Mohamed Boudiaf University - M'sila
Faculty of Technology
Department of Electrical Engineering
Module : Probabilities and Statistics

University Year : 2025/2026

Solutions to the Directed Works Series n°1

Exercice n°1 : (The different types of statistical variables)

1) Classify the variables below according to their type :

a) Number of people per household b) Height (cm) c) Sex.
d) marital status e) Weight (kg) f) Level of education

2) Specify the measurement or values they can take.

Solution :

Reminder :

— **Quantitative discrete :** Variables that take countable integer values (e.g., household size).
— **Quantitative continuous :** Variables measured on a continuous scale, can take any real value (e.g., height, weight).
— **Qualitative nominal :** Categorical variables without any natural order (e.g., sex, marital status).
— **Qualitative ordinal :** Categorical variables with a natural order or ranking (e.g., education level).

Var.	**Variable**	**Type**	**Modalities / Values**
a	Number of people per household	Quantitative discrete	2, 3, 4,...
b	Height (cm)	Quantitative continuous	Any real value (e.g., 150.2 cm, 172 cm, 189.5 cm)
c	Sex	Qualitative nominal	Male, Female
d	Marital status	Qualitative nominal	Single, Married, Divorced, Widowed
e	Weight (kg)	Quantitative continuous	Any real value (e.g., 52.7 kg, 68 kg, 95.4 kg)
f	Level of education	Qualitative ordinal	Primary, Secondary, University, Postgraduate
g	Mode of transport	Qualitative nominal	Car, Bus, Train, Bicycle, Walking, Motorcycle, etc.

Exercice n°2 : : (qualitative variable) The following table shows the blood groups of 2^{nd} year ST students

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

1) Indicate the population studied, the statistical unit, the character studied and its nature .

2) Draw up the table of the absolute frequencies n_i representing this statistical serie, calculating the relative frequencies f_i and percentages p_i.

3) Give two graphical representations suitable for this type of characteristic.
4) What is the mode of the series ?
Solution :

1) Study elements :
— Population : 2nd-year ST students.
— Statistical unit : One student.
— Character studied : Blood group.
— Nature : Qualitative nominal.

2) Frequency table :

Blood group	Absolute frequency n_i	Relative frequency f_i	Percentage p_i
A	21	$\frac{21}{54} \approx 0.3889$	38.89%
O	15	$\frac{15}{54} \approx 0.2778$	27.78%
B	10	$\frac{10}{54} \approx 0.1852$	18.52%
AB	8	$\frac{8}{54} \approx 0.1481$	14.81%
Total	54	1	100%

3) Graphical representations :

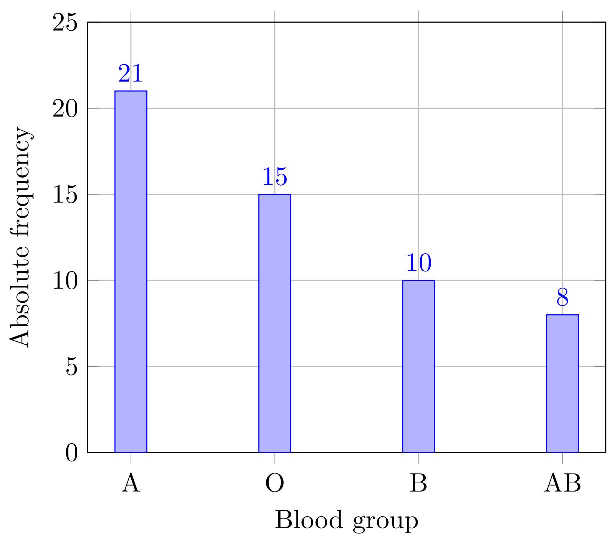

FIGURE 1 – Bar chart of blood groups

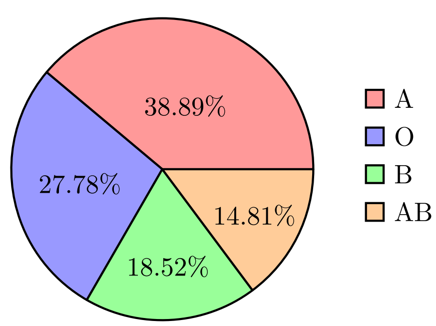

FIGURE 2 – Pie chart of blood groups

4) Mode of the series : The mode is **A**, since it has the highest frequency ($n_A = 21$).

