الجمهورية الجزائرية الديمقراطية الشعبية
People's Democratic Republic OF Algeria
وزارة التعليم العالي و البحث العلمي
MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH
جامعة الشهيد الشيخ العربي التبسي-تبسة
Echahid Cheikh Larbi Tebessi University-Tebessa
كلية العلوم و التكنولوجيا
Faculty of sciences and technology
قسم العلوم والتكنولوجيا
Department of science & technology

Probability and Statistics
Courses and Exercises

2nd Science & Technology Licene

*Dr:**Bouaziz KHELIFA***

2024/2025

Contents

Introduction

This course focuses on Probability and Statistics. It is intended for second year ST license students. It can serve as a teaching aid, because it contains almost all the elements of the basis of probability and descriptive statistics, all the propositions are demonstrated with illustrations by examples and figures, in addition to the exercises solved in each chapter.
The first chapter contains definitions and basic concepts in statistics.
The second chapter contains important theories in probability.
The last chapter is a series of application exercises around the content presented in the courses

Chapter 1

Statistics

1.1 Basic definitions

We define statistics as the science that categorizes and analyzes data

Terminology in Statistics

Statistical community: The elements that we want to study

Statistical unit: An element of the statistical population

Sample: A randomly selected subset of the statistical population

Qualitative statistical variable: Not related to numbers for example blood type, color, gender

Quantitative statistical variable: Related to numbers for example height measurement, students' results

Discrete statistical variable (x_i)Values are isolated (59, 70, ...)

Continuous statistical variable Values are in classes $[150cm - 160cm]$,

Frequency Distribution (f_i) The number of values that have the same attribute

Relative Frequency $\left(Rf_i = \frac{f_i}{N}, N = f_1 + f_2 + ...f_p\right)$

Example 1.1 *A box of coloring pencils*

Color	*Frequency*	*Relative Frequency*
Brown	18	$\frac{18}{80}$
Red	12	$\frac{12}{80}$
Yellow	10	$\frac{10}{80}$
Orange	20	$\frac{20}{80}$
Green	5	$\frac{5}{80}$
Blue	15	$\frac{15}{80}$
Total	*80*	$\frac{80}{80} = 1$

Decreasing cumulative frequency: It is the sum of the frequency of the value x_i and the frequencies of all values occurring after x_i
Growing cumulative frequency: It is the sum of the frequency of the value x_i and the frequencies of all preceding values of x_i

Example 1.2 *We have the following series*

Class interval	$[1-11[$	$[11-21[$	$[21-31[$	$[31-41[$	$[41-51[$	$[51-61[$	$[61-71[$
Frequency	6	12	22	37	17	8	5
Growing Cf	6	18	40	77	94	102	107
Decreasing Cf	107	101	89	67	30	16	5

Classes: Values are in domains $[150cm - 160cm]$, $[160cm - 170cm]$
Class boundaries: 150, 160, ...
Class width: 160-150=10
Midpoint: $\frac{150+160}{2} = 155$

1.2 One-variable statistical series

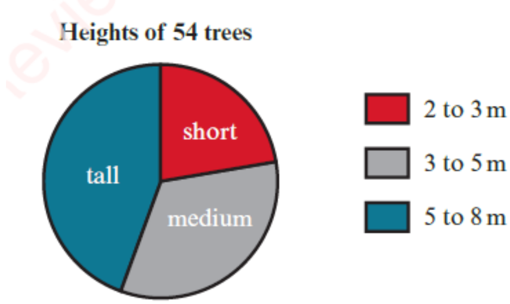
1.2.1 Graphical Representation of Data

1.2.2 Pie Chart

A pie chart is a circular chart divided into sectors.

Example 1.3 *University students measured the heights of the 41 trees in*

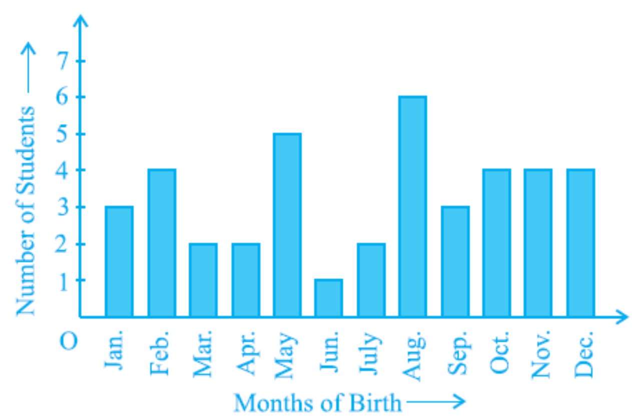
heights of trees	*Frequency*
short $[2 \text{ to } 3m[$	8
medium $[3 \text{ to } 5m[$	9
tall $[5 \text{ to } 8m[$	24



1.2.3 Bar charts

We define bar charts to represent the frequencies.

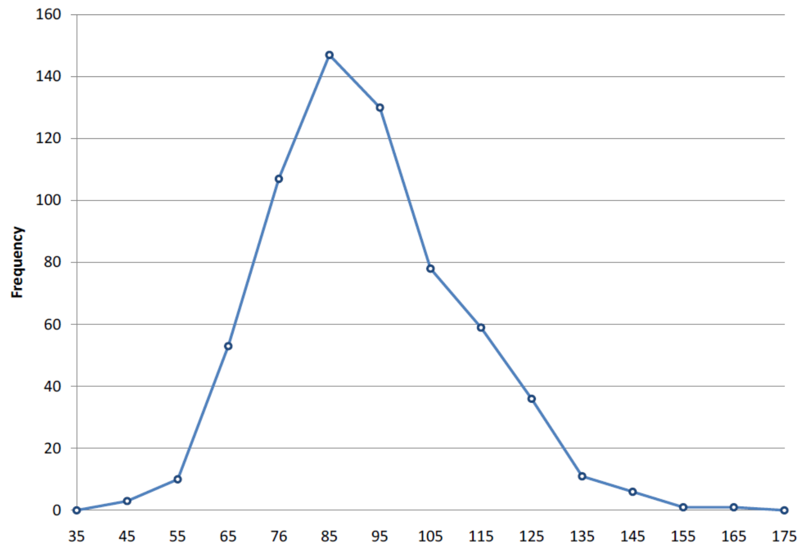
Example 1.4 *In a particular section of Class, 40 students were asked about the months of their birth and the following graph was prepared for the data so obtained.*

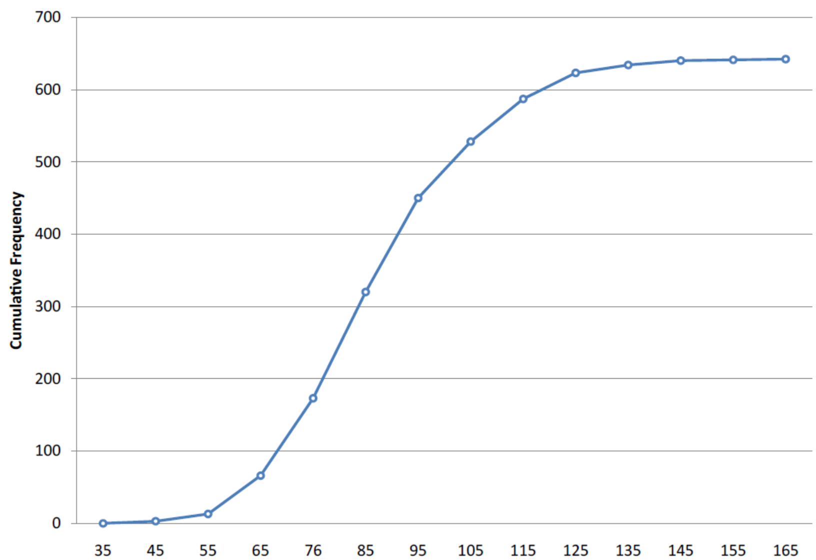


1.2.4 Frequency polygons

Example 1.5 *Suppose we are given the following series*

Values	35	45	55	65	75	85	95	105	115	125	135	145	155	165
Frequency	0	3	10	53	107	147	130	78	59	36	11	6	1	1
CF(less than)	0	3	13	66	173	320	450	528	587	623	634	640	641	642





1.2.5 Measures of Central tendency

1.2.6 Arithmetic mean ($\overline{X}$):

Arithmetic mean : is equal to dividing the sum of the values by their number

Example 1.6 *We have the following observations: 10, 7,30, 15, 42, 83 and 79*

$$\overline{X} = \frac{\sum x_i}{N} = \frac{(10+7+30+15+42+83+79)}{7} = 38$$

Remark 1.1 *In the case of frequencies*

$$\overline{X} = \frac{\sum_{i=1}^{i=p} x_i f_i}{N} \; / f_1 + f_2 + ... + f_p = N$$

Arithmetic mean of grouped data

If the data are given as a continuous frequency distribution, we take $x_i = m_{i,}$ where m_i the midpoint of the class.

Example 1.7 *the mean of the following distribution.*

Class	*6-10*	*11-15*	*16-20*	*21-25*	*26-30*	*31-35*	
mid-point m_i	*8*	*13*	*18*	*23*	*28*	*33*	
$frequency\ f_i$	*35*	*23*	*15*	*12*	*9*	*6*	*100*
$m_i f_i$	*280*	*299*	*270*	*276*	*252*	*198*	*1575*

$\overline{X} = \frac{\sum_{i=1}^{i=6} m_i f_i}{N} = \frac{1575}{100} = 15.75$

1.2.7 Median (Me)

We first arrange the values in ascending or descending order and then read the center value.

- If n is odd, $M_e = X_{\frac{n+1}{2}}$.
- If n is even, $M_e = \frac{1}{2}\left[X_{\frac{n}{2}} + X_{\frac{n}{2}+1}\right]$

Example 1.8 *The median of the following numbers.*

a) 5, 6, 2, 4, 9, 8.

b) 27, 14, 5, 6, 8,1, 2

Solution:

a) First order the data: 2, 4, **5, 6**, 8, 9

Here n=6

$$\begin{aligned} M_e &= \frac{1}{2}\left[X_{\frac{6}{2}} + X_{\frac{6}{2}+1}\right] \\ &= \frac{1}{2}\left[X_3 + X_4\right] \\ &= \frac{1}{2}[5+6] = 5.5 \end{aligned}$$

b) Order the data :1, 2, 5, **6**, 8, 14, 27

Here n=7

$$M_e = X_{\frac{7+1}{2}=X_4=6}$$

Remark 1.2 *the median in the case of classes is calculated by the formula:*

$Me = I_{1+}\frac{I_2-I_1}{P}(\frac{n}{2} - Q)$. *Where* M_e *the median*

I_1: *the lower limit of the median class.*

I_2: *the upper limit*

P: *the frequency*

n : *the total number of items*

Q: *the cumulative frequency of the class preceding the median class.*

Example 1.9

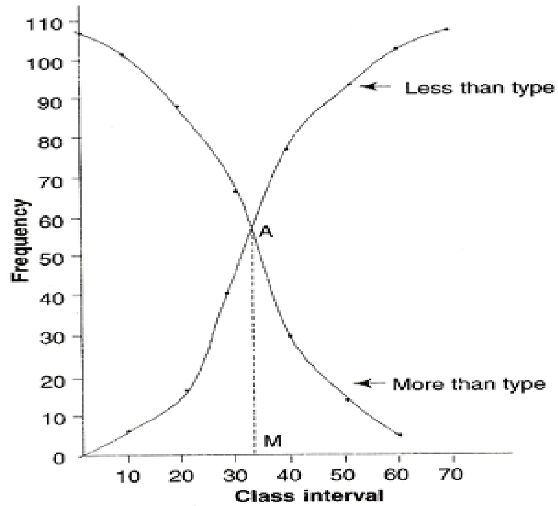
Monthly Wages	Frequency	Cumulative Frequency
800 -1,000	18	18
1,000 -1,200	25	43
1,200 -1,400	30	73
1,400 -1,600	34	107
1,600 -1,800	26	133
1.800 -2,000	10	143

$\frac{n}{2}=\frac{143}{2}$=*71.5.*

The median class is the class-interval 1,200 - 1,400

$$\begin{aligned} M_e &= 1200 + \frac{1400-1200}{30}(71.5-43) \\ &= 1389.81 \end{aligned}$$

Remark 1.3 *The median is the \(x \)-coordinate at the point where the ascending and descending* *curves intersect.*



1.2.8 Mode(Mo)

The mode is the most frequent value

Example 1.10 *1. The mode of 3, 8, 5, 5 and 9.*
Mode =5.
2. The mode of 8, 8, 9, 7, 9, 6 and 2.
It is a bimodal Data: 8 and 9
3. The mode of 12, 6, 3, 4 and 7.
No mode for this data.

