

TD SEMAINE 1 : SYSTÈMES LINÉAIRES

Considérons un système d'équations :

$$(S) : \begin{cases} 2x - 5y = 1 \\ x - 2y = -3 \end{cases}$$

On cherche à résoudre ce système, c'est-à-dire à trouver les valeurs de x et y qui satisfont ces deux équations simultanément.

RÉSOLUTION

La méthode par *substitution* est la plus simple, et est très efficace pour un système de deux équations à deux inconnues. On va utiliser une des deux équations pour exprimer une variable en fonction de l'autre. Dans cet exemple, la seconde équation peut être utilisée pour exprimer x en fonction de y :

$$x - 2y = -3 \quad \Leftrightarrow \quad x = -3 + 2y$$

Le système initial est donc équivalent à :

$$(S) \Leftrightarrow \begin{cases} 2x - 5y = 1 \\ x = -3 + 2y \end{cases}$$

On peut ensuite remplacer x par cette fonction de y dans la première équation (attention aux parenthèses!), et simplifier.

$$(S) \Leftrightarrow \begin{cases} 2(-3 + 2y) - 5y = 1 \\ x = -3 + 2y \end{cases} \Leftrightarrow \begin{cases} -6 + 4y - 5y = 1 \\ x = -3 + 2y \end{cases} \Leftrightarrow \begin{cases} -6 - y = 1 \\ x = -3 + 2y \end{cases}$$

La première équation ne fait alors intervenir que la variable y , et peut donc être résolue.

$$(S) \Leftrightarrow \begin{cases} -y = 7 \\ x = -3 + 2y \end{cases} \Leftrightarrow \begin{cases} y = -7 \\ x = -3 + 2y \end{cases}$$

On connaît donc la valeur de y , et on a exprimé x en fonction de y . On peut alors déterminer la valeur de x :

$$(S) \Leftrightarrow \begin{cases} -y = 7 \\ x = -3 + 2(-7) \end{cases} \Leftrightarrow \begin{cases} y = -7 \\ x = 11 \end{cases}$$

Le système a donc pour unique solution $x = 11$ et $y = -7$.

POUR ALLER PLUS LOIN

Tous les systèmes d'équations n'ont pas une unique solution. Certains systèmes ont une infinité de solution, par exemple :

$$\begin{cases} x + y = 1 \\ x + y = 1 \end{cases}$$

D'autres n'ont aucune solution :

$$\begin{cases} x + y = 1 \\ x + y = 2 \end{cases}$$

Ces cas particuliers apparaissent clairement lors de la résolution (on tombe des égalités triviales du type $0 = 0$, ou toujours fausses, du type $1 = 0$).

La méthode de substitution peut être utilisée pour des systèmes de trois équations à trois inconnues (ou plus), comme par exemple :

$$\begin{cases} 2x - 5y + 3z = 1 \\ x - y - z = 1 \\ x + y + z = 6 \end{cases}$$

Dans ce cas, on utilise une équation pour exprimer z en fonction de x et y . On substitue z par cette expression dans les deux autres équations. On retrouve alors un système de deux équations à deux inconnues, que l'on peut encore résoudre par substitution. Cependant, les calculs deviennent très vite trop compliqués pour être fait à la main ; il faut alors utiliser une calculatrice ou un ordinateur.

D'autres méthodes de résolution existent, comme la méthode de *combinaison linéaire*. Elle ne sera pas abordée dans ce cours, mais si vous la maîtrisez déjà, n'hésitez pas à l'utiliser !

TD SEMAINE 2 : MODÉLISATION PAR DES SYSTÈMES LINÉAIRES

L'objectif de la semaine est de travailler la modélisation, c'est-à-dire la capacité à traduire un problème concret en problème mathématique, dans le cadre des systèmes d'équations.

Considérons l'énoncé suivant :

Un bar offre de la bière blonde à 8€ le litre, et de la bière ambrée à 12€ le litre.
Un groupe d'amis consomme au total 2,5 litres de bière, et a payé 23€. Combien ont-ils consommé de chaque type de bière ?

QUE CHERCHE-T-ON ?

La première question à se poser est la suivante : que veut-on savoir ? Ici, ce sont la quantité de bière blonde consommée, et la quantité de bière ambrée. On va leur donner un nom :

- a : volume de bière ambrée consommée, en litres.
- b : volume de bière blonde consommée, en litres.

a et b seront nos inconnues. Si possible, il est utile de nommer l'inconnue de façon à faire le lien avec l'objet qu'elle désigne ; ici, on a pris les initiales d'"ambrée" et de "blonde". Cela évitera de les confondre par la suite.

QUE SAIT-ON ?

L'énoncé introduit deux types de grandeurs : des volumes et des prix. Ces grandeurs nous apportent des informations sur a et b . Je conseille de traiter chacune de ces grandeurs séparément, à l'aide d'identités comptables. Par exemple :

$$\text{Volume de bière} = \text{Volume de bière ambrée} + \text{Volume de bière blonde}$$

Afin de clarifier votre raisonnement, il peut être *très* utile d'écrire explicitement de telles égalités. Ensuite, on remplace chaque terme ou bien par des données de l'énoncé, ou bien par une expression dépendant des inconnues a et b . Ici :

- Volume de bière : 2,5 litres.
- Volume de bière ambrée : a litres.
- Volume de bière blonde : b litres.

