Cours d'été 2012

MATHÉMATIQUES

Pour les futurs bacheliers en :
Sciences économiques
Ingénieur de gestion

Syllabus conçu par : Hector Cordova Bulens
Bao Long Dang Van
Paul Arnaud Songhafouo Tsopméné
Maguy Tréfois

Table des matières

Chapitre 1

Limites, continuité et dérivées

Objectifs du chapitre : *Comprendre la notion de limite d'une fonction en un point, ainsi que la notion de continuité d'une fonction. Entrainer les étudiants au calcul de la dérivée. Connaître l'interprétation géométrique du nombre dérivé. Familiariser les étudiants avec l'utilisation de la tangente au graphe d'une fonction en un point pour approximer cette fonction autour de ce point.*

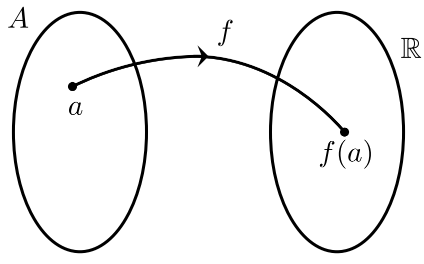
1.1 Fonctions

Cette section a pour objectif de familiariser les étudiants avec la notion de fonction réelle.

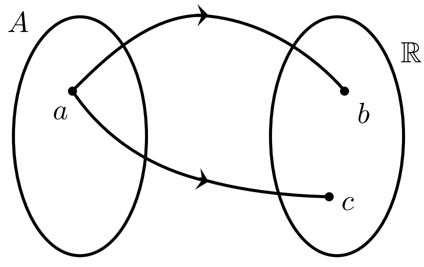
Définition : fonction
Une **fonction** $f : A \subset \mathbb{R} \to \mathbb{R}$ est une relation qui, étant donné un élément $a \in A$, lui associe exactement un et un seul élément de $\mathbb{R}$. L'élément associé à a est noté $f(a)$.

On appelle **domaine** de $f : A \subset \mathbb{R} \to \mathbb{R}$ l'ensemble A.

Une telle relation peut être illustrée de la manière suivante :



Considérons la relation illustrée ci-dessous :



Cette relation n'est pas une fonction, car elle associe à un nombre réel a deux nombres distincts b et c de $\mathbb{R}$.

On appelle **expression analytique** une expression du type :

$$f(x) = \frac{1}{x} \text{ ou } f(x) = \cos x.$$

Le **domaine de définition** est le plus grand sous-ensemble de $\mathbb{R}$ sur lequel l'expression analytique a du sens.

Exemple. Le domaine de définition de $f(x) = \frac{1}{x}$ est $\mathbb{R}_0$.

1.2 Limite d'une fonction en un point

Considérons une fonction f et un nombre $a \in \mathbb{R}$ tel que f est définie sur un voisinage de a, c'est-à-dire qu'il existe un intervalle $I \subset \mathbb{R}$ de longueur non nulle contenant le réel a et tel que la fonction f est définie en tous les points de $I \backslash \{a\}$. On dira que la limite de la fonction f quand x tend vers a vaut L si "quand x se rapproche sans cesse de a, alors $f(x)$ se rapproche sans cesse de L".

L'activité suivante devrait nous aider à comprendre la définition de limite d'une fonction en un point.

Activité 1.

(a) Considérons la fonction $f : \mathbb{R}_0 \to \mathbb{R} : x \mapsto x^3 + 2$. Donnez un intervalle $I \subset \mathbb{R}$ de la forme $[0 - \delta, 0 + \delta]$ (oË˜ δ est un nombre > 0) tel que pour tout $x \in I \backslash \{0\}$, on a que $1 \leq f(x) \leq 3$.

(b) Considérons la fonction $f : \mathbb{R} \to \mathbb{R}$ définie par :

$$f(x) = \begin{cases} x^2 + 2 & \text{si } x < 1 \\ -x^2 + 2 & \text{si } x \geq 1. \end{cases}$$

Est-il possible de trouver un intervalle $I \subset \mathbb{R}$ du type $[1 - \delta, 1 + \delta]$ (avec $\delta > 0$) tel que pour tout $x \in I \backslash \{1\}$, on ait que $0 \leq f(x) \leq 2$?

(c) Soient $a, b \in \mathbb{R}$. Comment exprimer la distance entre a et b à l'aide d'une expression mathématique ?

Essayons maintenant de comprendre la définition suivante.

