12

Informations chiffrées. Statistiques descriptives

Au programme

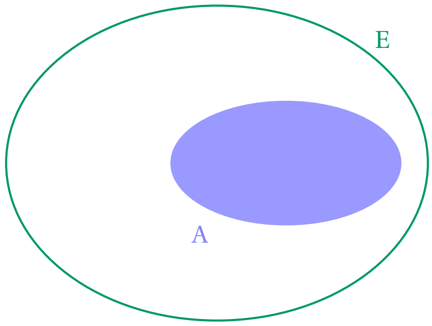
- √ Consolider les calculs de pourcentages
- √ Calculer des proportions de proportions ou des pourcentages de pourcentages
- √ Manipuler des taux d'évolution, des taux d'évolutions successives, des taux d'évolutions réciproques
- √ Manipuler différents paramètres statistiques : moyenne, médiane, écart-type, quartiles . . .

Table des matières

La **statistique descriptive** a pour but, comme son nom l'indique, de décrire par un certain nombre de données numériques, une **population** d'individus ou d'objets.

Le mot population a un sens très général. Cette population peut être l'**ensemble** des élèves d'une classe, l'ensemble des éléphants d'Afrique, l'ensemble des voitures du parc automobile français, l'ensemble des composants électroniques d'un ordinateur . . .

Puisque cette population est un ensemble, nous la noterons systématiquement E dans ce chapitre. Les parties A de E seront alors des **sous-populations**. Par exemple, A est l'ensemble des filles dans l'ensemble E des élèves de la classe.



On commence par une mise en jambes :

Exercice 1

Deux magasins A et B annoncent les prix de différents produits accompagnés parfois d'offres promotionnelles. Déterminer pour chaque produit quelle est l'offre la plus intéressante.

Produit	Magasin A	Magasin B
Sucre	Poids net : 750 g. Prix : $1,35$ €	Poids net : 600 g. Prix : $1,20$ €
Eau minérale	Pack de 4 bouteilles de $1,5$ l. Prix : $2,16$ €	Bidon de 5 l. Prix : $1,70$ €
Jus de pomme bio	La bouteille de $1,5$ l, $3,46$ €. Pour l'achat de deux bouteilles, la deuxième à moitié prix	La bouteille de $1,5$ l, $3,46$ €. Pour l'achat de deux bouteilles, la troisième gratuite
Machine à café	Prix : 70 €. 20% de remise immédiate en caisse	Prix : 75 €. 20 € de remise en caisse.

Solution 1 :

• Le sucre. Calculons le prix d'un kilogramme de sucre dans chacun des deux magasins.

Dans le magasin A, 750 grammes de sucre coûtent $1,35$ euro puis 1 gramme de sucre coûte $\dfrac{1,35}{750}$ € puis 1 kilogramme de sucre coûte $\dfrac{1,35}{750}\times 1\,000$ € ou encore $1,8$ €.

Dans le magasin B, 600 grammes de sucre coûtent $1,2$ euro puis 1 gramme de sucre coûte $\dfrac{1,2}{600}$ € puis 1 kilogramme de sucre coûte $\dfrac{1,2}{600}\times 1\,000$ € ou encore 2 €.

En ce qui concerne le sucre, l'offre du magasin A est plus avantageuse que l'offre du magasin B.

• L'eau minérale. Calculons le prix d'un litre d'eau dans chacun des deux magasins.

Dans le magasin A, puisque $4\times 1,5 = 6$, 6 litres d'eau coûtent $2,16$ euro puis 1 litre d'eau coûte $\dfrac{2,16}{6}$ € ou encore $0,36$ €.

Dans le magasin B, 5 litres d'eau coûtent $1,7$ euro puis 1 litre d'eau coûte $\dfrac{1,7}{5}$ € ou encore $0,34$ €.

En ce qui concerne l'eau minérale, l'offre du magasin B est plus avantageuse que l'offre du magasin A.

• Le jus de pomme bio. Calculons le prix d'un litre de jus de pomme bio dans chacun des deux magasins.

Dans le magasin A, le prix de trois litres de jus de pomme bio est $3,46+\dfrac{3,46}{2}$ € ou encore $5,19$ €. Le prix d'un litre de jus de pomme bio est donc $\dfrac{5,19}{3}$ euro ou encore $1,73$ €.

Dans le magasin B, $4,5$ litres de jus de pomme bio coûtent $3,46+3,46$ € ou encore $7,92$ €. Le prix d'un litre de jus de pomme bio est donc $\dfrac{7,92}{4,5}$ € ou encore $1,76$ €.

En ce qui concerne le jus de pomme bio, l'offre du magasin A est plus avantageuse que l'offre du magasin B.

• Dans le magasin A, le prix de la machine à café est $70\times\left(1-\dfrac{20}{100}\right)$ € ou encore $70\times 0,8$ € ou enfin 56 €.
Dans le magasin B, le prix de la machine à café est $75-20$ € ou encore 55 €.

En ce qui concerne la machine à café, l'offre du magasin B est (très légèrement) plus avantageuse que l'offre du magasin A.

■

I Informations chiffrées

A Effectifs

Définition 1

Soient E une population puis A une partie de E.

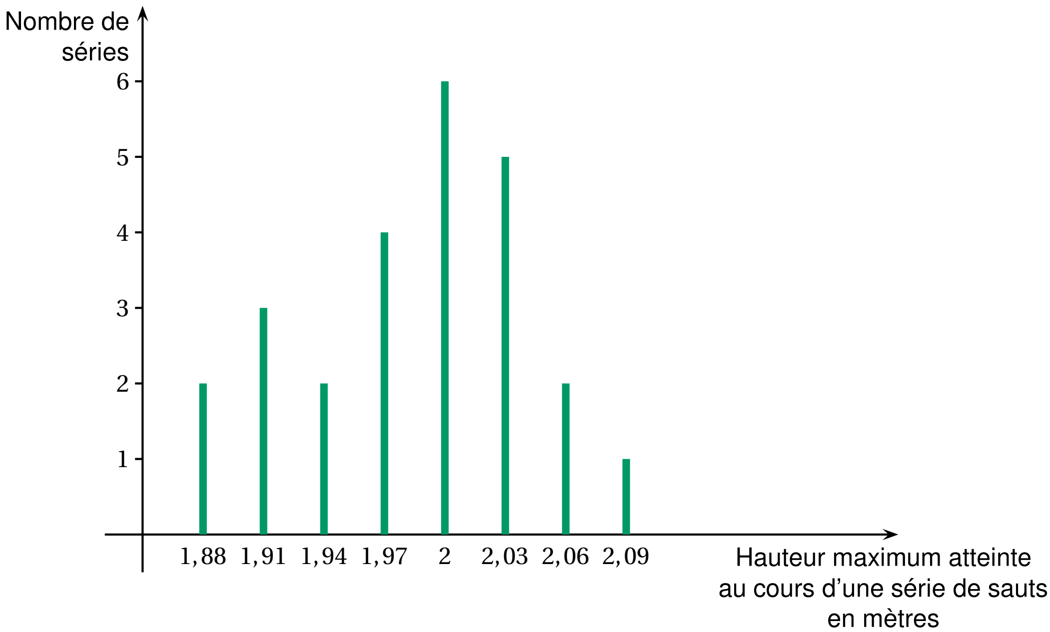
L'**effectif** de la sous-population A est le nombre d'individus de cette sous-population.
L'effectif de la population E est noté n_{E} et l'effectif de la sous-population A est noté n_{A}.
L'effectif de la population E est l'**effectif total** et celui de la sous-population A est un **effectif partiel**.

Par exemple, si une classe E est constituée de 37 élèves dont 14 sont des filles et si on note F l'ensemble des filles et G l'ensemble des garçons, $n_{\mathrm{E}}=37$, $n_{\mathrm{F}}=14$ et $n_{\mathrm{G}}=37-14=23$.