L'équation précédente devient donc :

$$a + b = 2,5$$

On fait de même avec les prix.

$$\text{Coût total} = \text{Coût de la bière ambrée} + \text{Coût de bière blonde}$$

De plus,

- Coût total : 23€.
- Coût de la bière ambrée : $12a$ € (on a a litres, à 12 euros le litre).
- Coût de bière blonde : $8b$ € (on a b litres, à 8 euros le litre).

Par conséquent :

$$12a + 8b = 23$$

Ainsi, a et b sont solutions du système d'équations :

$$\begin{cases} a + b = 2,5 \\ 12a + 8b = 23 \end{cases}$$

CONCLUSION

Il reste encore à résoudre ce système; en utilisant les méthodes de la semaine passée, on trouve $a = 0,75$ et $b = 1,75$. Enfin, on peut conclure : le groupe a consommé 0,75 litres de bière ambrée et 1,75 litres de bière blonde.

TD SEMAINE 2, BIS : RÉOLUTION DE SYSTÈME 3×3

Dans la première feuille de TD, certains systèmes de 3 équations à 3 inconnues apparaissent. Voici un exemple de résolution de tels systèmes. Au vu du temps nécessaire et du risque d'erreur, **il ne sera pas exigé de résoudre de tels systèmes.**

Le système que nous allons résoudre est le suivant (TD1, exercice facultatif) :

$$\begin{cases} 2x - 3y + 5z = 110 \\ 5x - 2y + 4z = 130 \\ 6x + 3y - 4z = 0 \end{cases}$$

Nous allons procéder par substitution. La méthode de combinaison linéaire peut aussi s'appliquer à cette situation (et est souvent plus efficace).

Nous allons donc choisir une première variable, que nous allons exprimer en fonction des deux autres. Par exemple, nous pouvons utiliser la première équation pour exprimer x en fonction de y et z :

$$\begin{aligned} 2x &= 110 + 3y - 5z \\ x &= 55 + \frac{3}{2}y - \frac{5}{2}z \end{aligned}$$

Nous allons ensuite remplacer x par cette expression dans les deux équations suivantes. La seconde équation devient :

$$\begin{aligned} 5 \left(55 + \frac{3}{2}y - \frac{5}{2}z \right) - 2y + 4z &= 130 \\ 275 + \frac{15}{2}y - \frac{25}{2}z - 2y + 4z &= 130 \\ \frac{11}{2}y - \frac{17}{2}z &= -145 \\ 11y - 17z &= -290, \end{aligned}$$

et la troisième :

$$\begin{aligned} 6 \left(55 + \frac{3}{2}y - \frac{5}{2}z \right) + 3y - 4z &= 0 \\ 330 + 9y - 15z + 3y - 4z &= 0 \\ 12y - 19z &= -330. \end{aligned}$$

Pour trouver y et z , il nous reste à résoudre le système d'équations :

$$\begin{cases} 11y - 17z = -290 \\ 12y - 19z = -330 \end{cases}$$

Là encore, on procède par substitution. En utilisant la première équation, on trouve :

$$\begin{aligned} 11y &= 17z - 290 \\ y &= \frac{17}{11}z - \frac{290}{11} \end{aligned}$$

En remplaçant y par cette expression dans la seconde équation, on trouve :

$$\begin{aligned}12 \left(\frac{17}{11}z - \frac{290}{11} \right) - 19z &= -330 \\ \frac{204}{11}z - \frac{3480}{11} - 19z &= -330 \\ 204z - 3480 - 209z &= -3630 \\ -5z &= -150 \\ z &= 30\end{aligned}$$

De là, on retrouve y et x :

$$\begin{aligned}y &= \frac{17}{11}z - \frac{290}{11} = \frac{17}{11} \times 30 - \frac{290}{11} = \frac{510 - 290}{11} = \frac{220}{11} = 20, \\ x &= 55 + \frac{3}{2}y - \frac{5}{2}z = 55 + \frac{3}{2} \times 20 - \frac{5}{2} \times 30 = 55 + 30 - 75 = 10.\end{aligned}$$

Le système a donc pour (unique) solution $x = 10$, $y = 20$ et $z = 30$.

TD SEMAINE 3 : ÉVOLUTION À TAUX CONSTANT

L'objectif de la semaine est de modéliser des phénomènes en faisant appel à des évolutions à taux constant.

Considérons l'énoncé suivant :

Une entreprise en ligne a 1 200 clients réguliers au 1er janvier 2020. Son nombre de clients augmente de 20% par an. Combien aura-t-elle de clients au 1er janvier 2023 ? Au 1er mars 2024 ? Combien avait-elle de clients au 1er juillet 2019 ?