Définition : limite

Soient $I \subset \mathbb{R}$ un intervalle de longueur non nulle, un nombre $a \in I$ et une application $f : I \backslash \{a\} \to \mathbb{R}$. La limite de la fonction f quand x tend vers a vaut L si :

$$\forall \epsilon > 0, \exists \delta > 0 \text{ tel que } \forall x \in I \backslash \{a\} : \text{si } |x - a| \leq \delta, \text{ alors } |f(x) - L| \leq \epsilon.$$

La limite L de la fonction f en a est notée $\lim\limits_{x \to a} f(x)$.

Voici quelques propriétés qui nous permettent de ramener le calcul d'une limite au calcul d'autres limites.

Propriétés

Soient $I \subset \mathbb{R}$ un intervalle de longueur non nulle, un nombre $a \in I$, un nombre quelconque $\lambda \in \mathbb{R}$, deux applications $f : I \backslash \{a\} \to \mathbb{R}$ et $g : I \backslash \{a\} \to \mathbb{R}$. Si f et g ont une limite en a, alors :

(a) la fonction $f + g$ a une limite en a et :

$$\lim_{x \to a} (f + g)(x) = \lim_{x \to a} f(x) + \lim_{x \to a} g(x),$$

(b) la fonction λf a une limite en a et :

$$\lim_{x \to a} \lambda f(x) = \lambda \lim_{x \to a} f(x),$$

(c) la fonction fg a une limite en a et :

$$\lim_{x \to a} (fg)(x) = \lim_{x \to a} f(x) \cdot \lim_{x \to a} g(x),$$

(d) si $f(x) \neq 0$ pour tout $x \in I \backslash \{a\}$ et $\lim_{x \to a} f(x)$ est différente de 0, alors la fonction $1/f$ a une limite en a et :

$$\lim_{x \to a} \frac{1}{f(x)} = \frac{1}{\lim_{x \to a} f(x)}.$$

1.3 Continuité d'une fonction en un point

Dans cette section, donnons-nous une fonction $f : I \to \mathbb{R}$ définie sur un intervalle $I \subset \mathbb{R}$ de longueur non nulle, et un nombre $a \in I$. Commençons par considérer la proposition suivante.

Proposition

La fonction f est continue en a si et seulement si $\lim_{x \to a} f(x)$ existe et vaut $f(a)$.

Activité 2. À partir de la proposition précédente et de la définition de la notion de limite, définissez la continuité de f au point a.

Voici quelques propriétés sur les fonctions continues.

Propriétés

Soient f, g deux fonctions définies sur un intervalle $I \subset \mathbb{R}$ de longueur non nulle, un nombre $a \in I$ et un réel $\lambda \in \mathbb{R}$. Si f et g sont continues en a, alors :

(a) la fonction $f + g$ est continue en a,

(b) la fonction λf est continue en a,

(c) la fonction fg est continue en a,

(d) si $f(x) \neq 0$ pour tout $x \in I$, alors $1/f$ est continue en a.

Finalement, nous allons donner un théorème important qui concerne les fonctions continues sur un intervalle fermé.

Théorème de Bolzano
Soit $f : [a, b] \to \mathbb{R}$ une fonction continue telle que $f(a)$ et $f(b)$ sont de signes contraires (c'est-à-dire que $f(a).f(b) < 0$), alors il existe un réel $c \in]a, b[$ tel que $f(c) = 0$.

Dans les exercices, nous verrons des applications de ce théorème.

1.4 Dérivées

Commençons par rappeler quelques formules bien connues :

Formules de dérivation de fonctions usuelles

(a) Si k est une constante, alors $(k)' = 0$,
(b) $(kf(x))' = kf'(x)$,
(c) $(x^k)' = kx^{k-1}$,
(d) $(f(x) + g(x))' = f'(x) + g'(x)$,
(e) $(f(x)g(x))' = f'(x)g(x) + f(x)g'(x)$,
(f) $\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x)g(x)-f(x)g'(x)}{g(x)^2}$,
(g) $(f \circ g)'(x) = f'(g(x))g'(x)$,
(h) $(\sin x)' = \cos x$,
(i) $(\cos x)' = -\sin x$
(j) $(\tan x)' = \frac{1}{\cos^2 x}$,
(k) $(e^x)' = e^x$,
(l) $(\ln x)' = \frac{1}{x}$.

Toutes les formules que nous venons de rappeler découlent de la définition du nombre dérivé que voici.

Définition : nombre dérivé
Soient $I \subset \mathbb{R}$ un intervalle de longueur non nulle, $a \in I$ et une application $f : I \to \mathbb{R}$. On dit que f est dérivable en a si

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

existe. Dans ce cas, cette limite est appelée le **nombre dérivé** de f en a et est notée $f'(a)$.