1.2.9 Measures of Variability

1.2.10 Range(R)

The range of a statistical series is the difference between the largest value and the smallest value.
$R = X_{max} - X_{min}$

Example 1.11 *The range of 10, 4, 7, 6, 5, 3 and 2 is :*
R= 10 - 2 = 8.

1.2.11 Variance

This variance is the "average squared deviation from the mean".

$$
\begin{aligned}
V &= \frac{1}{N}\sum f_i \left(X_i - \overline{X}\right)^2 \\
&= \left(\frac{1}{N}\sum f_i X_i^2\right) - \overline{X}^2
\end{aligned}
$$

1.2.12 Standard deviation (σ).

Where σ^2 (called sigma squared) is used to denote the variance.

$$\sigma = \sqrt{V}$$

Example 1.12 *We have the following distribution:*

class	f_i	m_i
$[0-10[$	1	5
$[10-20[$	3	15
$[20-30[$	6	25
$[30-40[$	10	35
$[40-50[$	12	45
$[50-60[$	11	55
$[60-70[$	6	65
$[70-80[$	3	75
$[80-90[$	2	85
$[90-100[$	1	95

$X_i = m_i$ *(Mid points)*

$\overline{X} = \frac{1}{N}\sum f_i m_i = \frac{1}{55}(5+45+150+350+540+605+390+225+170+95) = 46.81$

$V = \left(\frac{1}{N}\sum f_i X_i^2\right) - \overline{X}^2 = 2545 - (46.81)^2 = 353.82$

$\sigma = \sqrt{353.82} = 18.7$

1.2.13 Coefficient of variation

The coefficient of deviation expression is

$$\mathbf{CV} = \frac{\sigma \times 100}{\overline{X}}$$

Example 1.13 *An olive oil press has two machines* (A) *and* (B)*. Machine* A *produces an average of* 30 *bottles per day with a standard deviation of* 6 *bottles, while machine* B *produces an average of* 45 *bottles per day with a standard deviation of* 10*. Which machine shows greater regularity in its daily production?*

Solution:
We have the coefficient of variation for machine A is $CV_A = \frac{\sigma_A \times 100}{\overline{X_A}}$,
we get $CV_A = \frac{6 \times 100}{30}$
$CV_A = 20\%$,
the coefficient of variation for machine B is $CV_B = \frac{\sigma_B \times 100}{\overline{X_B}}$
$CV_B = \frac{10 \times 100}{45} = 22.2\%$.
The machine A is more consistent than machine B.

1.2.14 Measures of Skewness

We know the coefficient of Pearson

$$C_P = \frac{\overline{X} - M_o}{\sigma}$$

The curve is determined by the value of C_P

1. $C_P > 0$: the distribution is positively skewed.
2. $C_P = 0$: the distribution is symmetric.
3. $C_P < 0$: the distribution is negatively skewed.

1.3 Statistical series with two variables

X_i	x_1	x_2	...	x_n
Y_i	y$_1$	y$_2$	...	y$_n$

Binaries (X_i, Y_i) form a variables statistical series with two variables

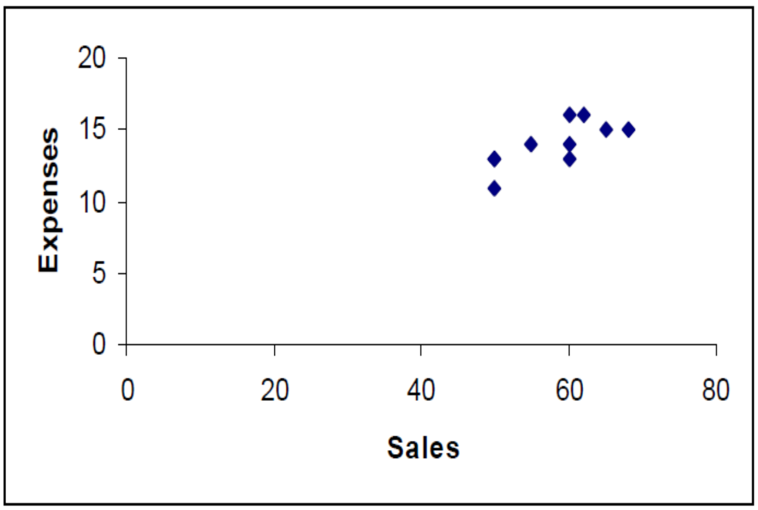
1.3.1 Scatter diagram

The scatter refers to the dots on the graph.
The middle point is the point $G\left(\overline{X}_G, \overline{Y}_G\right)$ where,

$$\overline{X}_G = \frac{\sum_{i=1}^{n} x_i}{N} \text{ and } \overline{Y}_G = \frac{\sum_{i=1}^{n} y_i}{N}$$

Example 1.14 *In a firm, The following data on sales and expenses for 10 month:*

$Month$	J	F	M	A	M	J	J	A	S	O
$Sales(in\ thousand)$	50	50	55	60	62	65	68	60	60	50
$Expenses(in\ thousand)$	11	13	14	16	16	15	15	14	13	13



The scatter diagram

$\overline{X} = \frac{50+50+55+60+62+65+68+60+60+50}{10} = 58,$
$\overline{Y} = \frac{11+13+....+13+13}{10} = 14.$
The middle point is $G(58, 14)$

1.3.2 Linear adjustment

Find a linear function that approximately expresses Y in terms of X

Linear regression

It is a straight line that includes the middle point $G\left(\overline{X}, \overline{Y}\right)$ of equation

$$y = ax + b,$$

where

$$\begin{aligned} a &= \frac{\left(\frac{1}{n}\sum x_i y_i\right) - \overline{XY}}{\left(\frac{1}{n}\sum x_i^2\right) - \overline{X}^2} \\ b &= \overline{Y} - a\overline{X} \end{aligned}$$

Example 1.15 *We have the table*

X	200	300	400	500	600
Y	30	35	41	44	47

$G\left(\overline{X},\overline{Y}\right) =?$

$\overline{X} = \frac{200+300+400+500+600}{5} = 400$, $\overline{Y} = \frac{30+35+41+44+47}{5} = 39.4$

$G\left(400, 39.4\right)$

$$\begin{aligned} a &= \frac{\left(\frac{1}{5}\right)\left(6000+10500+16400+22000+28200\right) - 15760}{\frac{1}{5}\left(40000+90000+160000+250000+360000\right) - 400^2} \\ &= \frac{860}{20000} = 0.043 \end{aligned}$$

and

$$\begin{aligned} b &= \overline{Y} - a\overline{X} \\ &= 39.4 - (0.043)\,400 = 22.2. \end{aligned}$$

The regression line is described by the equation: $y = 0.043x + 22.2$

Mayer Line Method

1) Arrange the coordinates in ascending order based on the x-values.
2) Split the data into two equal groups. If the number of coordinates is odd, assign the extra point to the group that is closest in value to it.
3) Calculate the mean for each group; denote these means as $M_1$and M_2..
4) Determine the equation of the line (in the form $y = ax + b$) that passes through the points $M_1$and M_2.

Example 1.16 : *The table below presents the coordinates of a scatter plot:*

X	10	6	5	1	7	7	12	6	16	15
Y	7	11	14	15	9	12	6	13	2	6

Step 1: Ensure the data is sorted in ascending order based on the x-values.

X	1	5	6	6	7	7	10	12	15	16
Y	15	14	13	11	12	9	7	6	6	2

Step 2: Divide the data into two equal groups (which is feasible with this dataset).

Here's a continuation of your text with the calculations for $M_1$and M_2.

Step 3: Calculate the Means

Point M_1: Calculate the mean of Group 1 $(1, 5, 6, 6, 7)$.

x*-coordinate* $= \frac{1+5+6+6+7}{5} = \frac{25}{5} =$ *5*

y*-coordinate* $= \frac{15+14+13+11+12}{5} = \frac{65}{5} =$ *13*

M_1 *(5, 13)*

Point M_2: Calculate the mean of Group 2 $(7, 10, 12, 15, 16)$.

x*-coordinate* $= \frac{7+10+12+15+16}{5} = \frac{60}{5} =$ *12*

y*-coordinate* $= \frac{9+7+6+6+2}{5} = \frac{30}{5} =$ *6*

M_2 *(12, 6)*

Step 4 Find the equation of the line ($y = ax + b$) that goes through points $M_1$and M_2 :

Slope: $a = \frac{y_2 - y_1}{x_2 - x_1} = \frac{6-13}{12-5} = -1$

Find the "b":

by substituting any one of your points in the equation: $y = -1x + b$

Use point M_1 $(5, 13) : 13 = (-1)(5) + b \Rightarrow b = 18$

The equation of the regression line (Mayer) is: $y = -x + 18$

Least squares

$$S = \sum (y_i - ax_i - b)^2$$

The Straight line closest to the scatter diagram is the straight line where S is small

Chapter 2

Probability

2.1 Combinatory analysis

Notation 2.1 *We denote n factorial by $n!$. as*

$n! = n \times (n-1) \times (n-2) \times ... \times 2 \times 1$

$n! = n \times (n-1)!$

By definition, $0! = 1.$

$1! = 1$

$4! = 4 \times 3 \times 2 \times 1 = 24$

2.1.1 List

(Arrangement and repetition):

The possible ordered selections of r elements from a total of n elements represent the number of lists n^r

2.1.2 Arrangement

(Arrangement without repetition):

are the possible ordered selections of r elements out of a total of n elements, The number of arrangements is denoted by A_n^r

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

Example 2.1 *: The digits $0, 1, 2, ..., 9$ are to be used in 3 digit identification card.*

How many different cards are possible if

a) Repetitions are permitted.

b) Repetitions are not permitted.

Solution:

a) There are three steps

1. Selecting the 1st digit, this can be made in 10 ways.
2. Selecting the 2nd digit, this can be made in 10 ways.
3. Selecting the 3rd digit, this can be made in 10 ways.

$10^3 = 10 \times 10 \times 10 = 1000$ different cards are possible.

b) There are three steps

5. Selecting the 1st digit, this can be made in 10 ways.
6. Selecting the 2nd digit, this can be made in 9 ways.
7. Selecting the 3rd digit, this can be made in 8 ways.

$$\begin{aligned} A_{10}^{3} &= \frac{10!}{(10-3)!} \\ &= 10 \times 9 \times 8 = 720, \end{aligned}$$

then 720 different cards are possible.

2.1.3 Combinations

(without arrangement without repetition):

are the possible selections of r elements from a group of n elements, The number of combinations is denoted by C_n^r

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

Example 2.2 *We select* 3 *students from a group of* 20 *students . How many different cases are possible*

Solution:

There are

$$C_{20}^{3} = \frac{20!}{(20-3)!3!} = \frac{20 \times 19 \times 18 \times 17!}{17! \times 3!} = 1140,$$

then 1140 possible cases.