En première approximation, on peut supposer que le nombre de clients suit une évolution à taux constant, au taux de 20% par **an**. On note C_n le nombre de clients dans n **années** (attention : il faut que les unités restent cohérentes !) après le 1er janvier 2020, c'est-à-dire au 1er janvier $2020+n$. Comme ce nombre de clients suit une évolution à taux constant,

$$(0.1) \quad C_n = C_0(1+t)^n.$$

D'après les données du problème, $C_0 = 1\,200$ et $t = 20\% = \frac{20}{100} = 0,2$. Par conséquent,

$$C_n = 1\,200(1+0,2)^n = 1\,200 \times 1,2^n.$$

Du 01/01/2020 au 01/01/2023, il se sera écoulé exactement 3 ans. Si la croissance de l'entreprise continue au même rythme, le nombre de clients au 01/01/2023 sera donc de

$$C_3 = 1\,200 \times 1,2^3 = 1\,200 \times 1,728 \simeq 2\,074.$$

Comme vu dans le cours, l'équation (0.1) reste valable pour des valeurs non entières de n . Du 01/01/2020 au 01/03/2024, il se sera écoulé exactement 4 ans et 2 mois. On convertit cette durée en années : comme 1 mois vaut $\frac{1}{12}$ ans, cette durée est de $4 + \frac{2}{12} \simeq 4,167$ ans. Le nombre de clients au 01/03/2024 sera donc de

$$C_{4,167} = 1\,200 \times 1,2^{4,167} \simeq 1\,200 \times 2,138 \simeq 2\,565.$$

Comme vu dans le cours, l'équation (0.1) reste aussi valable pour des valeurs négative de n . Cela correspond à des dates dans le passé, ici antérieures au 01/01/2020. Du 01/07/2019 au 01/01/2020, il s'est écoulé exactement 6 mois ; la première date est donc 6 mois = 0,5 an dans le passé. Le nombre de clients au 01/07/2019 était donc de

$$C_{-0,5} = 1\,200 \times 1,2^{-0,5} \simeq 1\,200 \times 0,913 \simeq 1\,095.$$

TD SEMAINE 4 : RECHERCHE ET CONVERSION DE TAUX

Les objectifs de la semaine sont de calculer et convertir des taux de croissance.

CALCUL DE TAUX

Considérons l'énoncé suivant :

Une entreprise a 2300 clients réguliers au 1er janvier 2015, et 1200 clients au 1er janvier 2020. Quel est son taux de croissance annuel moyen durant cette période ?

On suppose que le nombre de clients de l'entreprise suit une évolution à taux **annuel** constant t sur cette période. Soit C_n le nombre de clients n **années** après le 01/01/2015. Alors :

$$C_n = C_0(1 + t)^n.$$

On connaît $C_0 = 2300$, ainsi que $C_5 = 1200$ pour $n = 5$ (il s'est écoulé exactement 5 ans entre le 01/01/2015 et le 01/01/2020). Par conséquent,

$$1200 = 2300(1 + t)^5.$$

On cherche le taux t . Pour cela, on isole tout d'abord la quantité $(1 + t)^5$ pour se ramener à une équation puissance :

$$(1 + t)^5 = \frac{1200}{2300} \simeq 0,522.$$

Ensuite, on élève les deux membres de l'équation à la puissance $1/5$:

$$1 + t = ((1 + t)^5)^{\frac{1}{5}} = \left(\frac{1200}{2300}\right)^{\frac{1}{5}} \simeq 0,878.$$

Enfin, en soustrayant 1 de chaque côté, on trouve $t \simeq 0,878 - 1 = -0,122 = -12,2\%$. L'entreprise a donc un taux de croissance de $-12,2\%$ par an ; autrement dit, elle perd $12,2\%$ de sa clientèle chaque année.

CONVERSION DE TAUX

Ce type de calcul peut être utilisé pour convertir des taux annuels en taux mensuels, ou vice-versa.

Le chiffre d'affaires de l'entreprise A croît à un taux mensuel de -10% . Quel est son taux de croissance annuel ?

Le chiffre d'affaires de l'entreprise B croît à un taux annuel de -10% . Quel est son taux de croissance mensuel ?

Soient t_a et $t_m = -10\% = -0,1$ les taux de croissance respectivement annuel et mensuel du chiffre d'affaires de l'entreprise A . Après 1 an = 12 mois, le coefficient multiplicateur est de $1 + t_a = (1 + t_m)^{12}$. On trouve donc

$$1 + t_a = (1 - 0,1)^{12} = 0,9^{12} \simeq 0,28,$$

et donc $t_a \simeq 0,28 - 1 = -0,72 = -72\%$. L'entreprise A a un taux de croissance annuel de -72% .

Remarquons au passage que $t_a \neq 12 \times t_m = -120\%$.