Interprétation géométrique du nombre dérivé

Lorsqu'une fonction $f : I \to \mathbb{R}$ est dérivable en un point $a \in I$, le nombre dérivé $f'(a)$ est la pente de la tangente au graphe de f au point $(a, f(a))$. L'équation de la tangente en ce point est alors donnée par :

$$y = f(a) + f'(a)(x - a).$$

Si la fonction f est compliquée, une approche très répandue consiste à approximer f par une fonction qui est simple par excellence : un polynôme. En particulier, on peut rechercher le polynôme de degré au plus n qui approxime le mieux la fonction f autour d'un point a parmi tous les autres polynômes de degré au plus n. Ce polynôme, s'il existe, est appelé **polynôme de Taylor** de f d'ordre n autour du point a. Il se note :

$$T^n_{f,a}(x).$$

Si on est proche du point a, on peut donc "remplacer" f par le polynôme $T^n_{f,a}$, ce que l'on note $f(x) \approx T^n_{f,a}(x)$.

Dans le cas où $n = 1$, le polynôme de Taylor $T^1_{f,a}(x)$ est appelé **approximation linéaire** de f en a et n'est rien d'autre que la tangente au graphe de f au point $(a, f(a))$, c'est-à-dire le polynôme :

$$T^1_{f,a}(x) = f(a) + f'(a)(x - a).$$

Exemple. On peut trouver une approximation du nombre $\sin(0.1)$ à partir du polynôme de Taylor d'ordre 1 de f autour du point 0. Ce polynôme n'est rien d'autre que la tangente au graphe de f au point $(0, f(0))$ qui a comme équation :

$$\begin{aligned} y &= f(0) + f'(0)(x - 0) \\ &= \sin 0 + \cos 0(x - 0) \\ &= x. \end{aligned}$$

Par conséquent, $\sin(0,1) \approx 0.1$.

Pour terminer ce chapitre, voici une proposition montrant le lien entre la continuité et la dérivabilité d'une fonction.

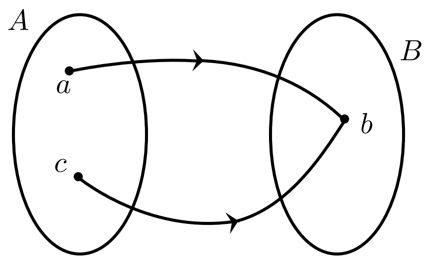
Proposition
Si une fonction est dérivable en un point a de son domaine, alors elle est continue en a.

Cependant, remarquons que la réciproque est fausse. En effet, la fonction $f : \mathbb{R} \to \mathbb{R} : x \mapsto |x|$ est continue en 0, mais n'est pas dérivable en 0.

1.5 Exercices

Fonctions

Exercice 1.1. Considérons la relation :



Cette relation est-elle une fonction ?

Exercice 1.2. Calculer le domaine de définition de :

(a) $f(x) = x^2 + 3$
(b) $f(x) = \frac{1}{x^2-x-6}$
(c) $f(x) = \sqrt{x+2}$
(d) $f(x) = \frac{1}{\sqrt{x^2-3x+2}}$

(e) $f(x) = \tan(3x)$
(f) $f(x) = \ln x^2$

Exercice supplémentaire 1.3. Calculer le domaine de définition de :
(a) $f(x) = \frac{1}{(x-2)^2}$
(b) $f(x) = \frac{x}{\sqrt{x^2-1}}$
(c) $f(x) = \frac{1}{\sin x}$
(d) $f(x) = \cos(3x)$
(e) $f(x) = \ln(\ln(\ln(x)))$

Limite d'une fonction en un point

Exercice 1.4. Calculer les limites suivantes.
(a) $\lim_{x\to 1} x^3 + x - 4$
(b) $\lim_{x\to 4} \frac{x^2-16}{x-4}$
(c) $\lim_{x\to 2} \frac{1}{x^2-4x+4}$
(d) $\lim_{x\to 1} \frac{x^2+3x-4}{x-1}$

Exercice supplémentaire 1.5. Calculer les limites suivantes.
(a) $\lim_{x\to 1} \frac{x^3+1}{x+2}$
(b) $\lim_{x\to 1} \frac{x^4-1}{x-1}$
(c) $\lim_{x\to 0} \frac{5x^2+x}{x}$
(d) $\lim_{x\to 2} \frac{x^2-4}{x-2}$
(e) $\lim_{x\to 5} \frac{1}{x^2-10x+25}$

Continuité d'une fonction en un point

Exercice 1.6. La fonction
$$f(x) = \begin{cases} \frac{x^2-9}{x-3} & \text{si } x \neq 3 \\ 6 & \text{si } x = 3 \end{cases}$$
est-elle continue en $x = 3$?

Exercice 1.7. Comment faudrait-il définir la fonction
$$f(x) = \frac{x^5-1}{x-1}$$
en $x = 1$ pour que $\lim\limits_{x\to 1} f(x) = f(1)$?

