) \in \mathbb{R}^2:$

6.3.2 Variance:

Definition 6.3.2:

Let X be a random variable. The variance of X is the expected value of the squared deviations of X from its mean:

$$Var(X) = \sum_{i=1}^{n} P_i \, (x_i - E(x))^2$$

Proposal 3: Let X be a discrete random variable and let a, b be two real numbers.

1-$Var(X) = E(X^2) - [E(X)]^2$

2/ $Var(aX + b) = a^2 Var(X)$

Example 6.3.1:

We roll a die and assume we win 100 DA if we get a 1; 50 DA if we get a 2 or 3; 20 DA if we get a 4; and we lose 50 DA if we get a 5 or 6.

Let X be the random variable representing the winnings.

We have the probability distribution of X:

x_i	-50	20	50	100
P_i	1/3	1/6	1/3	1/6

The expected value of the gain is:

$E(X)$ = -50*1/3 + 20*1/6 + 50*1/3 + 100*1/6 = 20 DA

So, the variance is:

Var(X) = 2500*1/3 + 400*1/6 + 2500*1/3 + 10000*1/6 = 3400

6.3.2 Standard Deviation:

The standard deviation, denoted by $\sigma(X)$, of a discrete random variable X is the square root of the variance of X.

$$\sigma(X) = \sqrt{Var(X)}$$

6.3.3 Covariance:

<u>*Definition 6.3.3:*</u>

Let Ω be a finite probability space and let X and Y be two random variables on Ω. The covariance of the random variables X and Y is the number, denoted $Cov(X,Y)$ (or σ_{XY}), defined by:

$Cov(X,Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]$.

<u>Properties :</u>

- $Cov(X, X) = Var(X)$
- If X and Y are independent, then $Cov(X, Y) = 0$. The converse is false.
- $Cov(X + a, Y + b) = Cov(X, Y)$
- $Cov(aX, bY) = ab.Cov(X, Y)$

Chapter 7:

Common Discrete Probability Distributions

7.1. Bernoulli Distribution

Definition 7.1.1:

A Bernoulli trial is a random experiment with only two possible outcomes. We are interested in whether or not a specific event occurs. In other words, we only study random experiments with two possible outcomes. The trial is considered a success if the event of interest occurs and a failure otherwise.

We associate with this trial a random variable X that takes the value 1 with probability "p" if the event occurs and the value 0 with probability "q" otherwise. This random variable therefore takes only two values (0 and 1) and its distribution is given by:

$$P[X = 1] = p, P[X = 0] = q = 1 - p$$

The Bernoulli distribution is summarized in the table below:

X	0	1
P_X	q	p

We write $X \rightarrow B(p)$ or $X \sim B(1 ; p)$

Examples 7.1.1:

1- A hospital patient either survives or does not survive,

2- A customer either signs a contract or does not,

3- Flipping a coin once, we call getting Heads a success and getting Tails a failure,

4- A voter votes either Democrat or Republican...

Proposal 1: If $X \sim B(1 ; p)$, then:

1- The cumulative distribution function:
$FX(0) = P(x \leq 0) = q = 1 - p$
$FX(1) = P(x \leq 1) = 1$

2- The mathematical expectation is :
$E(X) = p$

3- The variance is : $Var(X) = E(X^2) - (E(X))^2$

$= p - p^2$

$= p*(1 - p)$

$= p*q$

4- The Standard deviation is: $\sigma(x) = \sqrt{pq}$

<u>Example 7.1.2</u>:

We roll a six-sided die and are interested in obtaining a multiple of 2 or 3. Consequently, success S = "obtaining a multiple of 2 or 3"

Define X as a binary indicator: $X = \begin{cases} 1 \text{ if the outcome is a multiple of 2 or 3} \\ 0, \text{ otherwise} \end{cases}$

The distribution of X :

X	0	1
$P(X=x_i)$	1/3	2/3

Then, we have:

- The expected value: E(X) = p = 2/3
- The Variance: Var(X) = p*(1-p) = 1/2* 1/2=1/4
- The Standard deviation: $\sigma(x) = \sqrt{p*(1-p)} = \frac{1}{2}$