B Représentations graphiques

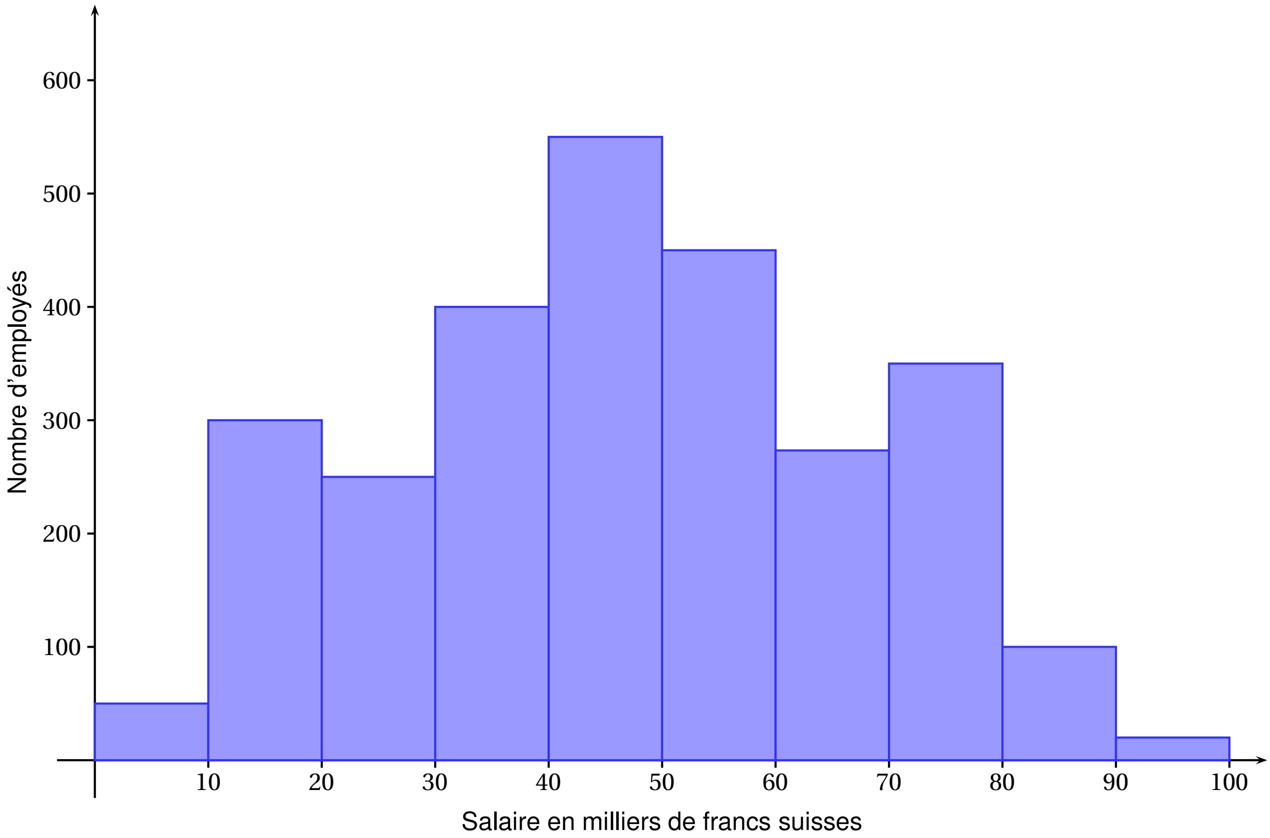
Il existe de très nombreuses manières de représenter graphiquement des données numériques. On en donne quelques unes.

Diagramme bâtons. Le graphique ci-dessous représente différentes séries de sauts en hauteurs effectués par un sauteur. On a placé en abscisse la hauteur maximum franchie par le sauteur au cours d'une série de sauts et en ordonnée le nombre de séries de sauts où le sauteur a atteint cette hauteur.

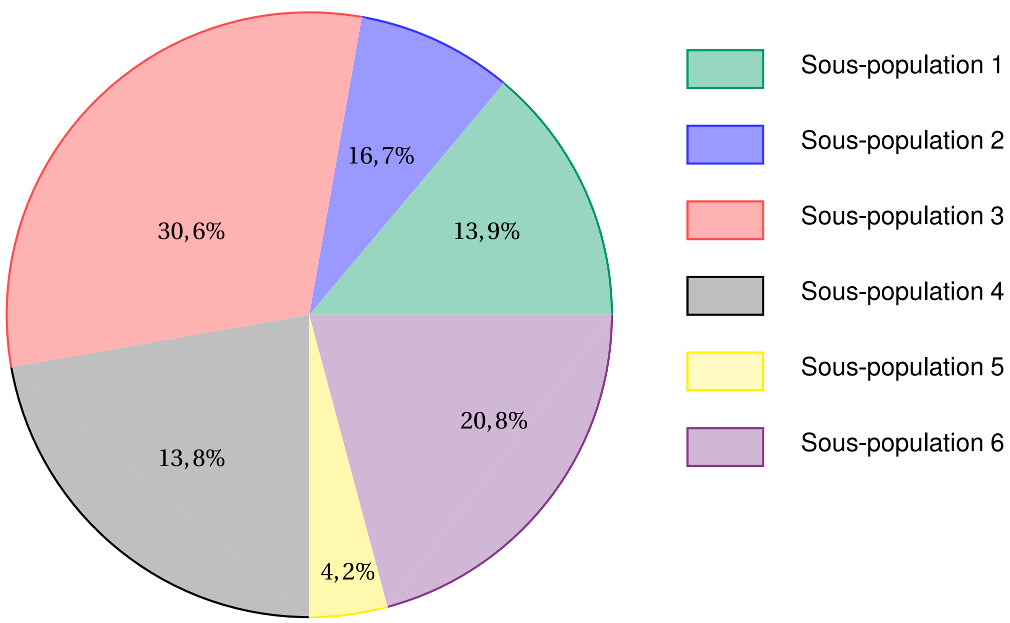


Sur graphique, on lit par exemple que le sauteur a effectué 4 séries de sauts durant lesquelles il a atteint $1,97$ m et pas plus. On voit aussi que la hauteur maximale qu'il atteint le plus couramment est 2 m (hauteur atteinte 6 fois).

Histogrammes. Ci-dessous, on a représenté les différents salaires annuels gagnés par les employés d'une certaine société suisse. En abscisse, les salaires sont regroupés en « tranches de salaire » ou encore en intervalles. Sur ce graphique, on voit par exemple qu'environ 550 employés (lu en ordonnée) perçoivent un salaire annuel compris entre 40 et 50 milliers de francs suisses (lu en abscisse). On note que pour un graphique statistiques, le repère choisi n'est quasiment jamais orthonormé. Le choix des unités est dicté par la lisibilité du graphique. Ce choix n'est pas anodin car il peut influer sur notre perception des choses et notre interprétation des données.



Diagrammes circulaires. Une certaine population se décompose en 6 sous-populations. Chaque secteur angulaire coloré représente l'une d'entre elles. Chacune des aires est proportionnelle à l'effectif ou au pourcentage de la sous-population correspondante. Par exemple, les sous-populations 5 et 6 ensembles font 25% de la population. Le secteur angulaire jaune+fuchsia occupe $90°$, c'est-à-dire 25% de $360°$.



C Proportions (ou fréquences), pourcentages

Définition 2

Soient E une population puis A une sous-population (ou un sous-ensemble) de E.

La **proportion** ou la **fréquence** de la sous-population A dans la population E est le quotient de l'effectif de la sous-population A par l'effectif de la population E.

Si on note t cette proportion (t est l'initiale de « taux »), alors $t = \frac{n_\mathrm{A}}{n_\mathrm{E}}$ (ou aussi si on note f la fréquence $f = \frac{n_\mathrm{A}}{n_\mathrm{E}}$).

Remarque. Nous avons adopté la notation t (pour « taux ») pour désigner une proportion pour garder à disposition la lettre p pour un pourcentage.

Par exemple, dans la classe de 37 élèves dont 14 sont des filles, la proportion de filles est $t = \frac{14}{37} = 0,37\ldots$

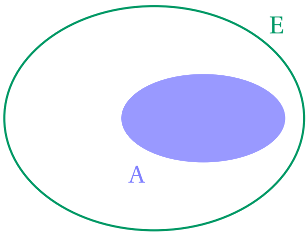
A partir de l'égalité $t = \frac{n_\mathrm{A}}{n_\mathrm{E}}$, en effectuant le « produit en croix », on obtient deux autres égalités :

Theoreme 1

Soient E une population puis A une sous-population de E. Soit $t = \frac{n_\mathrm{A}}{n_\mathrm{E}}$.

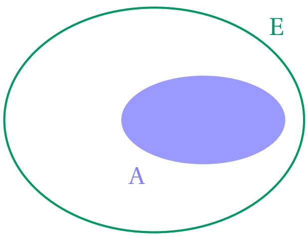
Alors, $n_\mathrm{A} = t \times n_\mathrm{E}$ et $n_\mathrm{E} = \frac{n_\mathrm{A}}{t}$.

• L'égalité $t = \frac{n_\mathrm{A}}{n_\mathrm{E}}$ est utilisée pour calculer la proportion t quand on connaît les effectifs n_E et n_A.



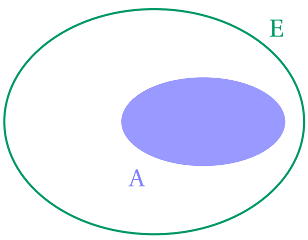
$\xrightarrow{\frac{n_\mathrm{A}}{n_\mathrm{E}}} \quad t = \frac{n_\mathrm{A}}{n_\mathrm{E}}$

• L'égalité $n_\mathrm{A} = t \times n_\mathrm{E}$ est utilisée pour calculer l'effectif n_A quand on connaît la proportion t et l'effectif n_E. Pour obtenir l'effectif de la sous-population A, on **multiplie** l'effectif de la population E par la proportion t.