Exercice supplémentaire 1.8. La fonction
$$f(x) = \begin{cases} \frac{12x^2-4x^3}{x-3} & \text{si } x \neq 3 \\ -28 & \text{si } x = 3 \end{cases}$$
est-elle continue en $x = 3$?

Exercice supplémentaire 1.9. La fonction
$$f(x) = \begin{cases} \frac{x^2-3x+2}{x-1} & \text{si } x \neq 1 \\ -1 & \text{si } x = 1 \end{cases}$$
est-elle continue en $x = 1$?

Exercice 1.10. Montrer que la fonction

$$f : [0,1] \to \mathbb{R} : x \mapsto e^x - 3x$$

admet au moins une racine.

Exercice 1.11. Montrer que le polynôme $P(x) = x^3 - x - 1$ possède une racine sur l'intervalle $[0,2]$.

Exercice 1.12. Montrer que les équations suivantes ont une solution.

(a) $x^{10} - 5x^2 + 1 = 0$
(b) $\sin x = x - 1$
(c) $x^2 = \sqrt{x+1}$
(d) $e^x = 2 - x$
(e) $\ln(x+1) = 1 - x$

Exercice 1.13. Montrer que l'équation

$$e^x + x^3 = 3$$

admet une solution réelle.

Dérivées

Exercice 1.14. Calculer la dérivée des fonctions suivantes.

(a) $f(x) = 2x^5 + x^2$
(b) $f(x) = \sqrt{x}$
(c) $f(x) = \sqrt{2x^3 + 5}$
(d) $f(x) = (x+2).(x^2+1)$
(e) $f(x) = \sqrt{1+x^3}.(3x+1)$
(f) $f(x) = \frac{x^3}{x^4+1}$
(g) $f(x) = \frac{\sqrt{x^2+1}}{x^4}$
(h) $f(x) = \cos(2x)$
(i) $f(x) = \exp(x^2)$
(j) $f(x) = \ln(3x+1)$

Exercice 1.15. Calculer le nombre dérivé de x^3 en $x = 2$ à l'aide de sa définition.

Exercice supplémentaire 1.16. Dériver les fonctions suivantes.

(a) $f(x) = (x^3 + 2x + 2)(x^2 + 3x)$
(b) $f(x) = \frac{x^3-3x+6}{x^2-1}$
(c) $f(x) = \frac{1}{x^3} + \frac{x}{x^2+1}$
(d) $f(x) = \frac{(x^2+1)^2+2}{(x^2+1)^2-2}$
(e) $f(x) = \frac{1}{(x^2-3)(x+2)^2}$
(f) $f(x) = \cos^2(3x)$
(g) $f(x) = \ln(\ln(x))$

Exercice supplémentaire 1.17. Calculer la dérivée de la fonction

$$f(x) = \begin{cases} x^2 \cos \frac{1}{x} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

en $x = 0$.

Exercice 1.18. Donner l'équation de la tangente au graphe de la fonction au point $(1, f(1))$.

(a) $f(x) = \sqrt{x} + \frac{1}{2.(x+1)}$

(b) $f(x) = \frac{2x+1}{x+1}$

Exercice 1.19. Calculer le polynôme de Taylor d'ordre 1 de la fonction

$$f : \mathbb{R} \to \mathbb{R} : x \mapsto x^2$$

en $a = 1$. Ensuite, calculer une valeur approximative de l'expression $(1,03)^2$.

Exercice 1.20. Calculer une valeur approximative des expressions suivantes en utilisant le polynôme de Taylor d'ordre 1 de la fonction $f(x) = \sqrt{x}$ autour du point $a = 9$.

(a) $\sqrt{9,02}$

(b) $\sqrt{10}$

(c) $\sqrt{8}$

Exercice supplémentaire 1.21. Trouver l'équation de la tangente au graphe de f au point $(x_0, f(x_0))$.

(a) $f(x) = 1 - x^2, \qquad x_0 = 2$

(b) $f(x) = x^2 - x, \qquad x_0 = 1$

(c) $f(x) = (x^2 - 7)\frac{3x}{x+2}, \qquad x_0 = 1$

(d) $f(x) = \frac{x}{1+x^2}, \qquad x_0 = 2$

Chapitre 2

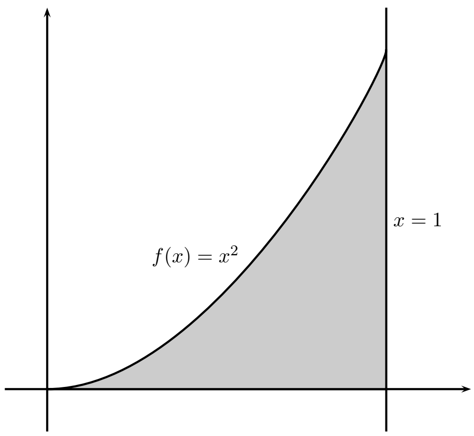
Calcul intégral

Objectifs du chapitre : *Familiariser les étudiants au calcul intégral. Introduction à certaines méthodes d'intégrations : intégration directe, intégration par parties et intégration par changement de variables.*