Soient $s_a = -10\% = -0,1$ et s_m les taux de croissance respectivement annuel et mensuel du chiffre d'affaires de l'entreprise B . Après 1 an = 12 mois, le coefficient multiplicateur est de $1 + t_a = (1 + t_m)^{12}$. On trouve donc

$$(1 + t_m)^{12} = 1 + t_a = 1 - 0,1 = 0,9,$$

d'où :

$$1 + t_m = (1 + t_a)^{\frac{1}{12}} = 0,9^{\frac{1}{12}} \simeq 0,9913.$$

Par conséquent, $t_m \simeq 0,9913 - 1 = -0,0087 = -0,87\%$. L'entreprise B a un taux de croissance mensuel de $-0,87\%$.

On peut aussi faire des conversions entre taux mensuel et trimestriel (en utilisant $(1 + t_t) = (1 + t_m)^3$), entre taux trimestriel et annuel (en utilisant $(1 + t_a) = (1 + t_t)^4$), entre taux hebdomadaire et annuel (en utilisant $(1 + t_a) = (1 + t_h)^{52}$)...

TD SEMAINE 5 : CALCUL DE DURÉES

L'objectif de la semaine est de calculer la durée pour atteindre un certain seuil dans des évolutions à taux constant.

Considérons l'énoncé suivant :

Au 1er janvier 2010, une ville a 135 000 habitants. Sa population croît au taux moyen de 2% par an. À quelle date atteindra-t-elle les 200 000 habitants ?

MODÉLISATION

On modélise la population de la ville comme suivant une évolution au taux constant de $t = 2\% = 0,02$ par an. Soit P_n la population n années après le 01/01/2010. Alors :

$$P_n = P_0(1 + t)^n = 135\,000 \times 1,02^n.$$

On cherche n tel que $P_n = 200\,000$, c'est-à-dire :

$$135\,000 \times 1,02^n = 200\,000.$$

RÉSOLUTION

Comme pour les recherches de taux, nous allons d'abord isoler le terme $1,02^n$. Ensuite, nous appliquerons le logarithme de chaque côté de l'équation pour faire descendre l'exposant.

$$135\,000 \times 1,02^n = 200\,000$$

$$1,02^n = \frac{200\,000}{135\,000} \simeq 1,48$$

$$\log(1,02^n) \simeq \log(1,48)$$

$$n \log(1,02) \simeq \log(1,48)$$

$$n \simeq \frac{\log(1,48)}{\log(1,02)} \simeq 19,85$$

La population de la ville atteindra les 200 000 habitants environ 19,85 ans = 19 ans et $0,85 \times 12 \simeq 10$ mois après le 01/01/2010, c'est-à-dire en novembre 2029.

TD SEMAINE 6 : ACTUALISATION

L'objectif de la semaine est d'utiliser les évolutions à taux constant pour actualiser des comptes financiers.

Considérons l'énoncé suivant :

Au 1er janvier 2010, M. Gates a 10 000 euros sur un compte d'épargne qui lui garantit 5% d'intérêts par an. Le 1er mai 2014, il dépose 5 000 euros supplémentaires sur ce compte ; le 1er septembre 2017, il en retire 1 000 euros.

Combien a-t-il d'argent sur ce compte au 1er janvier 2020 ?

Nous allons détailler deux méthodes : en actualisant le compte à chaque nouvelle opération, ou en calculant le montant actualisé au 1er janvier 2020 correspondant à chaque opération.

PREMIÈRE MÉTHODE

Chaque année, l'argent sur le compte est multiplié par $1 + 5\% = 1 + 0,05 = 1,05$. Si l'on attend une période de n années, alors cet argent est multiplié par $1,05^n$. On actualise à chaque nouvelle opération.

Date	Durée depuis l'opération précédente	Dépôt /retrait (€)	Capital total épargné (€)
01/01/2010	-	10 000	10 000
01/05/2014	4 ans 4 mois \simeq 4,33 ans	5 000	$10\,000 \times 1,05^{4,33} + 5\,000 \simeq 17\,354,36$
01/09/2017	3 ans 4 mois \simeq 3,33 ans	-1 000	$17\,354,36 \times 1,05^{3,33} - 1\,000 \simeq 19\,419,24$
01/01/2020	2 ans et 4 mois \simeq 2,33 ans	-	$19\,419,24 \times 1,05^{2,33} \simeq 21\,760,76$

Au 1er janvier 2020, le capital épargné est de 21 760,76€.

PREMIÈRE MÉTHODE

Chaque année, l'argent sur le compte est multiplié par $1 + 5\% = 1 + 0,05 = 1,05$. Si l'on attend une période de n années, alors cet argent est multiplié par $1,05^n$. On traite chaque somme déposée parallèlement, et actualise au 1er janvier 2020.