Definitions of some probability terms

Experiment: any process which generates well defined outcome.

Example 2.3 *The results obtained when rolling a regular dice once are :*$1, 2, 3, 4, 5, 6$

Event: It is a subset of sample space.

Example 2.4 *Considering the above experiment let A be the event of odd numbers, B be the event of even numbers, and C be the event of number 8.*

A=$\{1, 3, 5\}$

B=$\{2, 4, 6\}$

C =$\varnothing = \{\}$ *impossible event*

Complement of an Event: the complement of an event A means non occurrence of A and is denoted by $\overline{A}$,or C$_A$

Mutually Exclusive Events: Two events which cannot happen at the same time.

Independent Events: Two events are considered independent when the occurrence of one has no impact on the probability of the other happening.

Example 2.5 *The sample space for the following experiment*

a) Toss a dice one time.

b) Toss a coin two times.

Solution

a) $S = \{1, 2, 3, 4, 5, 6\}$

b) $S = \{(HH), (HT), (TH), (TT)\}$ where T = tail and H = head.

the probability that event A occur denoted $P(A)$ is defined as:

$$P(A) = \frac{the\ number\ of\ times\ the\ event\ to\ A}{Total\ number}$$

Example 2.6 *What is the probability of getting a fair dice is tossed once.*

a) A multiple of 3?

b) An even number?

c) Number 5?

d) Number 9?

Solution:

The sample space, $S = \{1, 2, 3, 4, 5, 6\}$

a) A the event of multiple of 3

$A = \{3, 6\} \Rightarrow P(C) = \frac{2}{6}$

b) B the event of **odd** numbers

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

c) C the event of number 5

$C = \{5\} \Rightarrow P(A) = \frac{1}{6}$

d) D the event of number 9

$D = \varnothing \Rightarrow P(D) = 0$

Exercise 1 *A box of* 24 *candles consists of* 10 *defective and* 14 *non defective candles. If* 6 *of this candles are selected at random, what is the probability*

a) All defective.

b) 4 will be non defective

c) All non defective

Solution 2.1 *Total selection is* C_{24}^{6}

$$C_{24}^{6} = \frac{24!}{(24-10)! \times 10!}$$

a) A "defective".

the number of times the event to A is C_{10}^{6}

$P(A) = \frac{C_{10}^{6}}{C_{24}^{6}}$

b) B "4 will be non defective".

the number of times the event to B *is* $C_{10}^{4}\times C_{14}^{2}$
$P\left(B\right)=\frac{C_{10}^{4}\times C_{14}^{2}}{C_{24}^{6}}$
c) C *"non defective".*
the number of times the event to C *is* $C_{10}^{0}\times C_{14}^{6}$
$P\left(C\right)=\frac{C_{10}^{0}\times C_{14}^{6}}{C_{24}^{6}}$

Proposition 2.1 *1.* $0\leq P(A)\leq 1$
2. $P(S)=1, S$ *is the sure event.*
3. $P(\varnothing)=0$*,* $\varnothing$ *is the impossible event.*
4. $P(\overline{A})=1-P(A)$
5. In general $P(A\cup B)=P(A)+P(B)-P\left(A\cap B\right)$
6. If A *and* B *are **mutually exclusive events***
$P(A\cup B)=P(A)+P(B)$ *because* $\left(P\left(A\cap B\right)=0.\right.$

Exercise 2 *Two balls are selected at random from a bag with four white balls and three black balls.*
- *What is the probability that both balls are white?*
- *What is the probability that both balls are black?*
- *What is the probability that one is white and one is black?*
- *What is the probability that at least one of them is white?*

Solution 2.2 *Total selection is* C_{7}^{2}

$$\begin{aligned} C_{7}^{2} &= \frac{7!}{(7-2)!\times 2!} \\ &= \frac{7\times 6\times 5!}{5!\times 2!} \\ &= 21 \end{aligned}$$

Let A *be the event that both balls are white.*
The number of times the event occurs to A *is* $C_{4}^{2}\times C_{3}^{0}$

$$\begin{aligned} C_{4}^{2}\times C_{3}^{0} &= \frac{4!}{2!2!}\times 1 \\ &= 6 \end{aligned}$$

$P(A)=\frac{C_{4}^{2}\times C_{3}^{0}}{C_{7}^{2}}=\frac{6}{21}$
Let B *be the event that both balls are black*
the number of times the event occurs to B *is* $C_{4}^{0}\times C_{3}^{2}$
$P\left(B\right)=\frac{C_{4}^{0}\times C_{3}^{2}}{C_{7}^{2}}$
Let E *be the event that one is white and one is black.*

the number of times the event occurs to E *is* $C_4^1 \times C_3^1$
$P(E) = \frac{C_4^1 \times C_3^1}{C_7^2}$
Let F *be the event that at least one of them is white*
the number of times the event occurs to F *is* $C_4^1 \times C_3^1 + C_4^2 \times C_3^0$
$P(F) = \frac{C_4^1 \times C_3^1 + C_4^2 \times C_3^0}{C_7^2}$.

2.2 Introduction to probability

2.3 Random varibales

Definition 2.1 *A random variable is a numerical description of the outcomes of the experiment*

A random variable may be...
◊ **Discrete** if it takes only a countable number of values. For example, number of dots on two dice, number of heads in three coin tossing, number of defective items,
number of boys in three births and so on.
◊ **Continuous** if can take on any value in an interval of numbers (i.e. its possible values are unaccountably infinite). For example, measured data on heights, weights,
temperature, and time and so on.

2.3.1 Discrete probability distribution

Example 2.7 *Flip a coin three times, let* X *be the number of heads in three tosses.*
$S = \{(HHH), (HHT), (HTH), (THH), (HTT), (THT), (TTH), (TTT)\}$
$X = \{0, 1, 2, 3\}$

The probability distribution

Example 2.8

$X = x$	0	1	2	3
$P(X = x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Mean and Variance of a random variable

Example 2.9 *The expected value of* X *is its mean*
Mean : $E(X) = \sum x_i P(X = x_i)$
The variance of X *is given by:*
Variance of $X = Var(X) = E(X^2) - [E(X)]^2$

where $E(X^2) = \sum x_i^2 P\left(X = x_i\right)$
The expected value $E(X) = \left(0 \times \frac{1}{8}\right) + \left(1 \times \frac{3}{8}\right) + \left(2 \times \frac{3}{8}\right) + \left(3 \times \frac{1}{8}\right) = 1.5$
$E(X^2) = \left(0^2 \times \frac{1}{8}\right) + \left(1^2 \times \frac{3}{8}\right) + \left(2^2 \times \frac{3}{8}\right) + \left(3^2 \times \frac{1}{8}\right) = 3$
The Variance $Var(X) = E(X^2) - [E(X)]^2$
$= 3^2 - (1.5)^2$

2.4 Conditioning and independence

2.4.1 Conditional probability of an event

The conditional probability of an event A given that B has already occurred, denoted $P(A/B)$ is

$$P(A/B) = \frac{P\left(A \cap B\right)}{P\left(B\right)}, \qquad P\left(B\right) \neq 0$$

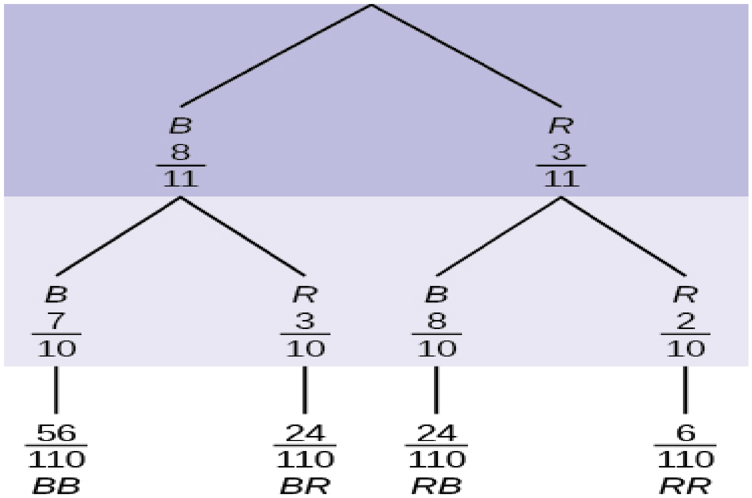
Remark 2.1 $P(\overline{A}/B) = 1 - P(A/B)$

Example 2.10 *Suppose a coin is tossed two times. The sample space is*
$S = \{HH, HT, TH, TT\}$ $(H : head,\ T : tail)$
Let A *be the event "two Heads",* $A = \{HH\}$
Let B *be the event "the first toss gives Heads",* $B = \{HH, HT\} \Rightarrow P\left(B\right) = \frac{2}{4}$
Then $A \cap B = \{HH\} \Rightarrow P\left(A \cap B\right) = \frac{1}{4}$
We have $P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{4}}{\frac{2}{4}} = \frac{1}{2}$

2.4.2 Tree Diagrams

Example 2.11 *In an urn, there are 11 balls. Three balls are red (R) and eight balls are blue (B). Selected two balls, one at a time, without replacement.*

The tree diagram using frequencies that show all the possible outcomes follows.



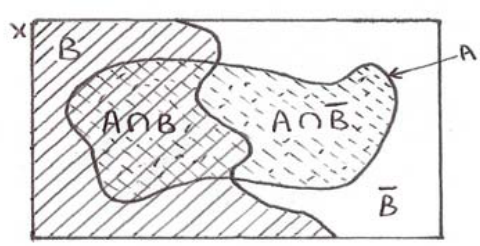
Partition of a Sample Space

A set of events $E_1, E_2, ..., E_n$ is said to represent a partition of a sample space S if
(a)$E_i \cap E_j = \emptyset, i \neq j; i, j = 1, 2, 3, ..., n$
(b) $E_1 \cup E_2 \cup ... \cup E_n = S, and$
(c) Each $E_i \neq \emptyset$, i. e, $P(E_i) > 0$ for all $i = 1, 2, ..., n$

2.4.3 The Law of Total Probability

We have $A = (A \cap B) \cup (A \cap \overline{B})$

We conclude that $P(A) = P(A \cap B) + P(A \cap \overline{B})$
or $P(A) = P(A/B) \times P(B) + P(A/\overline{B}) \times P(\overline{B})$



Theorem 2.1 *Let* $\{E_1, E_2, ..., E_n\}$ *be a partition of the sample space* S*. Let* A *be any event associated with* S*, then*

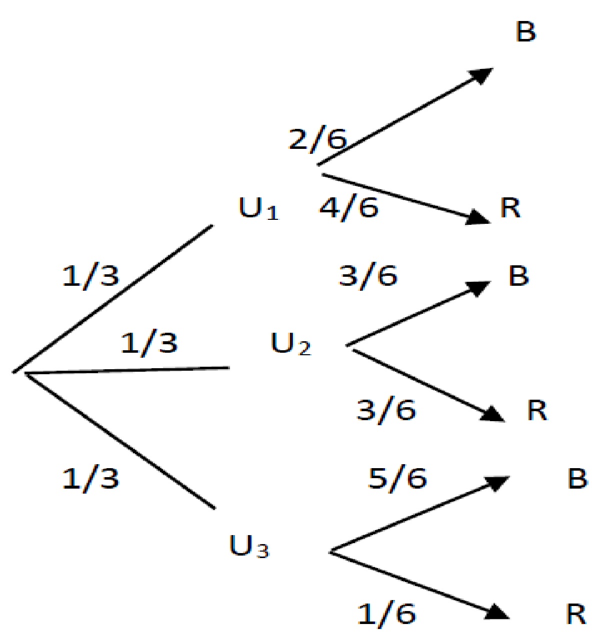
$$P(A) = \sum_{i=1}^{n} P(E_i) P(A/E_i)$$