$\xrightarrow{\times t} \quad n_\mathrm{A} = t \times n_\mathrm{E}$

• L'égalité $n_\mathrm{E} = \frac{n_\mathrm{A}}{t}$ est utilisée pour calculer l'effectif n_E quand on connaît la proportion t et l'effectif n_A. Pour obtenir l'effectif de la population E, on **divise** l'effectif de la sous-population A par la proportion t.



$\xrightarrow{\div t} \quad n_\mathrm{E} = n_\mathrm{A} \div t = \frac{n_\mathrm{A}}{t}$

Exercice 2

Dans une certaine classe, il y a un tiers de filles soit 12 filles. Combien y-a-t-il d'élèves dans la classe ?

Solution 2 : Soit n le nombre d'élèves de la classe. $\frac{1}{3} = \frac{12}{n}$ et donc $n = 3 \times 12 = 36$. Il y a 36 élèves dans la classe.

■

On peut décider d'exprimer cette proportion en pourcentage.

Définition 3

Soient E une population puis A une sous-population (ou un sous-ensemble) de E. Soit t (pour « taux ») la proportion de la sous-population A dans la population E.

Le **pourcentage** de la sous-population A dans la population E est le réel p tel que $t = \dfrac{p}{100}$.

Immédiatement, on a :

Theoreme 2

Soient E une population puis A une sous-population (ou un sous-ensemble) de E. Soient t (pour « taux ») la proportion de la sous-population A dans la population E et p le pourcentage de la sous-population A dans la population E.

Alors, $p = 100 \times t$.

☞ *On obtient le pourcentage en multipliant la proportion par* 100.

Connaissant le pourcentage p d'une sous-population A dans une population E, on peut retrouver son effectif :

Theoreme 3

Soient E une population puis A une sous-population (ou un sous-ensemble) de E. Soit t la proportion de la sous-population A dans la population E et p le pourcentage de la sous-population A dans la population E.

Alors, $n_{\mathrm{A}} = t \times n_{\mathrm{E}} = \dfrac{p}{100} \times n_{\mathrm{E}}$.

On peut bien sûr calculer une proportion ou un pourcentage de n'importe quelle quantité et pas seulement d'une population. Soient x et y deux quantités.
La proportion t de la quantité strictement positive y par rapport à la quantité strictement positive x est $t = \dfrac{y}{x}$. On a alors $y = tx$ et aussi $x = \dfrac{y}{t}$.
On note que si $y > x$, alors $t > 1$, si $\dfrac{y}{x}$, alors $t < 1$ et si $y = x$, alors $t = 1$.
La pourcentage p de la quantité y par rapport à la quantité x est $p = 100t$. On a alors $y = \dfrac{px}{100}$ et aussi $x = \dfrac{100y}{p}$.
Si $y > x$, le pourcentage est strictement supérieur à 100.

Par exemple, 200 € sont 200% de 100 €.

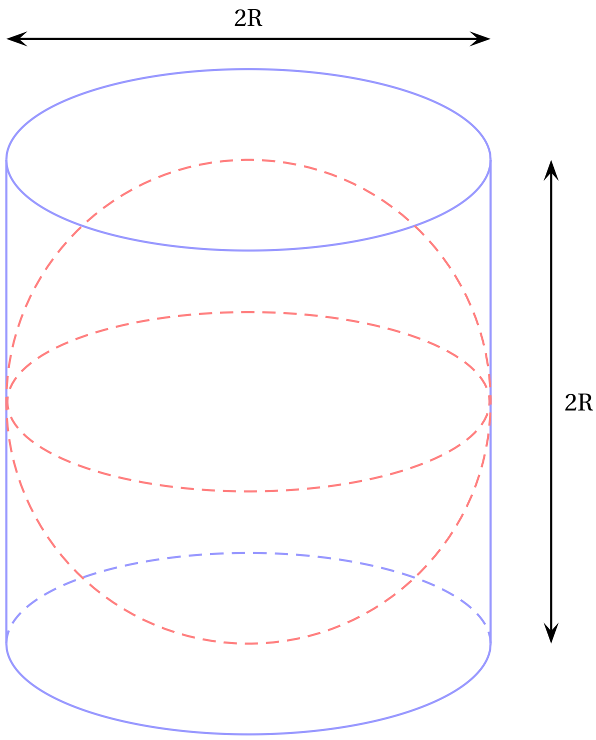
Exercice 3

Une boule de rayon $\mathrm{R} > 0$ est placée à l'intérieur d'un cylindre de rayon R et de hauteur $2\mathrm{R}$.

1) Quel pourcentage du volume du cylindre est occupé par la boule ?
(On rappelle que le volume d'une boule de rayon R est $\dfrac{4}{3}\pi\mathrm{R}^3$ et le volume d'un cylindre de révolution de rayon R et de hauteur h est $\pi\mathrm{R}^2 h$.)

2) Quel pourcentage de l'aire du cylindre représente l'aire la sphère ?
(On admet que l'aire d'une sphère de rayon R est $4\pi\mathrm{R}^2$.)

Solution 3 :



1) Le volume de la boule est $V_b = \dfrac{4}{3}\pi R^3$. Le volume du cylindre est $V_c = \pi R^2 \times 2R = 2\pi R^3$. La proportion du volume du cylindre occupé par la boule est

$$t = \frac{V_b}{V_c} = \frac{\frac{4}{3}\pi R^3}{2\pi R^3} = \frac{\frac{4}{3}}{2} = \frac{4}{3} \times \frac{1}{2} = \frac{2}{3}.$$

Le pourcentage du volume du cylindre occupé par la boule est $\dfrac{2}{3} \times 100\%$ ou encore $66,6\%$ arrondi à 10^{-1}.

2) L'aire de la sphère est $A_s = 4\pi R^2$. L'aire A_c du cylindre est la somme de l'aire d'un rectangle dont les côtés ont pour longueur $2R$ et $2\pi R$ et des aires de deux disques de rayon R. Donc,

$$A_c = 2R \times 2\pi R + 2\pi R^2 = 6\pi R^2.$$

La proportion de l'aire du cylindre demandée est

$$t = \frac{A_s}{A_c} = \frac{4\pi R^2}{6\pi R^2} = \frac{4}{6} = \frac{2}{3}.$$

On obtient la même proportion que pour le volume et donc le même pourcentage : $66,6\%$ arrondi à 10^{-1}.

■

D Proportions de proportions (ou pourcentages de pourcentages)

Dans une certaine classe, 40% des élèves sont des filles et d'autre part, il y a 30% des filles qui sont blondes. On se demande alors quelle est la proportion d'élèves qui sont des filles blondes dans la classe. On note respectivement n_E, n_F et n_{FB} le nombre d'élèves de la classe, le nombre de filles et le nombre de filles blondes. On note encore t_F la proportion de filles dans la classe, t_B la proportion de filles qui sont blondes puis t_{FB} la proportion d'élèves qui sont des filles blondes (c'est-à-dire le proportion que l'on veut calculer).

On a $t_F = \dfrac{n_F}{n_E}$ et $t_B = \dfrac{n_{FB}}{n_F}$. En **multipliant** ces égalités membre à membre, on obtient

$$t_F \times t_B = \frac{n_F}{n_E} \times \frac{n_{FB}}{n_F} = \frac{n_F \times n_{FB}}{n_E \times n_F} = \frac{n_{FB}}{n_E} = t_{FB}.$$

Ainsi, la proportion d'élèves qui sont des filles blondes dans la classe est

$$t_{\mathrm{FB}} = t_{\mathrm{F}} \times t_{\mathrm{B}} = \frac{40}{100} \times \frac{37}{100} = \frac{40 \times 37}{100 \times 100} = \frac{148}{1000} = 0,148.$$

En pourcentage, cela donne (en notant p_{FB} le pourcentage d'élèves de la classe qui sont des filles blondes) :

$$p_{\mathrm{FB}} = 100 \times t_{\mathrm{FB}} = 100 \times \frac{148}{1000} = \frac{148}{10} = 14,8.$$

Il y a donc $14,8\%$ des élèves de la classe qui sont des filles blondes. De manière générale,

Theoreme 4

Soit E une population. Soient A une sous-population de E puis B une sous-population de A.