2.1 Motivation : le calcul d'aires

Nous savons depuis le primaire que l'aire d'un rectangle est égale à sa longueur multipliée par sa largeur, que l'aire d'un cercle est égale au carré de son rayon fois π, que l'aire d'un triangle est égale à la moitié du produit de sa base par sa hauteur. Ces formules sont connues par coeur.

Considérons la région ci-contre :



Cette région est délimitée par l'axe des abscisses, les droites d'équations $x = 0$ et $x = 1$, et la courbe d'équation $y = x^2$. On a aussi envie de calculer l'aire exacte de cette région. Ce qui n'est pas une mince à faire. Il se trouve, par magie, que l'intégrale est l'outil le mieux adapté pour faire ce genre de calcul.

L'objectif principal de ce chapitre est de pouvoir calculer l'intégrale d'une fonction continue sur un intervalle. Dans la première partie, on donnera la définition d'une intégrale. Dans la deuxième partie, on étudiera les primitives et on expliquera pourquoi elles sont l'outil fondamental du calcul intégral. Dans la troisième partie, on parlera des propriétés que possède l'intégrale.

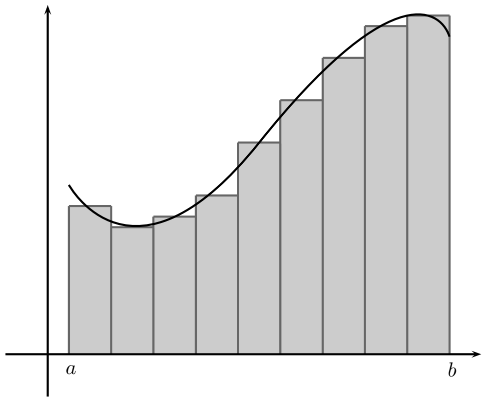
2.2 Définition de l'intégrale

Soit $f : [a, b] \to \mathbb{R}$ une fonction. Soit $n \geq 1$, $n \in \mathbb{N}$. On subdivise l'intervalle $[a, b]$ en n intervalles $[a_i, a_{i+1}]$ de même longueur (où $i \in \{1, \cdots, n\}$), avec $a_1 = a$ et $a_{n+1} = b$.

La longueur des sous-intervalles est $a_{i+1} - a_i = \frac{b-a}{n}$. Choisissons un point $c_i \in [a_i, a_{i+1}]$ dans chaque intervalle. L'expression $\frac{b-a}{n} f(c_i)$ donne l'aire du rectangle de base $[a_i, a_{i+1}]$ et de hauteur $f(c_i)$. Dès lors, la somme

$$S_n(f) = \sum_{i=1}^{n} \frac{b-a}{n} f(c_i)$$

donne l'aire de l'union des rectangles de base $[a_i, a_{i+1}]$ et de hauteur $f(c_i)$. Voir figure ci-dessous.



Cette somme d'aires approxime l'aire de la région délimitée par le graphe de f, l'axe des abscisses et les droites d'équation $x = a$ et $x = b$. On imagine que si on augmente le nombre n de sous-intervalles, notre approximation sera meilleure.

Définition : intégrale définie

L'**intégrale** d'une fonction f sur l'intervalle $[a, b]$, notée $\int_a^b f(x)\,dx$, est la limite pour n tendant vers l'infini de la somme $S_n(f)$. Autrement dit :

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} S_n(f).$$

Toutefois, cette définition est peu pratique pour calculer une intégrale. Heureusement, un autre moyen fort commode pour calculer des intégrales nous est donné par la notion de primitive.

2.3 Primitives et théorème fondamental

Définition : primitive d'une fonction

Soit une fonction $f : [a, b] \to \mathbb{R}$. On appelle **primitive** de f toute fonction $F : [a, b] \to \mathbb{R}$ dérivable telle que, pour tout $x \in [a, b]$, $F'(x) = f(x)$.

Remarquons que si une fonction f admet une primitive F, alors $F + k$, où $k \in \mathbb{R}$ est une constante, est aussi une primitive de f. Par conséquent, si une fonction admet une primitive, alors elle en admet une infinité.