Theorem 2.2 *If $E_1, E_2, \ldots, E_n$ are mutually exclusive and exhaustive events associated with a sample space, and A is any event of non zero probability, then*

$$P(E_i/A) = \frac{P(E_i) P(A/E_i)}{\sum_{i=1}^{n} P(E_i) P(A/E_i)}$$

Example 2.12 *Three bags, U_1, U_2 and U_3 each containing 6 balls.*
Bag U_1 contains 2 white balls and 4 red balls
Bag U_2 contains 3 white balls and 3 red balls
Bag U_3 contains 5 white balls and 1 red ball
We randomly choose a bag and then draw a ball from it
- Draw a tree diagram
- Calculate the probability of drawing a white ball from bag U_3
- Calculate the probability of drawing a white ball
- If that the drawn ball is white. What is the probability that it is from bag U_3

Solution:



the probability of drawing a white ball from bag U_3

$$P\left(B \cap U_3\right)=\frac{1}{3} \frac{5}{6}=\frac{5}{18}$$

the probability of drawing a white ball

$$\begin{aligned} P(B) &= P\left(B \cap U_1\right)+P\left(B \cap U_2\right)+P\left(B \cap U_3\right) \\ &= P(U_1)P(B/U_1)+P(U_2)P(B/U_2)+P(U_3)P(B/U_3) \\ &= \frac{1}{3}\frac{2}{6}+\frac{1}{3}\frac{3}{6}+\frac{1}{3}\frac{5}{6}=\frac{10}{18} \end{aligned}$$

$$\begin{aligned} P\left(U_3/B\right) &= \frac{P\left(B \cap U_3\right)}{P(B)} \\ &= \frac{5/18}{10/18} \\ &= \frac{5}{10} \end{aligned}$$

Example 2.13 *A store offers discounts of a portion of its merchandise include 3 types: A, B, and C*

"A" represents $\frac{1}{4}$ of its merchandise

"B" represents $\frac{1}{3}$of its merchandise

"C" represents the rest

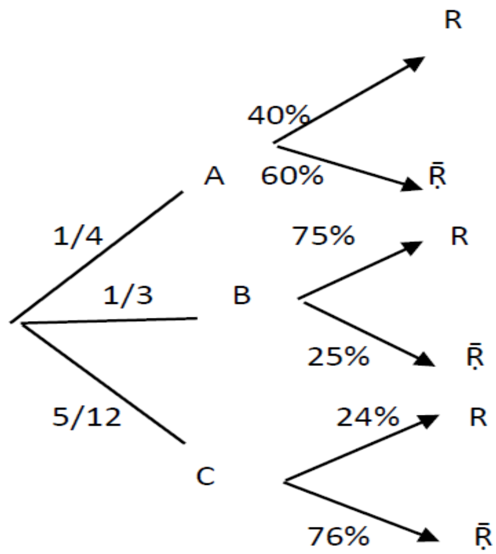
40% of type "A", 75% of type "B", and 24% of type "C" all with low prices.

The customer took a piece randomly,

- What is the probability that the piece is low

- What is the probability that the piece is from type B, knowing that it is low

Solution:



1) $P(R) = P(R \cap A) + P(R \cap B) + P(R \cap C)$
$= 0.4 \times \frac{1}{4} + 0.75 \times \frac{1}{3} + 0.24 \times \frac{5}{12}$
$= 0.45$

2) $P(B/R) = \frac{P(R \cap B)}{P(R)}$
$= \frac{0.75 \times \frac{1}{3}}{0.45} = 0.55$

2.4.4 Independent Events

Two events A and B are independent if

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

Here $P\left(A/B\right) = P\left(A\right), P\left(B/A\right) = P\left(B\right)$

Proposition 2.2 *If A and B are independent, then*
1) $\overline{A}$ and B are independent
2) A and $\overline{B}$ are independent
3) $\overline{A}$ and $\overline{B}$ are independent

Proof. 1) We have $A = (A \cap B) \cup \left(A \cap \overline{B}\right)$
We conclude that

$P\left(A \cap \overline{B}\right) = P(A) - P(A \cap B)$
$P\left(A \cap \overline{B}\right) = P(A) - P(A) \times P(B)$, becouse A and B are independent,

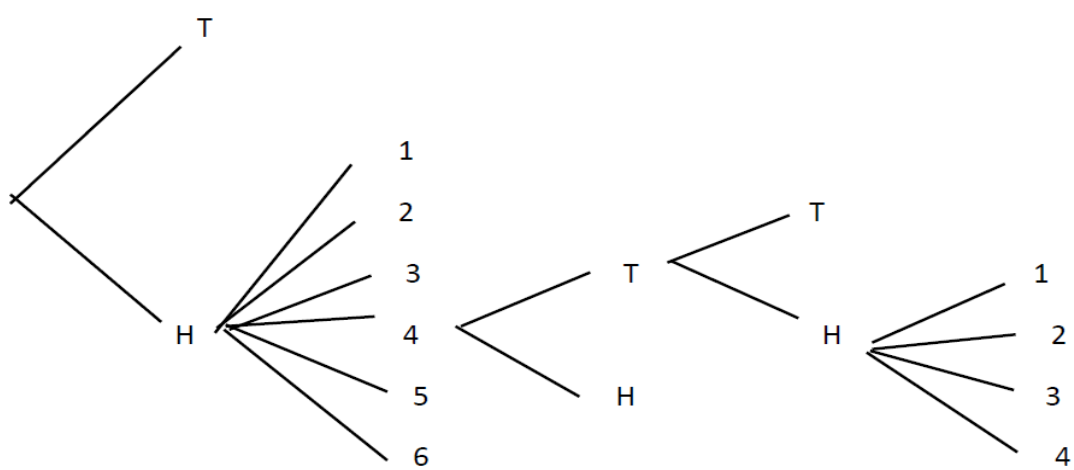
$$\begin{aligned} P\left(A \cap \overline{B}\right) &= P(A)\left[1 - P(B)\right] \\ &= P(A) \times P(\overline{B}) \end{aligned}$$

2) In the same way we prove A and $\overline{B}$ are independent.
3) We have

$$\begin{aligned} P\left(\overline{A} \cap \overline{B}\right) &= P\left(\overline{A \cup B}\right) \\ &= 1 - P\left(A \cup B\right) \\ &= 1 - \left[P(A) + P(B) - P\left(A \cap B\right)\right] \\ &= 1 - P(A) - P(B) + P\left(A\right) \times P(B) \\ &= \left(1 - P(A)\right)\left(1 - P(B)\right) \\ &= P\left(\overline{A}\right) \times P(\overline{B}) \end{aligned}$$

■

Example 2.14 *We toss a coin, then a dice numbered from 1 to 6, then a coin, then a coin, and finally a dice numbered from 1 to 4.*



Each result does not affect the other.
Probability of getting listed $(H4TH1)$ *is :* $\frac{1}{2} \times \frac{1}{6} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{4} = \frac{1}{192}$.

Example 2.15 *In one of the departments of the Faculty of Science and Technology, the students, who number* 895*, are distributed as follows*

class	*L_1*	*L_2*	*L_3*	*Total*
Resident	*50*		*85*	*195*
Non-resident	*285*	*220*		
Total			*280*	

1) Complete the table

2) A student was selected at random, R" The student is a resident"

L_1*"First-level student",* L_2*"Second-level student",* L_3*"Third-level student"*

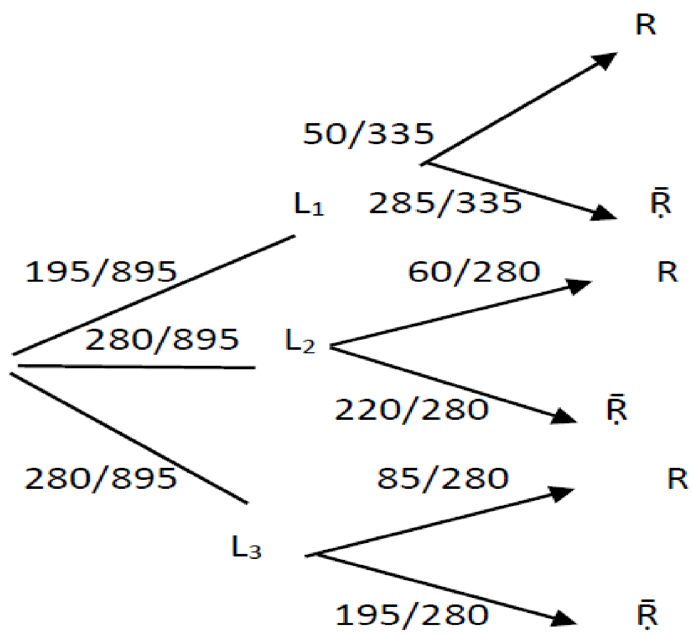
- Calculate the probabilities: $P(R \cap L_1), P(\overline{R} \cap L_2)$

- Are events R *and* L_3 *independent?*

3) Calculate the probabilities: $P(\overline{R}/L_1), P(L_3/R)$

Solution:

class	L_1	L_2	L_3	Total
Resident	50	60	85	195
Non-resident	285	220	195	700
Total	335	280	280	895



1)

$$\begin{aligned} P(R \cap L_1) &= \frac{195}{895} \times \frac{50}{195} \\ &\frac{10}{179} \end{aligned}$$

$$P\left(\overline{R} \cap L_2\right) = \frac{700}{895} \times \frac{220}{700}$$

$$P(R \cap L_3) = \frac{195}{895} \times \frac{85}{195}$$

2)

$$\begin{aligned} P(R) \times P(L_3) &= \frac{280}{895} \times \frac{195}{895} \\ &\neq P(R \cap L_3) \end{aligned}$$

R and L_3 are not independent.

3)

$$\begin{aligned} P\left(\overline{R}/L_1\right) &= \frac{P\left(\overline{R} \cap L_1\right)}{P(L_1)} \\ &= \frac{\frac{700}{895} \times \frac{285}{700}}{\frac{335}{895}} \\ &= \frac{285}{335} \end{aligned}$$

$$\begin{aligned} P(L_3/R) &= \frac{P(L_3 \cap R)}{P(R)} \\ &= \frac{\frac{195}{895} \times \frac{85}{195}}{\frac{195}{895}} \\ &= \frac{85}{195} \end{aligned}$$

2.5 Usual discrete and continuous probability laws

2.5.1 Binomial Distribution ($Law\ of\ Bernoulli$)

A binomial experiment is a probability experiment that satisfies the each trial has only one of the two possible mutually exclusive
outcomes, success (S) or a failure $\left(\overline{S}\right)$ with probabilities P and $(1-P)$ respectively

Law of Binomial Distrubution

Probability of getting x success on n trials:

$$P(x = k) = C_n^k P^k (1 - P)^{n-k}$$

And this is sometimes written as:

$$X\tilde{}Bin(n, p)$$

Example 2.16 *Flip a regular coin 4 times. What is the probability of getting 3 heads ?*

Solution: Let X be the number of heads in tossing a fair coin four times
$X\tilde{}Bin(n = 4, p = 0.50)$

$$\begin{aligned} P(x &= 3) = C_4^3 (0.5)^3 (1 - 0.5)^{4-3} \\ &= 4 \times (0.5)^4 \\ &= 0.25 \end{aligned}$$

Proposition 2.3 *If X is a binomial random variable with parameters n and p then* $E(X) = np,\ Var(X) = np(1 - p)$

Example 2.17 *In the Olympics, a runner crosses 5 obstacles in a row.*
The probability of passing a obstacle is 0.4.
- What is the probability that the rider passes only two obstacles?
- What is the probability that the rider passes at least three obstacles?