Soient t_{A} la proportion de la sous-population A dans la population E puis $t_{\mathrm{B/A}}$ la proportion de la sous-population B dans la population A. Alors, la proportion t_{B} de B dans E est $t_{\mathrm{A}} \times t_{\mathrm{B/A}}$:

$$t_{\mathrm{B}} = t_{\mathrm{A}} \times t_{\mathrm{B/A}}.$$

Démonstration : On note respectivement n_{E}, n_{A} et n_{B} l'effectif de la population E, l'effectif de la population A et l'effectif de la population B (puisque $\mathrm{B} \subset \mathrm{A}$, on a $n_{\mathrm{B}} \leqslant n_{\mathrm{A}}$).
On a $t_{\mathrm{A}} = \dfrac{n_{\mathrm{A}}}{n_{\mathrm{E}}}$, $t_{\mathrm{B/A}} = \dfrac{n_{\mathrm{B}}}{n_{\mathrm{A}}}$. En multipliant membre à membre ces deux égalités, on obtient (en tenant compte de $t_{\mathrm{B}} = \dfrac{n_{\mathrm{B}}}{n_{\mathrm{E}}}$) :

$$t_{\mathrm{A}} \times t_{\mathrm{B/A}} = \frac{n_{\mathrm{A}}}{n_{\mathrm{E}}} \times \frac{n_{\mathrm{B}}}{n_{\mathrm{A}}} = \frac{n_{\mathrm{A}} \times n_{\mathrm{B}}}{n_{\mathrm{E}} \times n_{\mathrm{A}}} = \frac{n_{\mathrm{B}}}{n_{\mathrm{E}}} = t_{\mathrm{B}}.$$

■

Si on veut le pourcentage d'éléments de B dans E, le plus simple est de faire comme dans l'exemple : on calcule d'abord la proportion de B dans E puis on multiplie par 100 pour avoir le pourcentage d'éléments dans E. On peut néanmoins donner un résultat général.

Theoreme 5

Soit E une population. Soient A une sous-population de E puis B une sous-population de A.

Soient p_{A} la pourcentage d'éléments de A dans E puis $p_{\mathrm{B/A}}$ la pourcentage d'éléments de B dans A. Alors, la pourcentage p_{B} d'éléments de B dans E est donné par :

$$p_{\mathrm{B}} = \frac{p_{\mathrm{A}} \times p_{\mathrm{B/A}}}{100}.$$

Démonstration : L'égalité $t_{\mathrm{B}} = t_{\mathrm{A}} \times t_{\mathrm{B/A}}$ s'écrit encore $\dfrac{p_{\mathrm{B}}}{100} = \dfrac{p_{\mathrm{A}}}{100} \times \dfrac{p_{\mathrm{B/A}}}{100}$ puis

$$p_{\mathrm{B}} = 100 \times \frac{p_{\mathrm{A}}}{100} \times \frac{p_{\mathrm{B/A}}}{100} = \frac{p_{\mathrm{A}} \times p_{\mathrm{B/A}}}{100}.$$

■

Ainsi, 10% de 10% d'une quantité est $\dfrac{10 \times 10}{100}\%$ de cette quantité c'est-à-dire 1% de cette quantité.

On ne se contente pas de multiplier les pourcentages. **On multiplie les pourcentages puis on divise par 100**. 10% de 10% ne font pas 100%.

Exercice 4

Au Scrabble, 45% des jetons (jokers non compris) sont des voyelles et d'autre part, 20% des voyelles sont des A.

Calculer le pourcentage de jetons portant la lettre A dans l'ensemble des jetons.

Solution 4 : Notons respectivement t_{V} et $t_{\mathrm{A/V}}$ la proportion de voyelles dans l'ensemble des lettres et la proportion de voyelles qui sont des A. On a $t_{\mathrm{V}} = \dfrac{45}{100} = 0,45$ et $t_{\mathrm{A/V}} = \dfrac{20}{100} = 0,2$. Mais alors, la proportion de A dans l'ensemble des lettres est

$$t_{\mathrm{A}} = t_{\mathrm{V}} \times t_{\mathrm{A/V}} = 0,45 \times 0,2 = 0,09.$$

Le pourcentage de A dans l'ensemble des lettres est donc

$$p_A = 100 \times t_A = 100 \times 0,09 = 9.$$

Il y a 9% des lettres qui sont des A.

■

II Evolutions

A Variation absolue et variation relative (ou taux d'évolution)

Définition 4

Une certaine quantité évolue d'une valeur initiale strictement positive V_i à une valeur finale strictement positive V_f.

La **variation absolue** de cette quantité est $V_f - V_i$ (valeur finale moins valeur initiale).

La **variation relative** ou **taux d'évolution** de cette quantité est $t = \dfrac{V_f - V_i}{V_i}$ c'est-à-dire le quotient de la variation absolue par la valeur initiale.

Par exemple, quand un prix passe de $P_i = 10$ euros à $P_f = 11$ euros, la variation absolue de ce prix est $11 - 10$ euros ou encore $+1$ euro. Et quand un prix passe de $10\,000$ euros à $11\,000$ euros, la variation absolue de ce prix est $11\,000 - 10\,000$ euros c'est-à-dire $+1\,000$ euros.

Maintenant, $\dfrac{11-10}{10} = \dfrac{1}{10} = 0,1$ et $\dfrac{11\,000 - 10\,000}{10\,000} = \dfrac{1\,000}{10\,000} = 0,1$. Dans les deux cas, le taux d'évolution est $t = 0,1$ (il revient au même de dire que les deux prix ont subi une augmentation de 10%).

Remarque. Si $t > 0$, la quantité subit une **augmentation**, si $t < 0$, la quantité subit une **diminution** et si $t = 0$, la quantité reste **constante**.

B Coefficient multiplicateur associé au taux d'évolution

Définition 5

Au cours d'une évolution, le coefficient multiplicateur est $C_M = \dfrac{V_f}{V_i}$.

Theoreme 6

$C_M = 1 + t$ puis $V_f = C_M \times V_i = (1+t)V_i$.

Démonstration : $t = \dfrac{V_f - V_i}{V_i} = \dfrac{V_f}{V_i} - \dfrac{V_i}{V_i} = C_M - 1$ puis $C_M = 1 + t$. Ensuite, l'égalité $\dfrac{V_f}{V_i} = C_M$ fournit, en effectuant le « produit en croix », $V_f = C_M \times V_i = (1+t)V_i$.

■

Remarque. Si $C_M > 1$, la quantité subit une augmentation, si $C_M < 1$, la quantité subit une diminution et si $C_M = 1$, la quantité ne varie pas.

C Evolution en pourcentage

On rappelle que le taux d'évolution peut être strictement négatif (quand il y a diminution), strictement positif (quand il y a augmentation) et nul (quand il n'ya pas de variation).

Définition 6

Soient V_i et V_f les valeurs initiale et finale d'une certaine quantité puis $t = \dfrac{V_f - V_i}{V_i}$ le taux d'évolution.

Le pourcentage d'évolution est $p = 100 \times t$.

Que l'on ait $t<0$ (et donc $p<0$), $t>0$ (et donc $p>0$) ou $p=0$, on a

Theoreme 7

$V_f=\left(1+\dfrac{p}{100}\right)V_i$.

Maintenant, on peut donner des résultats différenciés pour une augmentation et une diminution :

Theoreme 8

Soit p un réel strictement positif. Augmenter une valeur V_i de $p\%$, c'est multiplier V_i par $1+\dfrac{p}{100}$: $V_f=\left(1+\dfrac{p}{100}\right)V_i$.

Soit p un réel de $]0,100]$. Diminuer une valeur V_i de $p\%$, c'est multiplier V_i par $1-\dfrac{p}{100}$: $V_f=\left(1-\dfrac{p}{100}\right)V_i$.

Exercice 5

1) Un prix passe de 100 euros à 300 euros. Quel est le pourcentage p d'augmentation ?

2) Un prix passe de 100 euros à 50 euros. Quel est le pourcentage p de diminution ?

Solution 5 :

1) $V_f-V_i=300-100=200$ puis $t=\dfrac{V_f-V_i}{V_i}=\dfrac{200}{100}=2$ puis $p=100\times t=200$. Quand un prix passe de 100 euros à 300 euros, le pourcentage d'augmentation est 200%.

2) $t=\dfrac{V_f-V_i}{V_i}=\dfrac{50-100}{100}=-\dfrac{50}{100}=-0,5$ puis $|t|=0,5$ puis $p=100|t|=50$. Quand un prix passe de 100 euros à 50 euros, le pourcentage de diminution est 50%.