Voici quelques fonctions usuelles et leurs primitives :

$f(x)$	$F(x)$
x	$\frac{x^2}{2}+k$
x^2	$\frac{x^3}{3}+k$
x^n	$\frac{x^{n+1}}{n+1}+k$
$\sin x$	$-\cos x+k$
$\cos x$	$\sin x+k$
e^x	e^x+k
$\frac{1}{x}$	$\ln\lvert x\rvert+k$
$\frac{1}{x^2}$	$-\frac{1}{x}+k$

Le résultat suivant nous montre que la plupart des fonctions usuelles admettent une primitive.

Proposition

Toute fonction continue $f:[a,b]\to\mathbb{R}$ admet une primitive sur $[a,b]$.

Un lien, assez extraordinaire, entre intégrales et primitives est donné par le théorème suivant.

Théorème Fondamental du Calcul Différentiel et Intégral

Soient $f:[a,b]\to\mathbb{R}$ une fonction continue et F une primitive de f. Alors

$$\int_a^b f(x)\,dx = F(b)-F(a).$$

Ce théorème nous affirme que pour calculer l'intégrale d'une fonction il suffit de connaître une de ses primitives ! Dans la suite, une primitive de f sera notée simplement $\int f(x)dx$.

2.4 Méthodes de primitivation

2.4.1 Primitives immédiates

Dans le chapitre précédent nous avons appris à calculer la dérivée de la composée de deux fonctions. On sait, par exemple, que

$$(e^{x^2})' = 2xe^{x^2}$$

et, dès lors, e^{x^2} est une primitve de $2xe^{x^2}$. Ceci nous montre que grâce au connaissances qu'on a du calcul des dérivées, nous connaissons plusieurs primitives. Voici une liste des plus importantes :

$$\int f^r(x)f'(x)\,dx = \frac{f^{r+1}(x)}{r+1} + k \quad \text{pour } r \in \mathbb{R}\backslash\{-1\}$$

$$\int \frac{f'(x)}{f(x)}\,dx = \ln|f(x)| + k$$

$$\int e^{f(x)}f'(x)\,dx = e^{f(x)} + k$$

$$\int f'(x)\cos f(x)\,dx = \sin f(x) + k$$

$$\int f'(x)\sin f(x)\,dx = -\cos f(x) + k$$

$$\int \frac{f'(x)}{1+f^2(x)}\,dx = \arctan f(x) + k$$

$$\int \frac{f'(x)}{\sqrt{1-f^2(x)}}\,dx = \arcsin f(x) + k$$

$$\int \frac{-f'(x)}{\sqrt{1-f^2(x)}}\,dx, = \arccos f(x) + k.$$

Activité 3. Calculer les primitives des fonctions suivantes :

(a) $x\sqrt{1-x^2}$,

(b) $\frac{\ln^2(x)}{x}$,

(c) $x^2\sin(x^3+1)$,

(d) $\frac{x}{2+2x^2}$.

2.4.2 Linéarité de la primitivation

Nous savons que $(f+g)'(x) = f'(x) + g'(x)$. Dès lors, si F est une primitive de f et G une primitive de g, $F+G$ est une primitive de $f+g$. On a donc

$$\int (f+g)(x)\,dx = \int f(x)\,dx + \int g(x)\,dx.$$

Pour des raisons similaires, nous avons

$$\int \lambda f(x)\,dx = \lambda \int f(x)\,dx$$

où $\lambda \in \mathbb{R}$.

Activité 4. Calculer

$$\int (2\sin(x) + e^x + 4x^2)\,dx$$

Attention : en général, $\int_a^b f(x)g(x)\,dx$ n'est pas égal à $\left(\int_a^b f(x)\,dx\right) \cdot \left(\int_a^b g(x)\,dx\right)$!

2.4.3 Primitivation par parties

Soient $f, g : [a,b] \to \mathbb{R}$ deux fonctions dérivables. La formule de dérivation d'un produit nous dit que

$$(fg)'(x) = f'(x)g(x) + f(x)g'(x),$$

de là, on déduit qu'une primitive de $f'(x)g(x)+f(x)g'(x)$ est donnée par $(fg)(x)$. On en déduit la formule suivante :

Formule de primitivation par parties

$$\int f'(x)g(x)\,dx = f(x)g(x) - \int f(x)g'(x)\,dx.$$

Activité 5. Calculer les primitives de $f(x) = x\sin(x)$, $g(x) = e^x \sin(x)$ et $h(x) = \ln(x)$.