7.2. Binomial probability distribution:

<u>Definition 7.2.1:</u>

We perform n ($n \in \mathbb{N}^*$) independent repetitions of a Bernoulli trial with probability of success p. We define the random variable X as the number of successes among the n obtained results. Then X follows a binomial distribution with parameters n, p, and q such that q = 1 - p is the probability of failure:

we note X ~ B(n; p; q). Where X = {0; 1; 2;…….; n}

Remark: Generally, the binomial distribution B(n; p; q) is the distribution of a sum X of n independent random variables, each following the same Bernoulli distribution.

<u>Proposition</u>:

1- Let X ~ B(n, p, q), where q = 1 - p, then 1- the probability distribution of X:

$P(X = x) = C_n^x * p^x * (1-p)^{n-x}$, with x = 1; 2; 3; ...; n

2- The probability distribution function :

$$F_X(x) = \sum_{k=0}^{x} C_n^k * p^k * (1-p)^{n-k}$$

3- The expected value: E(X) = n*p

4- The variance: Var(X) = n*p*(1-p)

5- The standard deviation: $\sigma(x) = \sqrt{n * p * (1-p)}$

Example 7.2.1:

We toss 3 coins and observe the number of heads obtained. It is clear that X takes the value of: X → B(3 ; $\frac{1}{2}$; $\frac{1}{2}$); X = {0; 1; 2; 3}

- The expected value: E(X) = n*p = $\frac{3}{2}$
- The variance: Var(X) = n*p*q = ¾
- The standard deviation: $\sigma(x) = \sqrt{n * p * q} = \frac{\sqrt{3}}{2}$

Example 7.2.2:

A coin is tossed 5 times in a row. Let's denote the random variable indicating "the number of heads obtained".

1- What is the probability of getting exactly 3 heads?
2- 2- What is the probability of getting at least 1 head?

Answer:

X → B(5 ; $\frac{1}{2}$; $\frac{1}{2}$); X = {0; 1; 2; 3; 4; 5}

1- $P(X = 3) = C_5^3 * (\frac{1}{2})^3 * \left(1 - \frac{1}{2}\right)^{5-3} = \frac{5}{16}$

2- $P(X \geq 1) = 1 - P(X < 1)$

$= 1 - P(X = 0)$

$= 1 - C_5^0 * (\frac{1}{2})^0 * \left(1 - \frac{1}{2}\right)^{5-0}$

= 1 – 0.3125 = 0.9688

7.3. Poisson Distribution

Definition 7.3.1:

The Poisson distribution is a discrete probability distribution commonly used in the study of rare events under certain conditions. For example, X: the number of people over 100 years old in a population. The Poisson distribution also describes the behavior of the number of events occurring in a fixed interval of time, if these events occur with a known average or expected frequency.

This distribution is an approximation of the binomial distribution when np is small and n is large (in practice, n ≥ 50 and np ≤ 10).

Definition 7.3.2:

Let λ ∈ $R^{*,+}$. A discrete random variable X is said to follow a Poisson distribution with parameter λ, denoted $X{\sim}P(\lambda)$, when the set of values taken by X is the set of all non-negative integers: X(Ω) = ℕ. Additionally, we have:

$\forall\ x \in \mathbb{N}$, $P(X=x) = \frac{\lambda^x}{x!}\ e^{-\lambda}$ where (e =2.718).

Example 7.3.1:

An emergency service receives an average of 5 fractures per weekend. What is the probability of observing 3 fractures during the next weekend?

Answer: Poisson distribution with λ = 5 and x = 3

$P(X=3) = \frac{5^3}{3!}\ e^{-5} = 0.14$

Proposition: Let X be a discrete random variable following a Poisson distribution, $X{\sim}P(\lambda)$, with λ∈ R*+, then:

1- Probability distribution $P(X=x) = \frac{\lambda^x}{x!}\ e^{-\lambda}$, $\forall\ x \in \mathbb{N}$

2- Cumulative distribution function: $F_X(x) = \sum_{k=0}^{x} \frac{\lambda^x}{x!}\ e^{-\lambda}$

3- The mathematical expectation of X is: E(X) = λ

4- The variance of X is : Var(X) = λ

5- The standard deviation of X is: $\sigma(X) = \sqrt{\lambda}$

Example 7.3.2:

Consider the random variable X: number of typos per page in the Mathematics course. $X \sim P(\lambda)$, such that λ = 0.1.