■

D Evolutions successives

Quand une certaine quantité Q a une valeur initiale V_i et subit une première évolution dont le taux est noté t_1, la nouvelle valeur est $V_1=(1+t_1)V_i$. Si cette nouvelle quantité subit une nouvelle évolution dont le taux est noté t_2, alors la nouvelle valeur est

$$V_2=(1+t_2)V_1=(1+t_2)(1+t_1)V_i.$$

De manière générale :

Theoreme 9

Une certaine quantité Q a une valeur initiale (strictement positive) V_i et subit deux évolutions successives. On note t_1 et t_2 les taux d'évolution respectifs associés à chacune des deux évolutions. On note encore V_f la valeur finale de la quantité Q. Alors,

$$V_f=(1+t_1)(1+t_2)V_i.$$

Plus généralement, si la quantité Q subit n évolutions successives (où n est un entier supérieur ou égal à 2) et si $t_1, t_2, \ldots, t_n$, sont les taux d'évolution respectifs (taux positifs en cas d'augmentation et négatifs en cas de diminution) associés à chacune des n évolutions, alors

$$V_f=(1+t_1)(1+t_2)\ldots(1+t_n)V_i.$$

Si $p_1, p_2, \ldots, p_n$, sont les pourcentages d'évolution successifs (pourcentages positifs en cas d'augmentation et négatifs en cas de diminution),

$$V_f=\left(1+\frac{p_1}{100}\right)\left(1+\frac{p_2}{100}\right)\ldots\left(1+\frac{p_n}{100}\right)V_i.$$

Exercice 6

Les phrases suivantes sont-elle vraie ou fausse ?

Phrase 1 : « Quand on augmente de 10% puis qu'on rediminue de 10%, on revient au point de départ ».

Phrase 2 : « Il revient au même d'augmenter de 10% puis de rediminuer de 10% ou de diminuer de 10% puis de réaugmenter de 10% ».

Solution 6 : Soit x une certaine quantité (strictement positive).
La quantité x augmentée de 10% est $y = \left(1+\dfrac{10}{100}\right) = 1,1x$ puis la quantité y diminuée de 10% est

$$z = \left(1-\frac{10}{100}\right)y = 0,9y = 0,9\times 1,1x = 0,99x.$$

La quantité finale est différente de la quantité initiale et donc la phrase 1 est fausse.

Si maintenant on commence par diminuer x de 10% puis on réaugmente la quantité obtenue de 10%, la quantité finale est $1,1\times 0,9x = 0,99x$. La phrase 2 est donc vraie.

■

Exercice 7

Au premier janvier 2022, le cours d'une certaine action est $2,45$ €.

Le tableau ci-dessous donne l'évolution du cours de cette action au premier jour de chaque mois (par rapport au premier jour du mois précédent) du 1er janvier 2022 au 1er août 2022 :

Premier jour du mois de	Janvier	Février	Mars	Avril	Mai	Juin	Juillet	Août
Pourcentage d'évolution		$+3\%$	$+5\%$	$+1\%$	$+3\%$	-10%	$-0,5\%$	-1%

1) Calculer le cours de l'action au 2 mars 2022.

2) Calculer le pourcentage d'augmentation du cours de l'action du premier janvier au 2 avril.

3) Le cours de l'action a-t-il progressé au bout de ces huit mois ?

Solution 7 :

1) Au 2 mars 2022, le cours de l'action est $2,45\times\left(1+\dfrac{3}{100}\right)\left(1+\dfrac{5}{100}\right)$ € ou encore $2,45\times 1,03\times 1,05$ € ou enfin $2,649\,675$ euro, soit $2,65$ € arrondi à 10^{-2}. (En bourse, les arrondis sont dangereux. Si on achète 10 actions, ce n'est pas très grave. Mais si on en achète $100\,000$ ...).

2) Le coefficient multiplicateur du 1er janvier au 2 avril est $\left(1+\dfrac{3}{100}\right)\left(1+\dfrac{5}{100}\right)\left(1+\dfrac{1}{100}\right)$ ou encore $1,05\times 1,03\times 1,01$ ou enfin

$$C_M = 1,092\,315.$$

Le taux d'augmentation est $1,092\,315-1$ ou encore

$$t = 0,092\,315.$$

Enfin, le pourcentage d'augmentation est

$$p = 100t = 9,231\,5$$

Du 1er janvier au 2 avril, le pourcentage d'augmentation est $9,231\,5\%$.

3) Soient V_i et V_f les valeurs initiale et finale du cours de l'action.

$$\begin{aligned}\frac{V_f}{V_i} &= \left(1+\frac{3}{100}\right)\left(1+\frac{5}{100}\right)\left(1+\frac{1}{100}\right)\left(1+\frac{3}{100}\right)\left(1-\frac{10}{100}\right)\left(1-\frac{0,5}{100}\right)\left(1-\frac{1}{100}\right)\\ &= 1,03\times 1,05\times 1,01\times 1,03\times 0,9\times 0,995\times 0,99\\ &= 0,997\ldots\end{aligned}$$

Donc, $\dfrac{V_f}{V_i} < 1$ puis $V_f < V_i$. L'action a (légèrement) baissé.

■

☞ *Il est très important de noter qu'**on n'additionne pas les pourcentages successifs**. Par exemple, le pourcentage d'augmentation de janvier à avril est $9{,}231\,5\%$ et n'est pas 9% $(= 3\% + 5\% + 1\%)$. On peut aussi noter que la somme des pourcentages de hausse est 12% et la somme des pourcentages de baisse est $11{,}5\%$ $(< 12\%)$ et « pourtant », le cours de l'action a baissé. Mais encore une fois, **additionner des pourcentages successifs n'a pas de signification**.*

E Evolutions réciproques

Définition 7

Une quantité Q subit une évolution qui la fait passer d'une valeur V_1 à une valeur V_2. L'**évolution réciproque** est l'évolution qui refait passer la quantité Q de la valeur V_2 à la valeur V_1.

Theoreme 10

Soient t le taux de l'évolution qui fait passer d'une valeur V_1 strictement positive à une valeur V_2 strictement positive et t' le taux de l'évolution qui fait passer de la valeur V_2 à la valeur V_1 (évolution réciproque). Alors,

$$(1+t)(1+t') = 1$$

et donc aussi

$$t' = \frac{1}{1+t} - 1.$$

Démonstration : On a $V_2 = (1+t)V_1$ et $V_1 = (1+t')V_2$. Donc, $V_1 = (1+t')V_2 = (1+t)(1+t')V_1$. Puisque $V_1 \neq 0$, on peut simplifier par V_1 et on obtient $(1+t)(1+t') = 1$.

Déterminons alors t'. Supposons par l'absurde que $1+t = 0$, alors $(1+t)(1+t') = 0$ ce qui contredit $(1+t)(1+t') = 1$. Donc, $1+t \neq 0$ puis, après division des deux membres de l'égalité $(1+t)(1+t') = 1$ par $1+t$, on obtient $1+t' = \dfrac{1}{1+t}$ et finalement, $t' = \dfrac{1}{1+t} - 1$.

■

☞ *Si on note C_M et C'_M les coefficients multiplicateurs correspondants, on a $C_M \times C'_M = 1$ ou encore, **les coefficients multiplicateurs sont inverses l'un de l'autre** ($C'_M = \dfrac{1}{C_M}$).*

Exercice 8

Une certaine quantité (strictement positive) augmente de 10%. Quel pourcentage de diminution faut-il appliquer à la nouvelle quantité pour revenir au point de départ ?

Solution 8 : Soit x une certaine quantité (strictement positive). Quand on augmente de 10%, le taux d'évolution est $t = \dfrac{10}{100} = 0{,}1$. Soit t' le taux de l'évolution réciproque. L'égalité $(1+t)(1+t') = 1$ fournit

$$t' = \frac{1}{1+t} - 1 = \frac{1}{1{,}1} - 1 = \frac{1}{1{,}1} - \frac{1{,}1}{1{,}1} = -\frac{0{,}1}{1{,}1} = -\frac{1}{11}.$$

(t' est strictement négatif car on rediminue). Le pourcentage de diminution est alors

$$p = 100 \times |t'| = \frac{100}{11} = 9{,}0909\ldots$$

Pour revenir à la quantité initiale, il faut diminuer de $\dfrac{100}{11}\%$ c'est-à-dire environ $9{,}1\%$.

■

III Paramètres statistiques

On se donne une certaine série statistique, que l'on note x. On a deux descriptions possibles de cette série. Dans la première description, on suppose que la série prend n valeurs $x_1, x_2, \ldots, x_n$ où n est un entier naturel non nul et $x_1, x_2, \ldots, x_n$, sont n réels **pas nécessairement deux à deux distincts**.

Valeurs	x_1	x_2	$\ldots$	x_n

Par exemple, un groupe de 10 élèves a obtenu les notes suivantes (sur 20) à une interrogation :

Notes	17	8	15	8	12	12	20	5	17	12

Mais on peut aussi décider de regrouper les valeurs identiques en précisant l'**effectif** pour chacune des valeurs **deux à deux distinctes** de la série. On représente alors la série dans un tableau du genre :

Valeurs	x_1	x_2	$\ldots$	x_p
Effectifs	n_1	n_2	$\ldots$	n_p
Fréquences	f_1	f_2	$\ldots$	f_p

où p est un entier naturel non nul, $x_1, \ldots, x_p$, sont des nombres réels **deux à deux distincts**, $n_1, \ldots, n_p$, sont des entiers naturels non nuls (ce sont les **effectifs partiels**).