Date	Durée écoulée au 01/01/2020	Dépôt /retrait (€)	Valeur au 01/01/2020 (€)
01/01/2010	10 ans	10 000	$10\,000 \times 1,05^{10} \simeq 16\,288,95$
01/05/2014	5 ans 8 mois \simeq 5,67 ans	5 000	$5\,000 \times 1,05^{5,67} \simeq 6\,592,39$
01/09/2017	2 ans 4 mois \simeq 2,33 ans	-1 000	$-1\,000 \times 1,05^{2,33} \simeq -1\,120,58$

Au 1er janvier 2020, le capital épargné est de

$$16\,288,95 + 6\,592,39 - 1\,120,58 = 21\,760,76€.$$

TD SEMAINE 7 : EMPRUNTS À TAUX CONSTANT

L'objectif de la semaine est d'apprendre à utiliser la formule liant capital emprunté, remboursements, durée du prêt et taux d'intérêt pour des emprunts à taux constant :

$$C = p \left(\frac{1 - (1 + t)^{-n}}{t} \right),$$

où C est le capital emprunté, p le montant d'un remboursement, n la durée du prêt et t le taux d'intérêts sur une période (par exemple sur 1 mois si les remboursements sont mensuels).

Considérons l'énoncé suivant :

M. Travaux souhaite emprunter 20 000 € au taux d'intérêts de 5% par an. Il souhaite rembourser 300€ par mois. Quelle sera la durée du prêt ?

CONVERSION DE TAUX

Les remboursements étant mensuels, la période de référence est le mois. Nous allons donc convertir le taux d'intérêt en un taux mensuel. En notant t_m le taux mensuel et $t_a = 5\% = 0,05$ le taux annuel,

$$\begin{aligned} (1 + t_m)^{12} &= 1 + t_a = 1,05 \\ 1 + t_m &= 1,05^{\frac{1}{12}} \simeq 1,00407412 \\ t_m &= 1,05^{\frac{1}{12}} - 1 \simeq 0,00407412. \end{aligned}$$

Le taux d'intérêts mensuel est donc d'environ $0,00407412 = 0,407412\%$.

FORMULE DES EMPRUNTS

On part de la formule :

$$C = p \left(\frac{1 - (1 + t)^{-n}}{t} \right),$$

On connaît $C = 20000$, ainsi que $p = 300$ et $t = 0,00407412$. On cherche la durée du prêt en mois n . Par conséquent,

$$\begin{aligned} 20\,000 &= 300 \left(\frac{1 - 1,00407412^{-n}}{0,00407412} \right) \\ 1 - 1,00407412^{-n} &= \frac{20000 \times 0,00407412}{300} \simeq 0,271608 \\ 1,00407412^{-n} &\simeq 1 - 0,271608 = 0,728392. \end{aligned}$$

Ce type d'équation a déjà été travaillé (recherche de durée). On trouve directement :

$$\begin{aligned} -n &\simeq \frac{\log(0,728392)}{\log(1,00407412)} \\ n &\simeq -\frac{\log(0,728392)}{\log(1,00407412)} \simeq 77,95. \end{aligned}$$

Le prêt sera remboursé après 77,95 ans. En arrondissant au mois supérieur et en convertissant en années, on trouve une durée de 78 mois, soit 6 ans et 6 mois.

REMARQUE

S'il n'y avait pas d'intérêts, le nombre de remboursements aurait simplement été de :

$$\frac{20\,000}{300} = 66,7,$$

et donc le prêt aurait duré 67 mois, soit 5 ans et 7 mois. Même avec un taux d'intérêts relativement faible, les intérêts cumulés ont un effet significatif sur sa durée du prêt.

TD SEMAINE 8 : PRÉSENTATION DE DONNÉES

L'objectif de la semaine est de classer des variables statistiques et d'adopter des représentations adaptées pour celles-ci. Considérons l'énoncé suivant :

Une chaîne de cinémas a envoyé un questionnaire à ses clients, afin de connaître :

- leur sexe (F/M) ;
- leur âge ;
- leur département de résidence ;
- le genre de films qu'ils vont voir ("Comédie", "Drame", "Action") ;
- ce qu'ils ont pensé du service (de "Très mauvais" à "Très bon").

Quel type de représentation adopter pour chacune de ces variables ?

CLASSIFICATION DES VARIABLES STATISTIQUES

Le sexe est une variable qualitative (ce n'est pas un nombre) non-ordonnée (l'un n'est pas supérieur à l'autre). On pourra utiliser un **diagramme circulaire**.

L'âge est une variable quantitative (c'est un nombre, et on peut faire des moyennes d'âges) continue¹. Il faudra le trier par classe avant de le représenter par un **histogramme**.

Le département de résidence est une variable qualitative² non-ordonnée (il n'y a pas des départements supérieurs à d'autres). Cependant, il y a beaucoup trop de départements pour utiliser un diagramme circulaire. On pourra utiliser éventuellement un **diagramme en bâtons**, ou, pour sortir du cours, garder un tableau ou tracer une carte.

Le genre des films est une variable qualitative non ordonnée, présentant peu de modalités. On le représentera par un **diagramme circulaire**.

La qualité du service est une variable qualitative (ce n'est pas un nombre), discrète (il n'y a qu'un nombre fini de modalités), ordonnée (un service "Bon" est meilleur qu'un service "Moyen"). On utilisera un **diagramme en bâtons**.

1. Comme l'on donne en général seulement la partie entière de son âge en années, l'âge peut aussi être vu comme une variable discrète, mais avec un grand nombre de classes. Cela ne change pas la conclusion sur la représentation à adopter.