Solution:
1)

$$\begin{aligned} P(x &= 2) = C_5^2 (0.4)^2 (1 - 0.4)^{5-2} \\ &= 10 \times (0.4)^2 (0.6)^3 \end{aligned}$$

2)

$$\begin{aligned} P(x &\geq 3) = P(x = 3) + P(x = 4) + P(x = 5) \\ &= C_5^3 (0.4)^3 (1 - 0.4)^{5-3} + C_5^4 (0.4)^4 (1 - 0.4)^{5-4} + C_5^5 (0.4)^5 (1 - 0.4)^{5-5} \\ &= \left[10 \times (0.4)^3 (0.6)^2\right] + \left[5 \times (0.4)^4 (0.6)^1\right] + \left[1 \times (0.4)^5 (0.6)^0\right] \end{aligned}$$

2.5.2 Continuous probability distribution

2.5.3 Probability density function

The probabilities associated with a continuous random variable X are determined by the probability density function of the random variable. The function, denoted by $f(x)$, has the following properties:
1. $f(x) = 0$ for all x
2. The probability that X will be between two numbers a and b is equal to the area under $f(x)$ between a and b.

$$P(a < x < b) = \int_a^b f(x)dx$$

3.The total area under the entire curve of $f(x)$ is equal to 1.

$$\begin{aligned} P(-\infty < x < +\infty) &= \int_{-\infty}^{+\infty} f(x)dx \\ &= 1 \end{aligned}$$

4. Probability of a fixed value of a continuous random variable is zero.
$P(a \leq X \leq b) = P(a < X \leq b) = P(a < X < b) = P(a \leq X < b)$
*The cumulative distribution function of a continuous random variable:
$F(x) = P(X = x) =$ area under $f(x)$ between the smallest possible value of X (often $-\infty$) and point x

$$F(x) = \int_{-\infty}^{x} f(x)dx$$

*The expected value of a continuous random variable X, denoted by $E(X)$, and its variance, denoted by $Var(X)$, require the use of calculus for their computation. Thus

$$\begin{aligned} E(x) &= \int_{-\infty}^{+\infty} x.f(x)dx \\ Var(x) &= \int_{-\infty}^{-\infty} (x - E(x))^2 .f(x)dx \end{aligned}$$

Example 2.18 *A test is graded on the scale 0 to 1, with 0.55 needed to pass. Student scores are modeled by the following density:*

$$f(x) = \begin{cases} 4x \ \text{for} \ 0 \le x \le \frac{1}{2} \\ 4-4x \quad \text{for} \ \frac{1}{2} \le x \le 1 \\ 0 \quad \text{otherwise} \end{cases}$$

What is the probability that a random student passes the exam?

Solution: Let X = score of a random student.

Solution 2.3 $P(X \ge 0.55) = \int_{0.55}^{1} f(x)dx = \int_{0.55}^{1} (4-4x)dx = (4x - 2x^2)\mid_{0.55}^{1} = 0.405$

Exercise 3 *Suppose X is a random variable with cdf*

$$F(x) = \begin{cases} 0 & \text{for } x < 0 \\ x(2-x) & \text{for } 0 \le x \le 1 \\ 1 & \text{for } x > 1 \end{cases}$$

a) Find $E(X)$.
b)Find $P(X \le 0.4)$.

Solution 2.4 *a)* $f(x) = \acute{F}(x) = 2 - 2x$ *on* $[0, 1]$
therefore

$$\begin{aligned} E(x) &= \int_0^1 x f(x) dx \\ &= \int_0^1 x\,(2-2x)\,dx \\ &= x^2 - \frac{2}{3}x^3 \mid_0^1 \\ &= \frac{1}{3} \end{aligned}$$

b) $P(X \le 0.4) = F(0.4) = 0.4\,(2-0.4) = 0.64$

Exercise 4 *Compute the mean and variance of a random variable whose distribution is uniform on the interval* $[a, b]$.

Solution 2.5 *Let* $X \sim U(a, b)$. *The pdf of* X *is* $f(x) = \frac{1}{b-a}$ *on the interval [a, b]. Thus,*

$$
\begin{aligned}
E(x) &= \int_a^b x f(x) dx \\
&= \int_a^b x \frac{1}{b-a} dx \\
&= \frac{x^2}{2(b-a)} \Big|_a^b \\
&= \frac{b^2 - a^2}{2(b-a)} \\
&= \frac{b+a}{2}
\end{aligned}
$$

$$
\begin{aligned}
Var(x) &= \int_a^b (x - E(x))^2 f(x) dx \\
&= \int_a^b \left(x - \frac{b+a}{2}\right)^2 \frac{1}{b-a} dx \\
&= \frac{\left(x - \frac{b+a}{2}\right)^3}{3} \frac{1}{b-a} \Big|_a^b \\
&= \frac{1}{12} (b-a)^3 \frac{1}{b-a} \\
&= \frac{(b-a)^2}{12}
\end{aligned}
$$

Chapter 3

Series of exercises

3.1 Serie1

Exercise1: The glucose ratio for 30 people is as follows

$0.85, 0.87, 0.9, 0.93, 0.94, 0.94, 0.95, 0.97, 0.97, 0.98, 0.98, 0.99, 1.00, 1.01, 1.03,$
$1.03, 1.03, 1.03, 1.04, 1.08, 1.10, 1.10, 1.10, 1.11, 1.13, 1.14, 1.15, 1.17, 1.19, 1.20.$

1) Determine the range of the statistical series.

2) Draw a frequency table for the statistical series with glucose values (x_i) , Frequency (f_i) , Relative Frequency (R_{f_i}) and Percentage.

3) Draw a frequency table in the form classes of length 0.07 with the centers of the classes, Frequency, Relative Frequency, Growing C f and Decreasing C f.

Exercise2:

The following table represents the scores of 21 students in statistics

Score	7	9	11	14	15	17
Frequency	n_1	n_2	n_3	n_4	n_5	n_6

1) If you know that (n_i) is an arithmetic sequence with a base $r = 1$

- Calculate the frequencies $n_i, i = \overline{1, 6}$.

2) Draw bar charts.

Exercise3:

The following data represents the blood type of 30 students at a college.

A, B, O, AB, B, A, O, O, AB, B
B, B, A, O, O, AB, B, O, B, A

AB, A, O, A, A, B, O, A, A, B

1) Draw a frequency table for the data.

2) Plot the pie charts for the previous data.

Exercise4:

The following table shows the distribution of wages for 75 workers in a company.

$Wage(DA)$	$Frequency$	$RelativeFrequency$
$[24000 - 30000[$	6	0.08
$[30000 - 36000[$	9	.
$[36000 - 42000[$	.	0.24
$[42000 - 48000[$	24	0.32
$[48000 - 54000[$	15	.
$[54000 - 64000[$	.	.

Complete the table.

Serie2

Exercise1: Calculate the mean, median, and mode for the following series
series1: 12-14-16-7-8-15-12-12-14-8-12
series2: 42-36-36-57-36-24-18-42

Exercise2: We have the following data
5-4-7-9-12-11-5-12-6-7–6-5-7-10-6-8-10-4-8-12-9-11-16-14
1) Arrange and classify the previous values into table with four classes shown in it
classes -mid_point- frequency-cumulative frequency
2) Calculate the mean, median, and mode
3) Draw the frequency polygon and conclude the median graphically.

Exercise3: 1) Complete the following table

Classes	Relative frequency (R_{f_i})	Cumulative (R_{f_i})	Mid_point (m_i)	$m_i R_{f_i}$
$[10-20[$	0.18			
$[20-30[$	0.27			
$[30-40[$		0.54		
$[40-50[$		0.72		
$[50-60[$		1		

2) Calculate the arithmetic mean of the series.
3) Define the middle class and calculate the median
4) Draw the frequency polygon and conclude the median graphically.

Exercise4: We have the following distribution

Classes	$[7-9[$	$[9-11[$	$[11-13[$	$[13-15[$	$[15-17[$
frequency	10	70	30	70	50

1) Calculate the mean, median and mode

2) Calculate the range, variance, standard deviation and the coeffcient of variation

3) Calculate The coeffcient of Pearson. Determine the direction of Skewness for this distribution.

Exercise5: We have the following statistical series: 0-1-2-2-2-3-4

- Define the middle class and calculate the median.

- Is the distribution symmetrical ?

Exercise6: Below is the table of room rental for one of the hotels (In thousands)

Classes	$[5-8[$	$[8-11[$	$[11-14[$	$[14-17[$	$[17-20[$	$[20-23[$	$[23-26[$
frequency	9	11	15	35	45	35	30

1) Find the cumulative frequency and cumulative relative frequency

2) Draw a cumulative frequency polygon for the data ?

3) Calculate the mean, median and mode

4) Calculate the range, variance, standard deviation and the coeffcient of variation

5) Calculate The coeffcient of Pearson.

Determine the direction of Skewness for this distribution.

Serie3

Exercise1: A categorical variable has three categories (A, B, C) with the following frequencies of occurrence

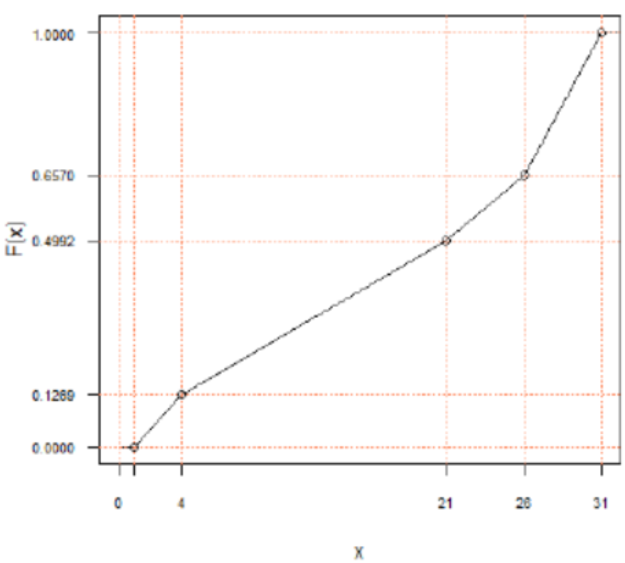
categories	frequency	relative frequency	percentage %
A	11	0.22	
B	30		
C			

a) Complete the table

b) Construct a pie chart.

c) Construct a bar chart.

Exercise2:134 firms of an industrial district have been classified according to the graphic below



a) Draw a frequency table with classes and cumulative frequency

b) Compute the mean, median and median graphically

c) Compute the range, variance, standard deviation and the coeffcient of variation.

Exercise3: Complete the following table representing an interrogation 400 person

class	frequency(f_i)	mid_point(m_i)	cumulative frequency	%	$f_i m_i$
$[14-16[$				5	
		17			1088
$[\ldots-22[$					
	84		280		
$[24-\ldots[$	40				
$[26-30[$		28			

Exercise4: The following table shows rainfall (in mm) in eastern Algeria between 2014 and 2019

X_i	2014	2015	2016	2017	2018	2019
Y_i	50	120	130	100	95	145

a) Find the middle point $G\left(\overline{X},\overline{Y}\right)$ and then draw a scatter plot

b) Fit a linear regression

c) What is the rainfall in 2024 and 2026?

Exercise5: We have the following distribution

X_i	8	14	22	24	30	34
Y_i	4.4	5.2	4.3	3.2	3.3	2.8

a) Compute the middle point $G\left(\overline{X},\overline{Y}\right)$

b) Find the equation of linear regression $(\Delta) : y = ax + b$

c) Find the equation of Straight Mayer (G_1G_2)

d) According to the principle of least squares, what is the closest line for the scatter points?