L'**effectif total** est alors $n = n_1 + n_2 + \ldots + n_p$ et les **fréquences** sont $f_1 = \dfrac{n_1}{n}, \ldots, f_p = \dfrac{n_p}{n}$.

Par exemple, pour le groupe de 10 élèves, en regroupant et ordonnant les valeurs, on obtient le tableau :

Valeurs	5	8	12	15	17	20
Effectifs	1	2	3	1	2	1
Fréquences	0,1	0,2	0,3	0,1	0,2	0,1

On conserve les notations précédentes jusqu'à la fin (en particulier, l'effectif total sera noté n et p sera le nombre de valeurs deux à deux distinctes de la série).

A Les indicateurs de position (ou de tendance centrale)

1 La moyenne

Dans la description où la série x a n valeurs pas nécessairement deux à deux distinctes, la définition de la moyenne est :

Définition 8

La **moyenne** de la série statistique x, notée $\overline{x}$ (ou aussi m ou aussi μ (lettre de l'alphabet grec qui se lit « mu »)) est

$$\overline{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}.$$

Dans la description où la série prend p valeurs deux à deux distinctes $x_1, \ldots, x_p$, d'effectifs associés $n_1, \ldots, n_p$, la définition de la moyenne est :

Définition 9

La **moyenne** de la série statistique x, notée $\overline{x}$ (ou aussi m ou aussi μ) est

$$\overline{x} = \frac{n_1 \times x_1 + n_2 \times x_2 + \ldots + n_p \times x_p}{n_1 + n_2 + \ldots + n_p} = f_1 \times x_1 + f_2 \times x_2 + \ldots + f_p \times x_p.$$

L'égalité $\dfrac{n_1 \times x_1 + n_2 \times x_2 + \ldots + n_p \times x_p}{n_1 + n_2 + \ldots + n_p} = f_1 \times x_1 + f_2 \times x_2 + \ldots + f_p \times x_p$ peut être détaillée :

$$\begin{aligned}\frac{n_1 \times x_1 + n_2 \times x_2 + \ldots + n_p \times x_p}{n_1 + n_2 + \ldots + n_p} &= \frac{n_1 \times x_1 + n_2 \times x_2 + \ldots + n_p \times x_p}{n} = \frac{n_1}{n} x_1 + \frac{n_2}{n} x_2 + \ldots + \frac{n_p}{n} x_p \\ &= f_1 \times x_1 + f_2 \times x_2 + \ldots + f_p \times x_p.\end{aligned}$$

Quand on calcule la moyenne comme dans la définition 9, les valeurs de la série sont précédés de coefficients, les fréquences $f_1, \ldots, f_p$ ou aussi les effectifs $n_1, \ldots, n_p$. On parle alors de **moyenne pondérée**, les « poids » étant les fréquences (ou les effectifs).

Exemple. Pour le groupe de 10 élèves du paragraphe précédent :

$$\overline{x} = \frac{17+8+15+8+12+12+20+5+17+12}{10} = \frac{126}{10} = 12,6$$

ou aussi

$$\overline{x} = \frac{1\times 5+2\times 8+3\times 12+1\times 15+2\times 17+1\times 20}{10} = \frac{5+16+36+15+34+20}{10} = \frac{126}{10} = 12,6.$$

Le résultat qui suit s'appelle « la **linéarité de la moyenne** ».

Theoreme 11

Soient a et b deux réels. La moyenne de la série $ax+b$ est $a\overline{x}+b$.

Démonstration :

• Soit $y = ax$. La moyenne de y est

$$\overline{y} = \frac{n_1\times(ax_1)+n_2\times(ax_2)+\ldots+n_p\left(ax_p\right)}{n_1+n_2+\ldots+n_p} = a\times\frac{n_1\times x_1+n_2\times x_2+\ldots+n_p\times x_p}{n_1+n_2+\ldots+n_p} = a\times\overline{x}.$$

• Soit $y = x+b$. La moyenne de y est

$$\begin{aligned}\overline{y} &= \frac{n_1\times(x_1+b)+n_2\times(x_2+b)+\ldots+n_p\left(x_p+b\right)}{n_1+n_2+\ldots+n_p} = \frac{\left(n_1\times x_1+n_2\times x_2+\ldots+n_p\times x_p\right)+\left(n_1\times b+n_2\times b+\ldots+n_p\times b\right)}{n_1+n_2+\ldots+n_p}\\ &= \frac{\left(n_1\times x_1+n_2\times x_2+\ldots+n_p\times x_p\right)+\left(n_1+n_2+\ldots+n_p\right)b}{n_1+n_2+\ldots+n_p} = \frac{n_1\times x_1+n_2\times x_2+\ldots+n_p\times x_p}{n_1+n_2+\ldots+n_p}+\frac{\left(n_1+n_2+\ldots+n_p\right)b}{n_1+n_2+\ldots+n_p}\\ &= \overline{x}+b.\end{aligned}$$

• Soient $y = ax$ puis $z = y+b = ax+b$. En cumulant les deux résultats précédents, on obtient

$$\overline{z} = \overline{y}+b = a\overline{x}+b.$$

■

Ainsi, si on ajoute 1 à chacune des notes d'une classe, la moyenne augmente de 1 et si on multiplie par 2, chacune des notes d'une classe, la moyenne est multipliée par 2.

2 La médiane

Pour définir correctement la médiane, on adopte la présentation où la série statistique x a n valeurs **pas nécessairement deux à deux distinctes** $x_1, x_2, \ldots, x_n$, **classées dans l'ordre croissant** $(x_1 \leqslant x_2 \leqslant \ldots \leqslant x_n)$.

Définition 10

Si n est impair, on pose $n = 2p+1$ où $p \in \mathbb{N}$. La **médiane** est x_{p+1} (ou encore $x_{\frac{n+1}{2}}$).
Si n est pair, on pose $n = 2p$ où $p \in \mathbb{N}^*$. La **médiane** est la moyenne des deux nombres x_p et x_{p+1} c'est-à-dire $\dfrac{x_p+x_{p+1}}{2}$ (ou encore $\dfrac{1}{2}\left(x_{\frac{n}{2}}+x_{\frac{n+1}{2}}\right)$). Ce n'est pas forcément une valeur de la série statistique.
Dans tous les cas, on notera $\mathrm{M}e$ la médiane.

Theoreme 12

Au moins 50% des valeurs de la série x sont inférieures ou égales à $\mathrm{M}e$ et au moins 50% des valeurs de la série x sont supérieures ou égales à $\mathrm{M}e$.

Démonstration :

• Si n est impair, on pose $n = 2k+1$ où $k \in \mathbb{N}$. La médiane est $\mathrm{M}e = x_{k+1}$.
De x_1 à x_k, il y a k valeurs (et de x_1 à x_{2k+1}, il y a $2k+1$ valeurs). Donc, de x_{k+2} à x_{2k+1}, il y a $(2k+1)-(k+1)$ valeurs ou encore il y a k valeurs.

$$\underbrace{x_1 \quad \dots \quad x_k}_{k \text{ valeurs}} \qquad x_{k+1} \qquad \underbrace{x_{k+2} \quad \dots \quad x_{2k+1}}_{k \text{ valeurs}}$$

Il y a $k+1$ valeurs (au moins) de la série x qui sont inférieures ou égales à la médiane $\mathrm{M}e = x_{k+1}$ (à savoir $x_1, \dots, x_k, x_{k+1}$) et il y a $k+1$ valeurs de la série x qui sont supérieures ou égales à la médiane $\mathrm{M}e = x_{k+1}$ (à savoir $x_{k+1}, x_{k+2}, \dots, x_{2k+1}$).
La proportion de valeurs de la série x qui sont inférieures ou égales à $\mathrm{M}e$ est au moins $\dfrac{k+1}{2k+1}$ et il y a (au moins) la même proportion de valeurs de la série x qui sont supérieures ou égales à $\mathrm{M}e$. Ensuite, $2k+1 \leqslant 2k+2$ puis $\dfrac{1}{2k+1} \geqslant \dfrac{1}{2k+2}$ puis $\dfrac{k+1}{2k+1} \geqslant \dfrac{k+1}{2k+2}$ ou encore $\dfrac{k+1}{2k+1} \geqslant \dfrac{k+1}{2(k+1)}$ ou enfin $\dfrac{k+1}{2k+1} \geqslant \dfrac{1}{2}$.
Ainsi, dans le cas où n est impair, au moins la moitié des valeurs sont inférieures ou égales à la médiane et au moins la moitié des valeurs sont supérieures ou égales à la médiane.