2. Attention : s'il y a des numéros de départements, ce sont bien des numéros, et non des nombres : faire des moyennes de départements n'a aucun sens.

TD SEMAINE 9 : HISTOGRAMME ET FRÉQUENCES CUMULÉES

L'objectif de la semaine est d'apprendre à représenter les variables continues à l'aide d'un histogramme, et de construire la courbe des fréquences cumulées. Considérons l'énoncé suivant :

On a mesuré la taille des étudiants dans une classe. Après un tri par classe, on a obtenu les données suivantes :

Taille (en cm)	[120,150[[150,160[[160,170[[170,180[[180,190[[190, 210[
Effectif	3	22	65	62	16	4

Quelle proportion d'étudiants mesure entre 165 et 175 centimètres, approximativement ?

À L'AIDE D'UN HISTOGRAMME

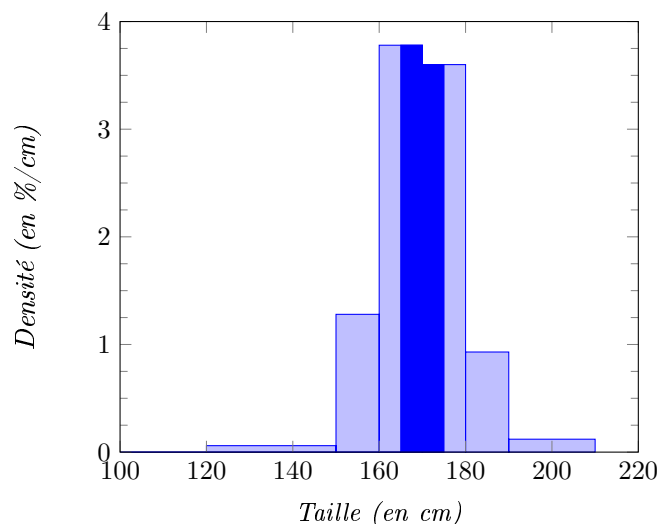
Tout d'abord, on suit la méthode du cours pour construire l'histogramme correspondant à ces données. Pour cela, il faut calculer les fréquences, les amplitudes et les densités associées à chaque classe. L'effectif total est de $3 + 22 + 65 + 62 + 16 + 4$ étudiants. On obtient le tableau suivant :

Taille (en cm)	[120,150[[150,160[[160,170[[170,180[[180,190[[190, 210[
Fréquence (en %)	1,7	12,8	37,8	36,0	9,3	2,3
Amplitude (en cm)	30	10	10	10	10	20
Densité (en %/cm)	0,06	1,28	3,78	3,60	0,93	0,12

Rappelons que la fréquence est l'effectif de la classe divisé par l'effectif total, que l'amplitude est la largeur de la classe de tailles, et que

$$\text{densité} = \frac{\text{fréquence}}{\text{amplitude}}.$$

À partir de là, on peut construire l'histogramme.



Dans un histogramme, la fréquence d'une modalité est égale à l'aire au-dessus de cette modalité. On recherche la proportion étudiants mesurant entre 165 et 175 centimètres. Cette fréquence est égale à l'aire en bleu foncé dans l'histogramme ci-dessus.

Cette aire est la somme des aires de deux rectangles. Celui de gauche a une largeur de $(170 - 165) = 5$ centimètres et une hauteur de 3,78% par centimètre, et donc une aire de $5 \times 3,78 = 18,9\%$. Celui de droite a une largeur de $(175 - 170) = 5$ centimètres et une hauteur de 3,60% par centimètre, et donc une aire de $5 \times 3,60 = 18,0\%$. La fréquence totale est donc de $18,9 + 18,0 = 36,9\%$: il y a environ 36,9% des étudiants de la classe qui mesurent entre 165 et 175 centimètres.

À L'AIDE DE LA COURBE DES FRÉQUENCES CUMULÉES

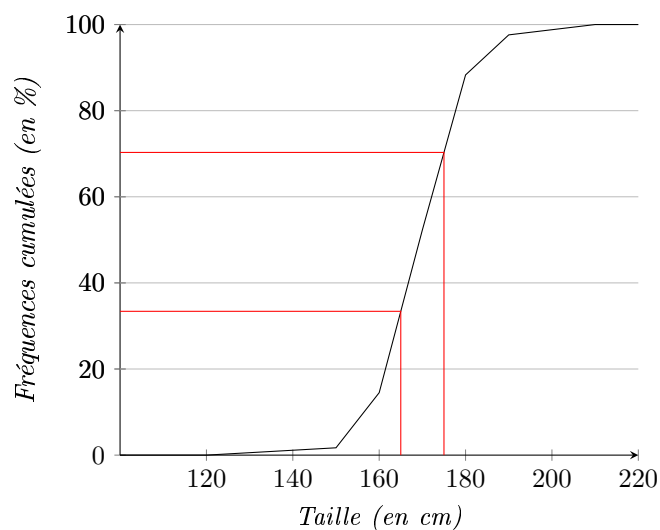
En repartant du tableau des fréquences :

Taille (en cm)	[120,150[[150,160[[160,170[[170,180[[180,190[[190, 210[
Fréquence (en %)	1,7	12,8	37,8	36,0	9,3	2,3

On peut construire le tableau des fréquences cumulées :

Taille (en cm)	120	150	160	170	180	190	210
Fréquences cumulées (en %)	0	1,7	14,5	52,3	88,3	97,6	100

et ensuite tracer la courbe des fréquences cumulées :



Une lecture graphique (segment en rouge nous donne :

- environ 33% des étudiants mesurent moins de 165 centimètres ;
- environ 70% des étudiants mesurent moins de 175 centimètres ;

en, par conséquent, environ $70 - 33 = 37\%$ des étudiants mesurent entre 165 et 175 centimètres.