The probability of having one error per page is: P(X=1) $= \frac{0.1^1}{1!} e^{-0.1} = 0.09$

7.4 Poisson Approximation to the Binomial Distribution

Let X be a discrete random variable following the binomial distribution X ~ B(n; p; q) with n large (n ≥ 30) and p small (p ≤ 0.1). This distribution can be approximated by the Poisson distribution with parameter λ = np ≤ 5.

Example 7.4.1:

In a population, one person in a hundred is a centenarian. We define the discrete random variable X: number of centenarians in a population of 200 people. X follows the binomial distribution with parameters n = 200, p = 0.01, and q = 1 - p = 0.99. Since n > 30 and p < 0.1, then we can approximate the distribution of X by the Poisson distribution P(λ) such that λ = np = 2.

Example 7.4.2:

We randomly draw 400 calls. Let X be the random variable that gives the number of calls with disturbances. The probability that a call is disturbed is p = 0.005.

1- Calculate the probability of finding 7 calls with disturbances.

Ans.

X ~ B (n = 400 ; p = 0.005)

$$P(X = 3) = C_{400}^{7} * (0.005)^3 * (1 - 0.005)^{400-7} = !!!!??$$

Since: n = 400 > 30, p = 0.005 < 0.1, and λ = np = 400 x 0.005 = 2 ≤ 5, then we can use the Poisson distribution.

P(X=x) $= \frac{2^x}{x!} e^{-2}$

P(X=7) $= \frac{2^7}{7!} e^{-2} = 0.0034$

2- Calculate the following probabilities:

a. There are exactly 2 calls with disturbance

b. There are between 2 and 4 calls with disturbance

c. There are at most 3 calls with disturbance

d. There are at least 4 calls with disturbance

Bibliography

[1] Pierre Dagnélie. Statistique théorique et appliquée. De BoeckUniversité, 1998.

[2] Rick Durrett. Elementary probability for applications. Cambridge university press, 2009.

[3] Richard Arnold Johnson etGouri K. Bhattacharyya. Statistics : principles and methods. Wiley, 1996.

[4] Aurelio Mattei. Inférence et décision statistiques : théorie et application à la gestion des affaires. P. Lang, 2000.

[5] Sheldon M. Ross. Initiation aux probabilités. Presses polytechniques et universitaires romandes, 2007.

[6] Gilbert Saporta. Probabilités, analyse des données et statistique. Technip, 1990

[7] Michael J. Evans and Jeffrey S. Rosenthal. Probability and Statistics: The Science of Uncertainty. University of Toronto

Overview

This document serves as a valuable resource for engineering students seeking to gain a deep understanding of fundamental concepts in probability and statistics. It covers statistical series for both one and two variables, enabling rigorous data analysis and interpretation.

The first part focuses on one-variable statistical series, emphasizing central tendency indicators (mean, median, mode) and dispersion measures (variance, standard deviation). The second part addresses two-variable statistical series, introducing key concepts such as covariance and correlation, which are essential for establishing relationships between different quantities.

Furthermore, the document explores probability theory in a finite framework, providing students with the necessary tools to quantify uncertainty and model random phenomena. Concepts such as conditional probabilities and event independence are also developed.

Finally, special attention is given to random variables, whether discrete or continuous, as well as their associated probability distributions, including the binomial, Poisson, and normal distributions. Through this pedagogical approach, engineering students acquire the foundational knowledge needed to model and analyze complex phenomena while developing a strong mathematical intuition.

Keywords :

One-variable statistical series, mean, median, mode, variance, standard deviation, two-variable statistical series, covariance and correlation, probability theory, random variables, probability distributions.