• Si n est pair, on pose $n = 2k$ où $k \in \mathbb{N}^*$. La médiane est $\mathrm{M}e = \dfrac{x_k + x_{k+1}}{2}$. On note que $x_k \leqslant \dfrac{x_k + x_{k+1}}{2} \leqslant x_{k+1}$.
De x_1 à x_k, il y a k valeurs. Donc, de x_{k+1} à x_{2k}, il y a $(2k) - (k)$ valeurs ou encore il y a k valeurs.

$$\underbrace{x_1 \quad \dots \quad x_k}_{k \text{ valeurs}} \qquad \tfrac{x_k + x_{k+1}}{2} \qquad \underbrace{x_{k+1} \quad \dots \quad x_{2k}}_{k \text{ valeurs}}$$

Il y a k valeurs (au moins) de la série x qui sont inférieures ou égales à la médiane $\mathrm{M}e = \dfrac{x_k + x_{k+1}}{2}$ (à savoir $x_1, \dots, x_k$) et il y a k valeurs de la série x qui sont supérieures ou égales à la médiane $\mathrm{M}e = \dfrac{x_k + x_{k+1}}{2}$ (à savoir $x_{k+1}, \dots, x_{k+2}, x_{2k}$).
La proportion de valeurs de la série x qui sont inférieures (ou supérieures) ou égales à $\mathrm{M}e$ est cette fois-ci au moins $\dfrac{p}{2p}$ ou encore au moins $\dfrac{1}{2}$.
Ainsi, dans le cas où n est pair, au moins la moitié des valeurs sont inférieures ou égales à la médiane et au moins la moitié des valeurs sont supérieures ou égales à la médiane.

■

Exemple. Pour la série $1, 2, 2, 5, 5, 7, 7$, la médiane est $\mathrm{M}e = 5$ et pour la série $1, 2, 2, 5, 5, 7$, la médiane est $\mathrm{M}e = \dfrac{2+5}{2}$ ou encore $\mathrm{M}e = 3{,}5$. De même, pour la série de 10 notes

Notes	5	8	8	12	12	12	15	17	17	20

la médiane est $\mathrm{M}e = \dfrac{12+12}{2}$ c'est-à-dire $\mathrm{M}e = 12$. On rappelle que la moyenne de cette série est $\overline{x} = 12{,}6$. Si on modifie les deux premières notes :

Notes	1	1	8	12	12	12	15	17	17	20

la moyenne change ($\overline{x} = 11{,}5$) mais la médiane reste inchangée. On dit alors que **la médiane n'est pas sensible aux valeurs extrêmes**.

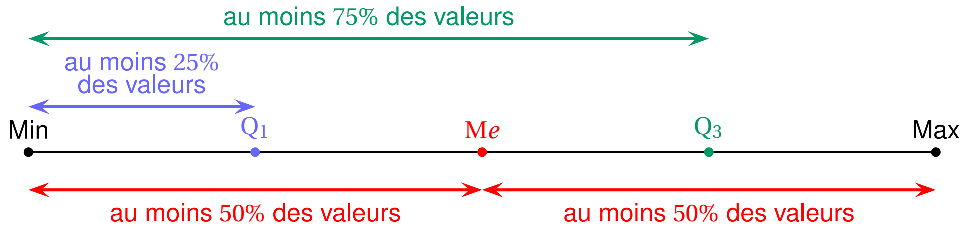
3 Les quartiles

Maintenant, on va partager la série statistique x en quatre paquets. De nouveau, on adopte la présentation où la série statistique x a n valeurs **pas nécessairement deux à deux distinctes** $x_1, x_2, \dots, x_n$, **classées dans l'ordre croissant** ($x_1 \leqslant x_2 \leqslant \dots \leqslant x_n$).

Définition 11

Le **premier quartile**, noté Q_1, est la première valeur (en partant de x_1) de la série telle que au moins 25% des valeurs de la série lui soient inférieures ou égales.
Le **troisième quartile**, noté Q_3, est la première valeur (en partant de x_1) de la série telle que au moins 75% des valeurs de la série lui soient inférieures ou égales.

On a donc le graphique suivant (on pose $x_1 = \mathrm{M}in$ et $x_n = \mathrm{M}ax$. x_1 et x_n sont respectivement la plus petite valeur et la plus grande valeur de la série) :



Reprenons la série de 10 notes :

Notes	5	8	8	12	12	12	15	17	17	20

25% de 10 est $2,5$. Le premier quartile est la troisième valeur (3 est le premier numéro dépassant $2,5$) ou encore $Q_1 = 3$. De même, 75% de 10 est $7,5$. Le troisième quartile est la huitième valeur (8 est le premier numéro dépassant $7,5$) ou encore $Q_3 = 17$.

Analysons le cas général.

• Si n est un multiple de 4, on peut poser $n = 4k$ où k est un entier naturel non nul. Le premier quartile est dans ce cas $Q_1 = x_k$ et le troisième quartile est $Q_3 = x_{3k}$.

• Si n est 1 de plus qu'un multiple de 4 et est donc de la forme $n = 4k+1$ avec $k \in \mathbb{N}$. Dans ce cas, le premier quartile est $Q_1 = x_{k+1}$ et le troisième quartile est $Q_3 = x_{3k+1}$.

• Si n est 2 de plus qu'un multiple de 4 et est donc de la forme $n = 4k+2$ avec $k \in \mathbb{N}$. Dans ce cas, le premier quartile est $Q_1 = x_{k+1}$ et le troisième quartile est $Q_3 = x_{3k+2}$.

• Si n est 3 de plus qu'un multiple de 4 et est donc de la forme $n = 4k+3$ avec $k \in \mathbb{N}$. Dans ce cas, le premier quartile est $Q_1 = x_{k+1}$ et le troisième quartile est $Q_3 = x_{3k+3}$.

B Les indicateurs de dispersion

La moyenne d'une série statistique est un **indicateur de position** (ou encore un **indicateur de tendance centrale**). Obtenir une moyenne de classe de $10/20$ à une interrogation donne une indication sur le comportement de la classe qui n'est pas la même que si la moyenne de classe est de $15/20$. Dans le deuxième cas, un(e) professeur(e) dira qu'il a une très bonne classe.

Mais si on sait que la moyenne est de $10/20$, on ne sait pas si les notes sont regroupées autour de la moyenne (par exemple, cinq notes égales à $9/20$ et cinq notes égales à $11/20$) ou bien s'il y a beaucoup d'écart entre les notes (par exemple, cinq notes égales à $0/20$ et cinq notes égales à $20/20$). Dans le premier cas, un(e) professeur(e) dira qu'il (ou elle) a une classe homogène et dans le deuxième cas, une classe hétérogène. Il nous manque un **indicateur de dispersion**.

1 Variance et écart-type

Dans la description où la série x a n valeurs **pas nécessairement deux à deux distinctes** $x_1, \ldots, x_n$:

Définition 12

La **variance** de la série statistique, notée V, est la moyenne des carrés des écarts à la moyenne c'est-à-dire

$$V = \frac{1}{n}\left(\left(x_1 - \overline{x}\right)^2 + \ldots + \left(x_n - \overline{x}\right)^2\right)$$

L'**écart-type** de la série, noté s ou σ (lettre s de l'alphabet grec qui se lit « sigma ») est la racine carrée de la variance :

$$s = \sqrt{V} = \sqrt{\frac{1}{n}\left(\left(x_1 - \overline{x}\right)^2 + \ldots + \left(x_n - \overline{x}\right)^2\right)}.$$

Dans la description où la série x a p valeurs **deux à deux distinctes**, d'effectifs respectifs $n_1, \ldots, n_p$:

Définition 13

La **variance** de la série statistique, notée V, est la moyenne des carrés des écarts à la moyenne c'est-à-dire

$$V = \frac{1}{n}\left(n_1 \times \left(x_1 - \overline{x}\right)^2 + \ldots + n_p\left(x_p - \overline{x}\right)^2\right) = f_1 \times \left(x_1 - \overline{x}\right)^2 + \ldots + f_p\left(x_p - \overline{x}\right)^2.$$

L'**écart-type** de la série, noté s ou σ est la racine carrée de la variance :

$$s = \sqrt{V} = \sqrt{\frac{1}{n}\left(n_1 \times \left(x_1 - \overline{x}\right)^2 + \ldots + n_p\left(x_p - \overline{x}\right)^2\right)} = \sqrt{f_1 \times \left(x_1 - \overline{x}\right)^2 + \ldots + f_p\left(x_p - \overline{x}\right)^2}.$$

Par exemple, en reprenant la liste des 10 notes, la variance est

$$\begin{aligned} V &= \frac{1}{10}\left(1 \times (5-12,6)^2 + 2 \times (8-12,6)^2 + 3 \times (12-12,6)^2 + 1 \times (15-12,6)^2 + 2 \times (17-12,6)^2 + 1 \times (20-12,6)^2\right) \\ &= 20,04 \end{aligned}$$

puis l'écart-type de la série de notes est

$$s = \sqrt{V} = \sqrt{20,04} = 4,4\ldots$$

L'écart-type indique si les valeurs de la série sont plutôt regroupées autour de la moyenne ou écartées de la moyenne. A titre de comparaison, si les dix notes sont composées de cinq $9/20$ et de cinq $11/20$, l'écart-type est $s = 1$ et si les dix notes sont composées de cinq $0/20$ et de cinq $20/20$, l'écart-type est $s = 10$ (pour une moyenne qui est dans les deux cas $m = 10$). Obtenir un écart-type petit par rapport à la moyenne signifie que les valeurs de la série sont regroupées et obtenir un écart-type grand par rapport à la moyenne signifie que les valeurs de la série sont dispersées.