Remarque : La méthode utilisée ici est une lecture graphique, par nature imprécise. On pourrait calculer exactement les valeurs de la courbe des fréquences cumulées à 165 et 175 centimètres, par exemple en calculant des équations de droites affines. La réponse finale serait alors la même que pour la méthode utilisant l'histogramme (36,9% des étudiants entre 165 et 175 centimètres). Ceci dit, le tri par classes introduit déjà son lot d'imprécisions, de telle sorte qu'une précision inférieure au pourcent est illusoire, et que la lecture graphique est en pratique largement suffisante.

TD SEMAINE 10 : HISTOGRAMME ET FRÉQUENCES CUMULÉES

L'objectif de la semaine est de savoir calculer des moyennes et des médianes. Considérons l'énoncé suivant :

Une étude statistique du lectorat d'une revue a permis d'obtenir les données suivantes :

Âge (en années)	[10,50[[50,60[[60,70[[70,90[
Fréquence (en %)	20	20	30	30

Quelles sont les âges moyen et médian du lectorat ? Commentez.

MOYENNE

Pour calculer la moyenne, on détermine la valeur centrale de chaque classe :

Âge (en années)	[10,50[[50,60[[60,70[[70,90[
Valeur centrale (en années)	30	55	65	80
Fréquence (en %)	20	20	30	30

La moyenne suit directement :

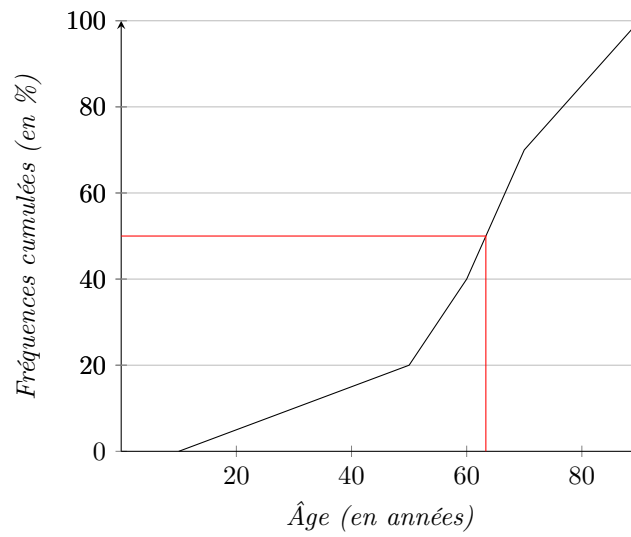
$$0,20 \times 30 + 0,20 \times 55 + 0,30 \times 65 + 0,30 \times 80 = 60,5.$$

L'âge moyen du lectorat est donc de 60,5 ans.

MÉDIANE

Nous allons lire la médiane sur la courbe des fréquences cumulées, qu'il faut d'abord construire.

Âge (en années)	10	50	60	70	90
Fréquences cumulées (en %)	0	20	40	70	100



Graphiquement, on lit que l'âge médian est d'environ 65 ans (une valeur plus précise, calculée, serait de 63,33 ans).

On remarque que l'âge moyen est significativement inférieur à l'âge médian (différence de 3 ans). C'est le signe que les âges sont concentrées vers les hautes valeurs, mais avec un minorité de lecteurs pouvant avoir des âges beaucoup plus bas (et qui tirent donc la moyenne vers le bas).

TD SEMAINE 11 : FRACTILES

L'objectif de la semaine est de savoir calculer des fractiles, et les représenter à l'aide d'un diagramme en boîte. Considérons l'énoncé suivant :

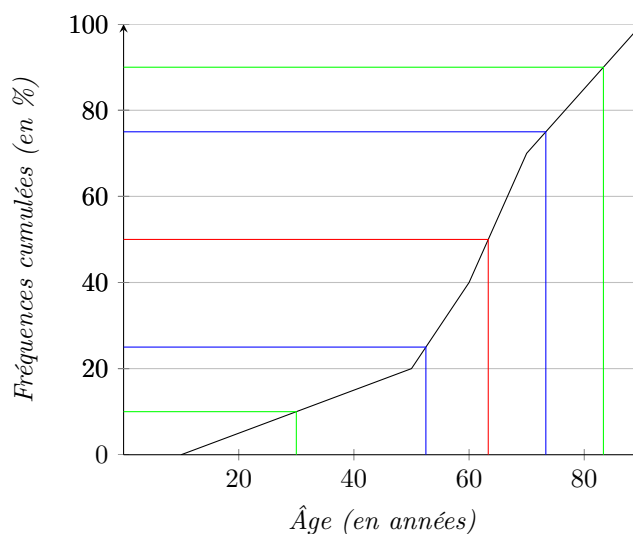
Une étude statistique du lectorat d'une revue a permis d'obtenir les données suivantes :

Âge (en années)	[10,50[[50,60[[60,70[[70,90[
Fréquence (en %)	20	20	30	30

Dessinez le diagramme en boîte associé à cette variable statistique. Commentez.