Serie4

Exercise1: **a)** Calculate: C_9^3, A_{11}^7

b) Define the value of the natural number $n : A_n^2 = 72$

c) Resolve in $\mathbb{N}^2$ the system $\begin{cases} C_{x+1}^{y} = C_x^{y-1} \\ C_{x+y}^{2} = 10 \end{cases}$

Exercise2: A box containing 10 balls (there is no difference between them when touched). We draw 6 balls randomly

- How many cases are possible?

a) If the draw is made at one time

b) If the draw by (Arrangement without repetition)

b) If the draw by (Arrangement and repetition).

Exercise3: Find expressions of the mean and variance of x for the mean and variance of $y = ax$

Exercise4: A box of 24 pens consists of 10 defective and 14 non defective pens. If 6 of this pens are selected at random, what is the probability

a) A"All will be defective".

b) B"3 will be non defective"

c) C" **at least** four of them is defective"

d) D"**at most** two of them is defective"

Exercise5: 4 balls are to be drawn without replacement from a box containing 8 red and 4 white balls. If X denotes the number of red ball drawn, find the probability distribution of X.

Exercise6: Determine variance and standard deviation of the number of heads in three tosses of a coin.

Exercise7: A committee of 4 people is to be formed from a group of 7 men and 6 women.
1) How many different committees are possible
2) what is the probability of events A" Among the members there is only one man"
B"Among the members there are at most two women"
3) Let X be the random variable to each choice the number of men
- Find the probability distribution of X.
- Determine mean, variance and standard deviation of X

Exercise8: Two dice are rolled.
$A =$ 'sum of two dice equals 3'
$B =$ 'sum of two dice equals 7'
$C =$ 'at least one of the dice shows a 1'
(a) What is $P(A)$, $P(B)$, $P(C), P(A \cap C)$ and $P(B \cap C)$?
(b) What is $P(B|C)$? What is $P(A|C)$, $P(B|C)$?
(c) Are A and C independent?
What about B and C?

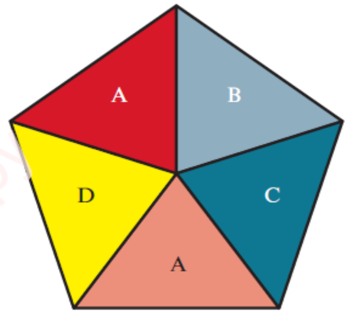
Serie5

Exercise1: A bag contains cards numbered from 1 to 14. One card is drawn at random.
Find the probability of
a) selecting a prime number or a multiple of four.
b) selecting a multiple of two or a multiple of three.

Exercise2: A man forgets his banker's card 10% of the time, he forgets his cheque book 5% of the time and he forgets both 2% of the time.
(a) What is the probability that, on any one day, he will have both his banker's card and his cheque book?
(b) What is the probability that he will find his banker's card in his pockets given that he has already found his cheque book?

Exercise3: In a competition, the student answers the numbers of questions. True (1) and False (0)
We consider two events A"The answers do not have the same symbol", B"At most one answer marked 0"
- If the number of questions is two. Are A and B independent?
- If the number of questions is three. Are A and B independent?

Exercise4: A regular pentagonal spinner is shown. Find the probability that 10 spins produce exactly three As.



Exercise5: The random variable $X \sim B(n, p)$. Given that $E(X) = 12$ and $Var(X) = 7.5$, find:
a) the value of n and of p
b) $P(X = 11)$.

Exercise6: In a particular country, 85% of the population has rhesus-positive (R+) blood.
- Find the probability that fewer than 39 people in a random sample of 40 have rhesus-positive blood.

- Calculate $E(X)$ and $Var(X)$.

Solution of serie1

Exercise1

1) The range of the statistical series=$X_{\max} - X_{\min} = 1.2 - 0.85 = 0.35$

2) The frequency table

values (x_i)	0.85	0.87	0.9	0.93	0.94	0.95	0.97	0.98	0.99	1.00
Frequency (f_i)	1	1	1	1	2	1	2	2	1	1
Relative Frequency (R_{f_i})	0.03	0.03	0.03	0.03	0.06	0.03	0.06	0.06	0.03	0.03
Percentage.$(\%)$	3	3	3	3	6	3	6	6	3	3

1.01	1.03	1.04	1.08	1.10	1.11	1.13	1.14	1.15	1.17	1.19	1.20	Total
1	4	1	1	3	1	1	1	1	1	1	1	30
0.03	0.12	0.03	0.03	0.09	0.03	0.03	0.03	0.03	0.03	0.03	0.03	1
3	12	3	3	9	3	3	3	3	3	3	3	100%

3) Draw a frequency table of the classes

Classes (C_i)	$[0.85, 0.92[$	$[0.92, 0.99[$	$[0.99, 1.06[$	$[1.06, 1.13[$	$[1.13, 1.20[$	-
Centres(m_i)	0.885	0.995	1.025	1.095	1.165	-
Frequency (R_{f_i})	3	8	8	6	5	30
Relative Frequency (R_{f_i})	0.1	0.366	0.366	0.2	0.166	1
Growing C f	3	11	19	25	30	-
Decreasing C f	30	27	19	11	5	-

Exercise2:

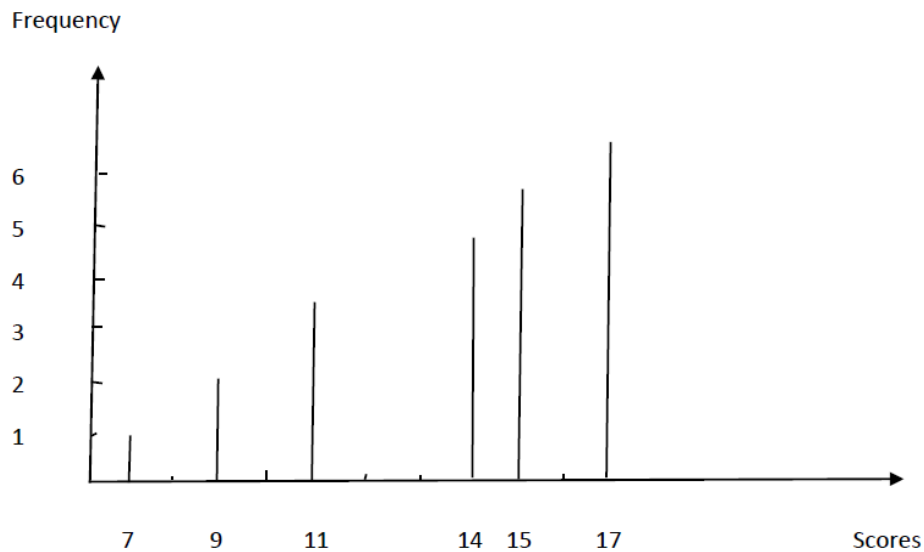
(n_i) is an arithmetic sequence with a base $r = 1$ then

$n_6 = n_1 + 5r \Rightarrow n_6 = n_1 + 5...(1)$

We have $\sum_{i=1}^{6} n_i = 21 \Rightarrow \frac{6(n_1+n_6)}{2} = 21$,
then $n_1 + n_6 = 7 ... (2)$

$$(1), (2) \Rightarrow \begin{cases} n_6 = n_1 + 5 \\ n_1 + n_6 = 7 \end{cases}$$
$$\Rightarrow n_1 = 1 \text{ and } n_6 = 6,$$

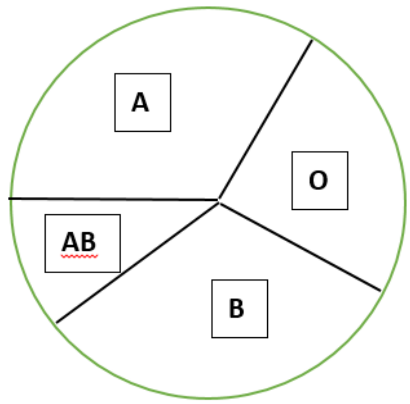
we obtain $n_2 = 2, n_3 = 3, n_4 = 4, n_5 = 5$



Bar charts

Exercise3

The blood	$Frequency$	$RelativeFrequency$
A	9	$9/30$
B	9	$9/30$
AB	4	$4/30$
O	8	$8/30$
Total	30	1.000



Pie charts

Exercise4

$Wage(DA)$	$Frequency$	$RelativeFrequency$
$[24000 - 30000[$	6	0.08
$[30000 - 36000[$	9	0.12
$[36000 - 42000[$	18	0.24
$[42000 - 48000[$	24	0.32
$[48000 - 54000[$	15	0.2
$[54000 - 64000[$	3	0.04
Total	75	1.000

24/Total = 0.32

cell/75 = 0.24

Total = 24/0.32 = 75

cell = 0.24×75 = 18

Solution of serie2

Exercise1

Serie1: we arrange the values.

7,8,8,12,12,**12**,12,14,14,15,16

The mean $\bar{X} = \frac{\sum n_i x_i}{N} = \frac{7+8\times2+12\times4+14\times2+15+16}{1+2+4+2+1+1}$

The median $Me = X_{\frac{11+1}{2}} = X_6 = 12$ (the number of values is odd)

The mode $Mo = 12$ (the most frequent value)

Serie2 : we arrange the values.

18,24,36,36,36,42,42,57

The mean $\bar{X} = \frac{\sum n_i x_i}{N} = \frac{18+24+36\times3+42\times2+57}{1+1+3+2+1}$

The median $Me = \frac{X_{\frac{8}{2}}+X_{\frac{8}{2}+1}}{2} = \frac{X_4+X_5}{2} = \frac{36+36}{2} = 36$ (the number of values is even)

The mode $Mo = 36$

Exercise2

Number of classes $= \frac{\text{range}}{\text{class midth}} \Rightarrow 4 = \frac{16-4}{E}$

$E = \frac{16-4}{4} = 3$

class	$[4-7[$	$[7-10[$	$[10-13[$	$[13-16[$
Frequency	8	9	5	2
Cumulative F2	8	17	22	24
Mid point (mi)	5.5	8.5	11.5	14.5

$\bar{X} = \frac{\sum n_i f_i}{N} = \frac{(5.5\times8) + (8.5\times9) + (11.5\times5) + (14.5\times2)}{24}$

$= 8.025$

$\frac{N}{2} = \frac{24}{2} = 12$, It means median lies in the class-interval $[7-10[$

$M_e = l_1 + \frac{(l_2-l_1)}{f}\left(\frac{N}{2}-c\right)$, l_1 : the lower limit of the class in which the median lies

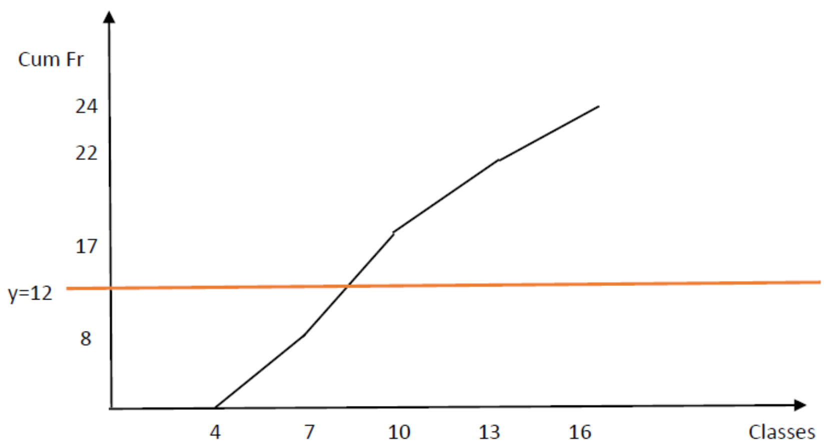
l_2 : the upper limit of the class in which the median lies