Exercice n°3 : Show that :

i) The sum of the deviations from the mean is 0 :
$\Sigma_{i=1}^{k} n_i(x_i - \bar{x}) = 0$

ii) Linearity property : If
$\forall i \in \{1, 2, ..., k\}, a, b \in \mathbb{R}$ we have : $y_i = ax_i + b$, then :$\bar{y} = a\bar{x} + b$.

iii) * Property of partial means : Consider two statistical series with respective means $\bar{x_1}$ and $\bar{x_2}$ and respective numbers N_1 and N_2.
The mean of the two series is

$$\bar{x} = \frac{N_1\bar{x_1} + N_2\bar{x_2}}{N_1 + N_2}$$

Solution :

Properties of the mean

i) Sum of deviations from the mean equals zero.

Let x_i be the modalities, n_i their frequencies, $N = \sum_{i=1}^{k} n_i$, and $\bar{x} = \frac{1}{N}\sum_{i=1}^{k} n_i x_i$. Then :

$$\sum_{i=1}^{k} n_i(x_i - \bar{x}) = \sum_{i=1}^{k} n_i x_i - \bar{x}\sum_{i=1}^{k} n_i = \sum_{i=1}^{k} n_i x_i - \bar{x}N = \sum_{i=1}^{k} n_i x_i - \left(\frac{1}{N}\sum_{i=1}^{k} n_i x_i\right) N = 0.$$

ii) Linearity property.

If $y_i = ax_i + b$ for all i, then

$$\bar{y} = \frac{1}{N}\sum_{i=1}^{k} n_i y_i = \frac{1}{N}\sum_{i=1}^{k} n_i(ax_i + b) = a\left(\frac{1}{N}\sum_{i=1}^{k} n_i x_i\right) + b\left(\frac{1}{N}\sum_{i=1}^{k} n_i\right) = a\bar{x} + b.$$

iii) Property of partial means.

For two series with means $\bar{x_1}, \bar{x_2}$ and sizes N_1, N_2, their totals are $S_1 = N_1\bar{x_1}$ and $S_2 = N_2\bar{x_2}$. The combined mean is

$$\bar{x} = \frac{S_1 + S_2}{N_1 + N_2} = \frac{N_1\bar{x_1} + N_2\bar{x_2}}{N_1 + N_2}.$$

Exercice n°4 :

The variance is the quantity :

$$V(x) = \Sigma_{i=1}^{k} \frac{n_i(x_i - \bar{x})^2}{N} = \Sigma_{i=1}^{k} f_i(x_i - \bar{x})^2 \tag{1}$$

Show that :

1) $V(x) = \Sigma_{i=1}^{k} \frac{n_i x_i^2}{N} - \bar{x}^2 = \Sigma_{i=1}^{k} f_i x_i^2 - \bar{x}^2$

2) If $\forall i \in \{1, 2, ..., k\}, a, b \in \mathbb{R}$ we have : $y_i = ax_i + b$, $a, b \in R$, then :
$V(y) = a^2 V x$,
$\sigma_y = |a|\sigma_x$.

3) $V(x) \geq 0$.

Solution :

Variance properties

The variance is defined as :

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

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

1) Show that :

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

Proof :

$$\begin{aligned}
V(x) &= \frac{1}{N} \sum n_i (x_i - \bar{x})^2 \\
&= \frac{1}{N} \sum n_i (x_i^2 - 2x_i\bar{x} + \bar{x}^2) \\
&= \frac{1}{N} \sum n_i x_i^2 - \frac{2\bar{x}}{N} \sum n_i x_i + \frac{\bar{x}^2}{N} \sum n_i \\
&= \frac{1}{N} \sum n_i x_i^2 - 2\bar{x} \cdot \bar{x} + \bar{x}^2 \cdot 1 \\
&= \frac{1}{N} \sum n_i x_i^2 - \bar{x}^2.
\end{aligned}$$

2) If $y_i = ax_i + b$, with $a, b \in \mathbb{R}$, then :

$$V(y) = a^2 V(x), \quad \sigma_y = |a| \sigma_x.$$

Proof :

$$\begin{aligned}
V(y) &= \frac{1}{N} \sum n_i (y_i - \bar{y})^2 \\
&= \frac{1}{N} \sum n_i \big(a(x_i - \bar{x})\big)^2 \\
&= a^2 \cdot \frac{1}{N} \sum n_i (x_i - \bar{x})^2 \\
&= a^2 V(x).
\end{aligned}$$

Thus $\sigma_y = \sqrt{V(y)} = \sqrt{a^2 V(x)} = |a| \sigma_x$.

3) Show that $V(x) \geq 0$.

Proof : Each term $(x_i - \bar{x})^2 \geq 0$ and $n_i \geq 0$. Therefore :

$$V(x) = \frac{1}{N} \sum n_i (x_i - \bar{x})^2 \geq 0,$$

with equality iff all x_i are equal.