2.4.4 Primitivation par substitution

Soit $f : [a,b] \to \mathbb{R}$ une fonction ayant une primitive F sur $[a,b]$. Soit $g : [a,b] \to \mathbb{R}$ une fonction dérivable dont la dérivée g' est continue. La formule de dérivation de fonctions composées nous dit que :

$$\begin{aligned}(F(g(x)))' &= F'(g(x))\,g'(x)\\ &= f(g(x)\,)g'(x).\end{aligned}$$

Cela nous donne la formule suivante :

Primitivation par substitution

$$\int f(g(x))\,g'(x)\,dx = F(g(x)).$$

Un moyen mnémotechnique pour retenir cette formule consiste, face à une primitive à calculer de la forme $\int f(g(x))g'(x)\,dx$, à poser $u = g(x)$. On a que $\frac{du}{dx} = g'(x)$. En faisant “comme si” $\frac{du}{dx}$ était une fraction, on en tire $du = g'(x)\,dx$. Dès lors, la primitive à calculer $\int f(g(x))g'(x)\,dx$ devient $\int f(u)\,du$ et comme F est une primitive de f :

$$\int f(g(x))\,g'(x)\,dx = \int f(u)\,du = F(u) = F(g(x)).$$

Pour comprendre l’utilité de cette formule regardons un exemple : pour calculer la primitive de $\frac{\ln(x)}{x}$, on pose $u = g(x) = \ln(x)$, ce qui donne $du = (\ln x)'\,dx = \frac{1}{x}\,dx$. On a alors

$$\int \frac{\ln(x)}{x}\,dx = \int u\,du = \frac{u^2}{2} + k = \frac{(\ln x)^2}{2} + k.$$

Ces différentes techniques de primitivation nous donnent, grâce au théorème fondamental, des techniques d’intégration.

Activité 6. Grâce au théorème fondamental et aux techniques de primitivation, calculez les intégrales suivantes :

(a) $\int_{\pi}^{2\pi} \cos x \sin x dx$,

(b) $\int_0^1 x^2 e^x dx$,

(c) $\int_{-1}^1 \frac{4}{3+2x^2} dx$.

2.5 Propriétés des intégrales

Citons quelques propriétés inhérentes à la notion d’intégrale :

(a) $\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx$,

(b) $\int_a^a f(x)\,dx = 0$,

(c) pour deux fonctions $f, g : [a,b] \to \mathbb{R}$ et un nombre $\lambda \in \mathbb{R}$, on a que

$$\int_a^b (f(x)+g(x))\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx,$$
$$\int_a^b \lambda f(x)\,dx = \lambda \int_a^b f(x)\,dx,$$

(d) pour un nombre $c \in [a,b]$, on a que

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx,$$

(e) si $f : [a,b] \to \mathbb{R}$ est ≥ 0 pour tout $x \in [a,b]$, alors

$$\int_a^b f(x)\,dx \geq 0,$$

(f) si $f, g : [a,b] \to \mathbb{R}$ sont telles que $f(x) \leq g(x)$ pour tout $x \in [a,b]$, alors

$$\int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx.$$

Activité 7. Calculez l'aire comprise entre les courbes $y = x^2 + 2x + 1$, $y = x^2 - 2$, $x = 0$, $y = 0$ et $x = 2$.

2.6 Exercices

Exercice 2.1. Calculez une primitive des fonctions suivantes.

(a)	$5x^6$	(f)	$6x^2+8x+3$	(k)	$x(x+2)(x-3)$
(b)	$\sqrt{2x}$	(g)	$\frac{1}{\sqrt{x}}$	(l)	$\frac{1-\sqrt{x}}{\sqrt{x}}$
(c)	$x^3 \ln x$	(h)	$\frac{1}{x^2+7}$	(m)	$\frac{1}{\sqrt{8-x^2}}$
(d)	3^x	(i)	$3^x e^x$	(n)	$\tan(x)$
(e)	$\sin x \cos x$	(j)	xe^{x^2}	(o)	$\frac{ln^2 x}{x}$

Exercice 2.2. Calculer les primitives suivantes par parties :

(a)	$\int \ln x\,dx$	(e)	$\int x2 \ln x\,dx$	(i)	$\int x \arctan x\,dx$
(b)	$\int \frac{x}{e^x}\,dx$	(f)	$\int e^x \sin x\,dx$	(j)	$\int x \sin x \cos x\,dx$
(c)	$\int \arctan x\,dx$	(g)	$\int x^2 e^{3x}\,dx$		
(d)	$\int \arcsin x\,dx$	(h)	$\int x \sin x\,dx$		

Exercice 2.3. Calculer les primitives suivantes par substitution :

(a)	$\int (\tan x)^2\,dx$	(e)	$\int \frac{(\cos x)^3}{(\sin x)^4}\,dx$	(i)	$\int \frac{\cos\sqrt{x}}{\sqrt{x}}\,dx$
(b)	$\int xe^{x^2}\,dx$	(f)	$\int (\sin x)^2\,dx$	(j)	$\int \sqrt{\frac{x}{1-x}}\,dx$
(c)	$\int \frac{1}{1+e^x}\,dx$	(g)	$\int \frac{1+x}{1+\sqrt{x}}\,dx$	(k)	$\int \frac{\sqrt{1-x^2}}{x^2}\,dx$
(d)	$\int \frac{1}{x(\ln x)^2} dx$	(h)	$\int (\sin 2x)\sqrt{1+\cos 2x}\,dx$	(l)	$\int \frac{1}{x^2\sqrt{1+x^2}}\,dx$