$l_1 = 7, l_2 = 10, c = 8,$

$M_e = 7 + \frac{3}{7}(12-8) \simeq 8.3...$

The mode $Mo \in [7-10[\,, Mo = \frac{7+10}{2} = 8.5$

The x coordinute at the intersection of the ascinding curve and the right $\frac{N}{2}$ is the medion



Exercise3 :

class	R Frequency (fi)	Cumulative R$_{fi}$	Mid point (mi)	mi R$_{fi}$
$[10-20[$	0.18	0.18	15	2.7
$[20-30[$	0.27	0.18+0.27=0.45	25	6.75
$[30-40[$	0.54-0.45=0.09	0.54	35	3.15
$[40-50[$	0.72-0.54=0.18	0.72	45	8.1
$[50-60[$	1-0.72=0.28	1	55	15.3

Arethmetic mean : $\bar{X} = \sum miR_{fi} = 2.7 + 6.75 + ... + 15.3$

=36.2

Middle class . $[30-40[$

Median :M$_e = l_1 + \frac{l_2-l_1}{f}\left(\frac{1}{2}-c\right)$

$l_1 = 30$ $\qquad c = 0.45$

$f = 0.09$ $\qquad l_2 = 40$

M$_e = 30 + \frac{40-30}{0.09}(0.5-0.45) \simeq 35.55$

Exercise4 :

class	$[7-9[$	$[9-11[$	$[11-13[$	$[13-15[$	$[15-17[$	
Frequency fi	10	70	30	40	15	$\sum f_i = 200$
Mid point (mi)	8	10	12	14	16	
CF↗	10	80	110	150	200	

Mean $\quad \bar{X} = \frac{(10\times8)+(70\times10)+(30\times12)+(40\times14)+(50\times16)}{10+70+30+40+50}$

=12.5

Median :M$_e \in [11-13[$

$$M_e = l_1 + \frac{l_2 - l_1}{f_i}\left(\frac{N}{2} - c\right)$$
$$= 11 + \frac{2}{30}(100-80) \simeq 12.33$$

Node : Mo $\in [9-11[\implies Mo = \frac{9+11}{2} = 10$

ronge = 17-7 = 10 $(X_{\max} - X_{\min})$

Varience

Var = $\frac{1}{N}\left(\sum f_i e i^2\right) - \bar{X}^2$

$=\frac{1}{200}\left[10\,(8)^2 + 20\,(10)^2 + 30\,(12)^2 + 40\,(14)^2 + 50\,(16)^2\right] - (12.5)^2$

$\simeq 6.75$

Standard deviation

$\sigma = \sqrt[2]{var} \implies \sigma = \sqrt{6.75} = 2.59$

The coefficieut of variation (CV)

CV = $\frac{\sigma}{\bar{X}} \times 100\% = 20.78\%$

The coefficieut of pearson

S = $\frac{\bar{X} - Mo}{\sigma} = \frac{12.5-10}{2.59} \rangle\ 0$

The distubution is negatively skewed

Exercise5 :

Serie : 0 , 1 , 2 , 2 , 2 , 3 , 4

$\bar{X} = 2$

M$_e = 2$

Mo =2

$\bar{X} = M_e = Mo$,the distubution symmetrical

Solution of serie3

Exercise 01 :

we have $Rf_i = \frac{f_i}{\sum f_i} \implies 0.22 = \frac{11}{\sum f_i}$ $\sum f_i = N$

$\implies N = \frac{11}{0.22} = 50$

Category	Frequency	Relative Fres	%
A	11	0.22	22%
B	30	0.60	60%
C	9	0.18	18%
Total	50	1.00	100%

Exercise 2 :

Class	Relative Fr Rf$_i$	cumulative relative r	mi	mi^2
$[1-4[$	0.1269	0.1269	2.5	6.25
$[4-21[$	0.3723	0.4992	12.5	156.25
$[21-26[$	0.1578	0.657	23.5	552.25
$[26-31[$	0.3430	1	28.5	812.25

$\bar{X} = \sum m_i R_{fi} = (0.1269 \times 2.5) + (0.3727 \times 12.5) + (0.1578 \times 22.5) + (0.3430 \times 28.5) = 18.45$

Median : $M_e = l_1 + \frac{l_2 - l_1}{R_f}\left[\frac{1}{2} - C\right]$

Middle class : $\frac{1}{2} \in [21-26[$,$l_1 = 21$, $l_2 = 26$

$R_f = 0.1578$, C=0.4992

$M_e = 21 + \frac{5}{0.1578}(0.5 - 0.4992) = 21 + 0.0253 = 21.0.0253$

C) The range =31-1=30

Variance: $V_5 = \frac{1}{N}\sum(f_i m_{i^2})$-$\bar{X}\ ^2-$

$=\sum R_{fi} m_{i^2} - \bar{X}\ ^2$

$=(0.1269 \times 6.25 + 0.3723 \times 156.25 + ... + 0.3430 \times 812.25) - (18.45)^2$
=424.7118-$(18.45) \simeq 84.1999$

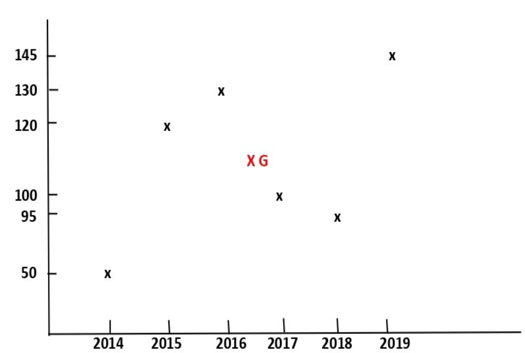
Standard deviation $\sigma = \sqrt{V_r} = \sqrt{84.125} \simeq 9, ...$
Coefficent of variation= CU=$\frac{6\times100\%}{\bar{X}}$

Exercice 03:

Class	Frequency (G_i)	mid point (m_i)	cumulative F_2	%	f_i m_i
$[14-16[$	20	15		5	300
$[16-18[$	64	17		16	1088
$[18-22[$	112	20	196	28	
$[22-24[$	84	23	280	21	
$[24-26[$	40	25	320	10	
$[26-30[$	80	28	400	20	

Exercice 04: a) $\bar{X} = \frac{2014+2015+.....+2019}{6} = 2016.5$
$\bar{Y} = \frac{50+120+130+500+95+45}{6} = 106.66$

The middle point $\sigma\ (2016.5.106.66)$



The scatter plot

b) the linear regression, $y = ax + b$

$$a = \frac{\frac{1}{N}\sum x_i y_i + \bar{x}\ \bar{y}}{\frac{1}{N}\sum x_i^2 + \bar{x}^2}$$

$a = \frac{\frac{1}{6}(100700+241800+262080+201700+191710+292755)-210079.89}{\frac{1}{6}(4056196+4060225+4064256+4068289+4072324+4075361)-4066772.25}$
$= \frac{44.27}{2.91} = 15.21$

$b = 106.66 - 15.21 \times 2016.5$

$b = -30564.305$

$y = 15.21x - 30564.305$

c) The rain fall in 2024: $y = 15.21 \times 2024 - 30564.305 = 220.735$

The rain fall in 2026: $y = 15.21 \times 2026 - 30564.305 = 251.155$

Exercice 05: a) $G\left(\bar{X}, \bar{Y}\right)$,
$\bar{X} = \frac{8+14+22+24+30+34}{6} = 22$
$\bar{Y} = \frac{4.4+5.2+4.3+3.2+3.3+2.8}{6} = 3.86$
We obtain $G\left(22, 3.86\right)$

b) The linear regression $(\Delta) : y = ax + b$

$a = -0.0779$, $b = 5.579$, then $y = +0.0779x + 5.579$

c) The linear of Mayer $(G_1 G_2)$

G_1 : the midlle point of M_1, M_2, M_3

$G_1\left(14.66, 4.63\right)$

G_2 : the midlle point of M_4, M_5, M_6

$G_2\left(29.33, 3.1\right)$,

we obtain $(G_1 G_2) : y = -0.104x + 6.154$.

The least squars : $S = \sum \left(y_i - ax_i - b\right)^2$

$S_1 = 1.284$ of (Δ), $S_2 = 1.604$ of $(G_1 G_2)$

$S_1 < S_2 \Rightarrow$the closest line for scatter points it is (Δ)

Solution of serie4

Exercice 01: 1)

$$\begin{aligned} C_9^3 &= \frac{9!}{(9-3)!3!} = 84, \\ A_{11}^7 &= \frac{11!}{(11-7)!} \end{aligned}$$

2)

$$\begin{aligned} A_n^2 &= 72 \Rightarrow \frac{n!}{(n-2)!} = 72 \\ &\Rightarrow \frac{n\,(n-1)\,(n-2)!}{(n-2)!} = 72 \\ &\Rightarrow n^2 - n - 72 = 0, \end{aligned}$$

$\Delta = 289, n_1 = -8$ or $n_2 = 9$

3)

$$\left\{ \begin{array}{c} C_{x+1}^{y} = C_{x}^{y-1} \\ C_{x+y}^{2} = 10 \end{array} \right. \Rightarrow \left\{ \begin{array}{c} \frac{(x+1)!}{(x+1-y)!} = \frac{x!}{(x-y+1)!(y-1)!} \\ \frac{(x+y)!}{(x+y-2)!2!} = 10 \end{array} \right.$$

$$\Rightarrow \left\{ \begin{array}{c} (x+1)!\,(y-1)! = x!y! \\ (x+y)\,(x+y-1) = 20 \end{array} \right.$$

$$\Rightarrow \left\{ \begin{array}{c} x+1 = y \\ (x+y)\,(x+y-1) = 20, \end{array} \right.$$

we find $2x\,(2x+1) = 20 \Rightarrow 2x^2 + x - 10 = 0, \Delta = 81,$

$x_1 = \frac{-10}{4}$ or $x_2 = \frac{-8}{4} = 2,$

We obtain $(x, y) = (2, 3)\,.$

Exercice 2: a) The draw is made at one time $(Combination)$

$$C_{10}^{6} = \frac{10!}{(10-6)!6!}$$

b) The draw by Arrangement without repetition $(Arrangement)$

$$A_{10}^{6} = \frac{10!}{(10-6)!}$$

c) The draw by Arrangement without repetition $(List)$
the number of lists is 10^6.

Exercice 3: We have $y = ax$

$$\begin{aligned} y &= \frac{\sum y_i}{N} \\ &= \frac{\sum ax_i}{N} \\ &= a\bar{X}, \end{aligned}$$

and

$$\begin{aligned} V_y &= \frac{\sum \left(y_i - \bar{Y}\right)^2}{N} \\ &= \frac{\sum \left(ax_i - a\bar{X}\right)^2}{N} \\ &= \frac{\sum a^2 \left(x_i - \bar{X}\right)^2}{N} \\ &= a^2 V_x \end{aligned}$$

Exercice 4: The are $C_{24}^{6} = \frac{24!}{(24-6)!6!} = 29260$
A"all will be defective"

$$P(A) = \frac{C_{10}^{6} \times C_{14}^{0}}{C_{24}^{6}}$$

B"3 will be non defective"

$$P(B) = \frac{C_{10}^{3} \times C_{14}^{3}}{C_{24}^{6}}$$

C" at least 4 of them is defective"

$$P(C) = \frac{C_{6}^{4} \times C_{14}^{2} + C_{6}^{5} \times C_{14}^{1} + C_{6}^{6} \times C_{14}^{0}}{C_{24}^{6}}$$

D" at meost 2 of them is defective"

$$P(D)=\frac{C_6^2\times C_{14}^4+C_6^1\times C_{14}^5+C_6^0\times C_{14}^6}{C_{24}^6}$$

Exercice 5: Sine 4 balls have to be drawn

$X=\{0,1,2,3,4\}$

$P(X=0)=\frac{C_8^0C_4^4}{C_{12}^4}=\frac{1}{495}$ ($X=0$ no red balls, 4white balls)

$P(X=1)=\frac{C_8^1C_4^3}{C_{12}^4}=\frac{32}{495}$

$P(X=2)=\frac{C_8^2C_4^2}{C_{12}^4}=\frac{168}{495}$ (2 red balls 2 white balls)

$P(X=3)=\frac{C_8^3C_4^1}{C_{12}^4}=\frac{224}{495}$

$P(X=4)=\frac{C_8^4C_4^0}{C_{12}^4}=\frac{70}{495}$.