MÉDIANE, QUARTILES ET DÉCILES

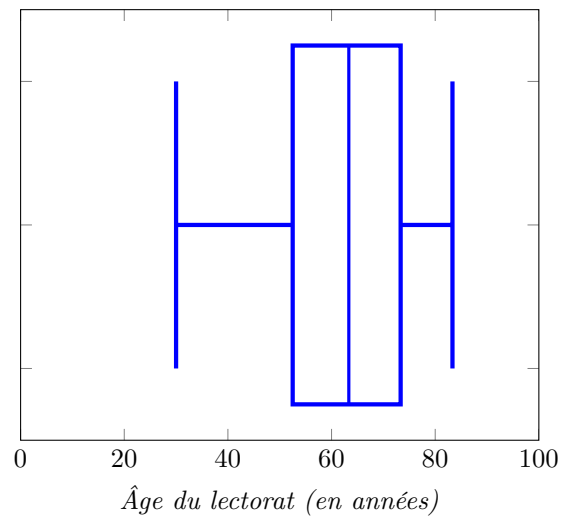
Nous allons lire la médiane, les quartiles et les déciles extrêmes sur la courbe des fréquences cumulées, construite la semaine passée.



Graphiquement, on lit :

- $D_1 \approx 30$ ans (courbe verte du bas).
- $Q_1 \approx 52$ ans (courbe bleue du bas).
- Médiane : environ 65 ans (courbe rouge).
- $Q_3 \approx 75$ ans (courbe bleue du haut).
- $D_9 \approx 85$ ans (courbe verte du haut).

Ceci permet de dessiner enfin le diagramme en boîte associé :



Comme remarqué la semaine dernière, la distribution de âges est très asymétrique : le lectorat est globalement concentré vers des âges élevés (entre 52 et 85 ans), mais avec une "queue" assez longue de lecteurs moins âgés.

TD SEMAINE 12 : PARAMÈTRES DE DISPERSION

L'objectif de la semaine est de savoir calculer des paramètres de dispersion, et en particulier l'écart-type. Considérons l'énoncé suivant :

Une étude statistique du lectorat d'une revue a permis d'obtenir les données suivantes :

Âge (en années)	[10,50[[50,60[[60,70[[70,90[
Fréquence (en %)	20	20	30	30

Quel est l'écart-type de l'âge du lectorat ?

MOYENNE

Commençons par calculer la moyenne, on détermine la valeur centrale de chaque classe :

Âge (en années)	[10,50[[50,60[[60,70[[70,90[
Valeur centrale (en années)	30	55	65	80
Fréquence (en %)	20	20	30	30

La moyenne suit directement :

$$0,20 \times 30 + 0,20 \times 55 + 0,30 \times 65 + 0,30 \times 80 = 60,5.$$

L'âge moyen du lectorat est donc de 60,5 ans.

VARIANCE

Nous allons maintenant calculer la variance de l'âge. Pour cela, nous remplaçons là encore chaque classe par sa valeur centrale, et calculons successivement les écarts à la moyenne, les carrés des écarts à la moyenne, et la moyenne des carrés des écarts à la moyenne (c'est-à-dire la variance).

Âge (en années)	[10,50[[50,60[[60,70[[70,90[
Valeur centrale (en années)	30	55	65	80
Écart à la moyenne (en années)	$30 - 60,5 = -30,5$	-5,5	4,5	19,5
Carrés des écarts à la moyenne (en années ²)	$(-30,5)^2 = 930,25$	30,25	20,25	380,25
Fréquence (en %)	20	20	30	30

On calcule donc la variance, qui est la moyenne de l'avant-dernière ligne du tableau :

$$\begin{aligned} \sigma^2 &= 0,2 \times 930,25 + 0,2 \times 30,25 + 0,3 \times 20,25 + 0,3 \times 380,25 \\ &= 312,25 \end{aligned}$$

La variance est donc de 312,25 années².

ÉCART-TYPE

L'écart-type est la racine carrée de la variance :

$$\sigma = \sqrt{312,25} = 17,7$$

L'écart-type est donc de 17,7 ans.

Remarque : Si l'on ajoutait le terme correctif apparaissant dans le polycopié de cours, on trouverait un variance de 353,08 années², et donc un écart-type de 18,8 ans. La différence avec la valeur calculée précédemment est faible. Il est donc justifié de ne pas ajouter ce terme.