2 L'écart interquartile

Définition 14

L'**écart interquartile** de la série est la distance entre le premier et le troisième quartile c'est-à-dire $Q_3 - Q_1$.

Par exemple, pour la liste de notes du groupe de dix élèves, nous avions obtenu $Q_1 = 8$ et $Q_3 = 17$. L'écart interquartile est donc $Q_3 - Q_1 = 9$.

On termine ce chapitre par un exercice qui permet de revoir tous les paramètres statistiques.

Exercice 9

Le tableau suivant donne l'espérance de vie en 2020, à l'unité près par défaut, dans les différents pays de l'union européenne (d'après un document de l'Ined (Institut national d'études démographiques)).

Pays	Allemagne	Autriche	Belgique	Bulgarie	Chypre	Croatie	Danemark
Espérance de vie	81	82	82	75	82	78	81
Pays	Espagne	Estonie	Finlande	France	Grèce	Hongrie	Irlande
Espérance de vie	84	79	82	83	81	76	82
Pays	Italie	Lettonie	Lituanie	Luxembourg	Malte	Pays-Bas	Pologne
Espérance de vie	83	75	76	82	82	82	78
Pays	Portugal	Roumanie	Slovaquie	Slovénie	Suède	Tchéquie	
Espérance de vie	81	75	77	81	83	79	

1) Construire le tableau des effectifs et des fréquences (on donnera les arrondis à 10^{-2}).

2) Représenter la série par un diagramme bâtons (donnant les effectifs).

3) Déterminer l'espérance de vie moyenne m dans l'union européenne en 2020. On arrondira à 10^{-1}.

4) Déterminer l'écart-type s de la série. On arrondira à 10^{-1}.

5) Déterminer la médiane de la série.

5) Déterminer l'écart interquartile de la série.

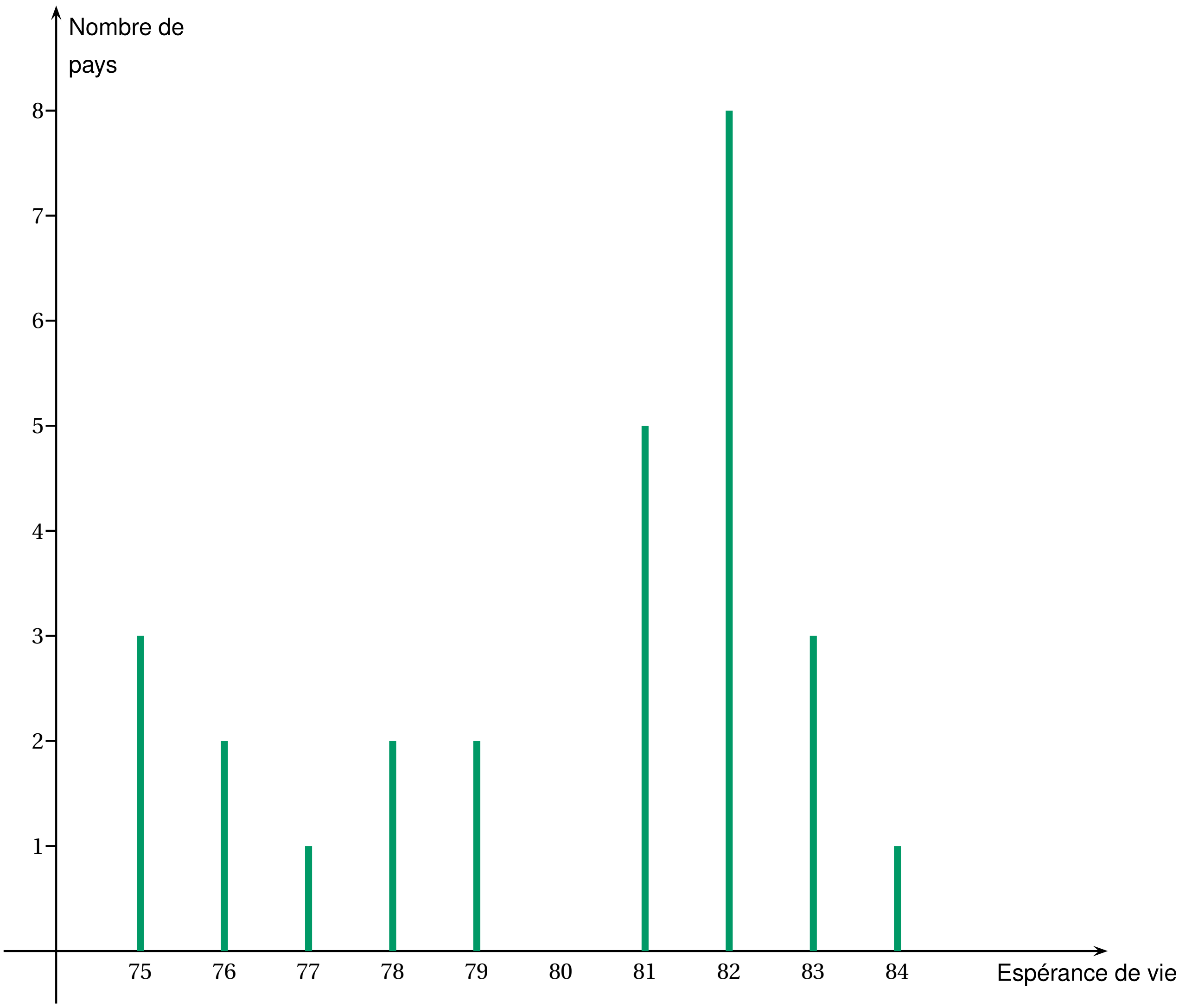
Solution 9 :

1) Le tableau des effectifs et des fréquences arrondies à 10^{-2} est :

Espérance de vie	75	76	77	78	79	81	82	83	84
Nombre de pays	3	2	1	2	2	5	8	3	1
Fréquences	0,11	0,07	0,04	0,07	0,07	0,19	0,30	0,11	0,04

(On note que la somme des effectifs est égale à 27 et la somme des fréquences est égale à 1.)

2) Diagramme bâtons.



3) L'espérance de vie moyenne dans l'union européenne en 2020, arrondie à 10^{-1}, était :

$$m = \frac{3\times 75+2\times 76+1\times 77+2\times 78+2\times 79+5\times 81+8\times 82+3\times 83+1\times 84}{27} = \frac{2\,162}{27}$$
$$= 80,1 \text{ arrondie à } 10^{-1}.$$

4) a) La variance est

$$V = \frac{1}{27}\left(3\times\left(75-\frac{2\,162}{27}\right)^2+2\times\left(76-\frac{2\,162}{27}\right)^2+1\times\left(77-\frac{2\,162}{27}\right)^2+2\times\left(78-\frac{2\,162}{27}\right)^2+2\times\left(79-\frac{2\,162}{27}\right)^2\right.$$
$$\left.+5\times\left(81-\frac{2\,162}{27}\right)^2+8\times\left(82-\frac{2\,162}{27}\right)^2+3\times\left(83-\frac{2\,162}{27}\right)^2+1\times\left(84-\frac{2\,162}{27}\right)^2\right)$$
$$= 7,624\,142\ldots.$$

L'écart-type est donc

$$s = \sqrt{V} = \sqrt{7,624\,142\ldots} = 2,8 \text{ arrondi à } 10^{-1}.$$

5) 27 est impair. Plus précisément, $27 = 13 + 1 + 13$. La médiane de la série est la 14-ème valeur de la série. Les 14 premières valeurs de la série sont

$$75 \quad 75 \quad 75 \quad 76 \quad 76 \quad 77 \quad 78 \quad 78 \quad 79 \quad 79 \quad 81 \quad 81 \quad 81 \quad 81\ .$$

La médiane de la série est $\mathrm{M}e = 81$.

5) 25% de 27 est $6,75$. Le premier numéro qui dépasse $6,75$ est 7. Le premier quartile est la 7ème valeur ou encore $Q_1 = 78$.
75% de 27 est $20,25$. Le premier numéro qui dépasse $20,25$ est 21. Le troisième quartile est la 21ème valeur ou encore $Q_3 = 82$.

L'écart interquartile est $Q_3 - Q_1 = 4$.

■