The probability distrubution of X

X	0	1	2	3	4
$P(X)$	$\frac{1}{495}$	$\frac{32}{495}$	$\frac{168}{495}$	$\frac{224}{495}$	$\frac{70}{495}$

Exercice 6: Let X denote the number of heads tossed.

So, $X=\{0,1,2,3\}$

$S=\{HHH,HHT,HTH,THH,THT,TTH,HTT,TTT\}$

$X=0\,(TTT)\,,X=1\,(HTT,THT,TTH)\,,X=2\,(HHT,HTH,THH)\,,X=3(HHH)$

$P(X=0)=\frac{1}{8}$

$P(X=1)=\frac{3}{8}$

$P(X=2)=\frac{3}{8}$

$P(X=3)=\frac{1}{8}$

The probability distrubution of X

X	0	1	2	3
$P(X)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Varience of X, $V_X=\sum X_i^2P_i(X)-E^2(X)$,

where $E(X)$ is the mean of X given by

$E(X)=\sum X_iP_i(X)=0\times\frac{1}{8}+1\times\frac{3}{8}+2\times\frac{3}{8}+3\times\frac{1}{8}=\frac{3}{2}$

$V_X=\left[0^2\times\frac{1}{8}+1^2\times\frac{3}{8}+2^2\times\frac{3}{8}+3^2\times\frac{1}{8}\right]-\left(\frac{3}{2}\right)^2=\frac{3}{4}$.

The standard deviation is $\sigma=\sqrt{V_X}=\frac{\sqrt{3}}{2}$

Exercice 7: Total selections is

$C_{13}^4=\frac{13!}{9!4!}=715$

$P(A) = \frac{C_7^1 C_6^3}{C_{13}^4}, P(B) = \frac{C_6^1 C_7^3 + C_6^2 C_7^2 + C_7^4}{C_{13}^4}$
$X = \{0, 1, 2, 3, 4\}$
$P(X = 0) = \frac{C_6^4 C_7^0}{C_{13}^4} = \frac{15}{715} = 0.021$
$P(X = 1) = \frac{C_6^3 C_7^1}{C_{13}^4} = \frac{140}{715} = 0.2$
$P(X = 2) = \frac{C_6^2 C_7^2}{C_{13}^4} = \frac{315}{715} = 0.44$
$P(X = 3) = \frac{C_6^1 C_7^3}{C_{13}^4} = \frac{210}{715} = 0.29$
$P(X = 4) = \frac{C_6^0 C_7^4}{C_{13}^4} = \frac{35}{715} = 0.049$
The probability distribution is

X	0	1	2	3	4
$P(X)$	0.021	0.2	0.44	0.29	0.049

$E(X) = \sum X_i P_i(X) = 2.146$
$V(X) = \sum X_i^2 P_i(X) - E^2(X) = 0.754$
$\sigma = \sqrt{V(X)} = 0.86.$

Exercice 8: The sample space is:

$$\begin{aligned} \Omega &= \left\{(i, j)\ /\ i, j = \overline{1, 6}\right\} \\ &= \{(1, 1), (1, 2), \dots (6, 6)\} \end{aligned}$$

1/ $A = \{(1, 2), (2, 1)\}$
$B = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$
$C = \{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6), (2, 1), (3, 1), (4, 1), (5, 1), (6, 1)\}.$
2/ $P(A) = \frac{2}{36}, P(B) = \frac{6}{36}, P(C) = \frac{11}{36}$
$P(A \cap B) = 0, P(A \cap C) = \frac{2}{36}, P(B \cap C) = \frac{2}{36}$
3/ $P(A/C) = \frac{P(A \cap C)}{P(C)} = \frac{2/36}{11/36} = \frac{2}{11}$
$P(B/C) = \frac{P(B \cap C)}{P(C)} = \frac{2/36}{11/36} = \frac{2}{11}$
4/ $P(A) = \frac{2}{36} \neq P(A/C)$, so they are not independent .
Remark: two events A and B are independent if
$P(A \cap B) = P(A) \times P(B)$
or $P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \times P(B)}{P(B)} = P(A)$
$P(B) = \frac{6}{36} \neq P(B/C)$, so they are not independent .

Solution of serie5

Exercice 1: a) A“ Prime number”, $A = \{2, 3, 5, 7, 11, 13\}$.
B“multiple of 4”, $B = \{4, 8, 12\}$
$P(A \cup B) = \frac{6}{14} + \frac{3}{14} = \frac{9}{14}$
b) C“ multiple of 2”, $C = \{2, 4, 6, 8, 10, 12, 14\}$
D“ multiple of 3”, $D = \{3, 6, 9, 12\}$, $C \cap D = \{6, 12\}$
$P(C \cup D) = P(C) + P(D) - P(C \cap D)$
$P(C \cup D) = 7/14 + 4/14 - 2/14 = 9/14$

Exercice 2: The probability of forgetting the banker’s card is $P(A) = 0.1$
The probability of forgetting the cheque book is $P(B) = 0.05$
The probability of forgetting both book and card is $P(A \cap B) = 0.02$
- we know that

$$\begin{aligned} P\left(\bar{A} \cap \bar{B}\right) &= 1 - P\overline{\left(\bar{A} \cap \bar{B}\right)} \\ &= 1 - P(A \cup B) . \end{aligned}$$

Also: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
$= 0.1 + 0.05 - 0.02 = 0.13.$
Therofore $P\left(\bar{A} \cap \bar{B}\right) = 1 - 0.13 = 0.87$

$$\begin{aligned} P\left(\frac{\bar{A} \cap \bar{B}}{\bar{B}}\right) &= \frac{P\left(\bar{A} \cap \bar{B}\right)}{P\left(\bar{B}\right)} \\ &= \frac{0.87}{1 - 0.05} = 0.916 \end{aligned}$$

Exercice 3: So $X \backsim B\left(10, \frac{2}{5}\right)$

$$P\left(x = k\right) = C_n^k P^k \left(1 - P\right)^{n-k}$$

$$\begin{aligned} P\left(x = 3\right) &= C_{10}^3 \left(\frac{2}{5}\right)^3 \left(1 - \frac{2}{5}\right)^7 \\ &= C_{10}^3 \left(0.4\right)^3 \left(0.6\right)^7 \\ &\simeq 0.215 \end{aligned}$$

Exercice 4: 1/ The sample space S_1

$S_1 = \{(1,1), (1,0), (0,1), (0,0)\}$

$A = \{(1,0), (0,1)\} \Rightarrow P\left(A\right) = \frac{2}{4}$

$B = \{(1,1), (1,0), (0,1)\} \Rightarrow P\left(B\right) = \frac{3}{4}$

$A \cap B = \{(1,0), (0,1)\} \Rightarrow P\left(A \cap B\right) = \frac{2}{4}$

$P\left(A\right) \times P\left(B\right) = \frac{6}{16} \neq P\left(A \cap B\right)$

A and B are not independent.

2/ The sample space S_2

$S_2 = \{(1,1,1), (1,1,0), (1,0,1), (0,1,1), (1,0,0), (0,1,0), (0,0,1), (0,0,0)\}$

$\begin{cases} P\left(A\right) = \frac{6}{8} = \frac{3}{4}, \\ P\left(B\right) = \frac{4}{8} = \frac{1}{2} \end{cases} \Rightarrow P\left(A\right) \times P\left(B\right) = \frac{3}{4} \times \frac{1}{2} = \frac{3}{8}$

$A \cap B = \{(1,0,1), (0,1,1), (1,1,0)\} \Rightarrow P\left(A \cap B\right) = \frac{3}{8}$

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

A and B are independent.

Exercice 4: $X \backsim B\left(40, 0.85\right)$

$$\begin{aligned} P\left(x \leq 39\right) &= 1 - \left[P\left(x = 39\right) + P\left(x = 40\right)\right] \\ &= 1 - \left[C_{40}^{39} \left(0.85\right)^{39} \left(0.15\right)^1 + C_{40}^{40} \left(0.85\right)^{40} \left(0.15\right)^0\right] \\ &= 0.98 \end{aligned}$$

$E\left(X\right) = nP = 40\left(0.85\right)$

$V(X) = nP(1-P) = 40(0.85)(0.15)$

Exercice 5: $X \backsim B(n, P)$

$E(X) = 12, V(X) = 7.5,$

we use $\begin{cases} E(X) = nP \\ V(X) = nP(1-P) \end{cases}$

a)

$$\begin{aligned} V(X) &= E(X)(1-P) \Rightarrow 1 - P = \frac{V(X)}{E(X)} \\ &\Rightarrow P = 1 - \frac{7.5}{12} = 0.375 \end{aligned}$$

$n = \frac{E(X)}{P} = \frac{12}{0.375} = 32$

b) $P(x = 11) = C_{32}^{11}(0.375)^{11}(1 - 0.375)^{21}$

$= 0.138$

Sample exam

Exercise1 (6points): 1) If the mode of the following values: $6, 2, \frac{\omega}{3}, 4, 8$ equals 4. Find the value of ω

2) Let A and B be two events with $P(A) = 0.25$, $P(B) = 0.55$, and $P(A \cup B) = 0.7$.
What is $P(A \cap \overline{B})$?

3) Are the distribution of the following series symmetrical ? 2-4-5-5-5-6-6-7

Exercise2 (7 points): Complete the following table representing an interrogation 300 person

class	frequency(f_i)	mid_point(m_i)	cumulative frequency	%	$f_i m_i$
$[4-6[$				5	
		7			308
$[...-12[$					
	102		225		
$[14-...[$	40				
$[16-20[$		18			

2) Find the median (Me)

Exercise3 (7 points): In a factory, Three machines M_1, M_2, M_3 produce $40\%, 35\%$ and 25%, respectively, of the total daily production of plastic bottles. It is known that $9\%, 5\%$, and 7% of the bottles produced in each of the M_1, M_2, and M_3 machines are known to be defective. If one bottle is picked up at random from a day's production.

1) Draw a tree diagram

2) Calculate the probability that it is defective

3) If that the tbottle is defective.
What is the probability that it is from M_3

Bibliography

[1] CHAlMER, Dean. Cambridge International AS and A Level Mathematics, Probability & Statistics 1 Coursebook. Vol. 1. Cambridge University Press, 2018.

[2] CLAUDIA Von Brömssen, Jesper Rydén, et al. Collection of Exercises in Basic Statistics, Swedish university of agricultural sciences, 2019

[3] DEKKING, Frederik Michel, et al. A Modern Introduction to Probability and Statistics, Understanding why and how. London: springer, 2005.

[4] HON, Keone. An introduction to statistics. alt2. mathlinks. ro, no. February, 2010, 1-29.

[5] ILLOWSKY, Barbara. DEAN, Susan. Introductory statistics. Openstax, 2018.

[6] WATKINS, Joseph C. An introduction to the science of statistics, From theory to implementation, 2019.