Exercice 2.4. Calculez les intégrales suivantes :

(a)	$\int_0^{\frac{\pi}{2}} \sin\left(t + \dfrac{\pi}{4}\right) dt$	(e)	$\int_0^{\pi} 2\sin^2(x)\, dx$
(b)	$\int_0^1 \frac{2t}{\sqrt{1+t^2}}\, dt$	(f)	$\int_{-1}^3 \lvert x^2 - 4\rvert\, dx$
(c)	$\int_{-1}^0 (2u+1)\, e^{-u}\, du$	(g)	$\int_0^1 \frac{4x^2+5}{x+1}\, dx$
(d)	$\int_0^1 (x^2-4)\, e^{-x}\, dx$		

Exercice 2.5. Soit $I = \displaystyle\int_0^1 \frac{\mathrm{d}x}{\sqrt{x^2+2}}$.

(a) Calculer la dérivée de la fonction $f(x) = \sqrt{x^2+2}$.

(b) En déduire la dérivée de la fonction g définie sur $[0,1]$ par $g(x) = \ln(x + \sqrt{x^2+2})$.

(c) Calculer I.

Exercice 2.6. Soit f la fonction définie sur $\mathbb{R}$ par $f(x) = 2x - \frac{x}{x^2+1}$ et C sa courbe représentative. La droite Δ d'équation $y = 2x$ est asymptote à C à l'infini. Soit E le domaine limité par C, Δ et les droites d'équation $x = 0$ et $x = 2$.

(a) Sans tracer C situer C par rapport à Δ sur $[0,2]$.

(b) Calculer l'aire du domaine E.

Chapitre 3

Systèmes d'équations linéaires et calcul matriciel

Objectifs du chapitre : *Familiariser les étudiants au calcul matriciel dans le but de leur apprendre à résoudre des systèmes d'équations linéaires.*

3.1 Introduction

Activité 8. Considérons la devinette suivante :

> Un frère s'adresse à sa soeur :
> – Sais-tu que si je te donnais deux de mes années, tu aurais le double des années qu'il me resterait ?
> – Mais bien évidemment. Et, ai-je envie de rajouter, sais-tu, cher frère, que si je te donnais trois des années qui sont miennes, toi tu aurais le triple de ce qu'il me resterait comme âge ?
> – Tu as raison.
> Quel est l'âge des deux enfants ?

Voici une marche à suivre pour résoudre le problème :

(a) Montrez que le problème peut se réécrire sous la forme

$$\begin{cases} x - 3y &= -12 \\ 2x - y &= 6 \end{cases}$$

où x représente l'âge du petit garçon et y l'âge de sa soeur.

(b) Utilisez cette réécriture du problème pour trouver l'âge des deux enfants.

Le problème ci-dessus est celui de la résolution d'un **système d'équations linéaires**. Dans ce chapitre du cours, nous verrons comment résoudre ce type de problèmes.

Systèmes d'équations à 2 inconnues

Pour commencer, notre objectif sera de résoudre des systèmes de la forme

$$\begin{cases} ax + by &= u \\ cx + dy &= v \end{cases}$$

où a, b, c, d, u et v sont des nombres réels fixés et x et y sont les inconnues. Une solution de ce système est un couple $(s_1, s_2) \in \mathbb{R}^2$ qui vérifie les deux équations, c'est-à-dire tel que :

$$\begin{cases} as_1 + bs_2 &= u \\ cs_1 + ds_2 &= v. \end{cases}$$

Par exemple, le couple $(6,6)$ est solution du système :

$$\begin{cases} x-3y & =-12 \\ 2x-y & =6. \end{cases}$$

Une première façon de résoudre est d'utiliser la deuxième équation pour obtenir $y=2x-6$ et d'injecter ceci dans la première équation et ainsi obtenir que $x=6$. Ayant maintenant la valeur de x, on trouve la valeur $y=2x-6=2\times 6-6=6$.

Une façon qui sera plus facile à généraliser est d'utiliser les différentes propriétés qu'on connaît sur les équations et de remplacer la deuxième équation par "elle-même moins 2 fois la première". On obtient le système :

$$\begin{cases} x-3y & =-12 \\ \quad 5y & =30. \end{cases}$$

où la solution est assez simple à tirer. La deuxième équation nous dit maintenant que $y=6$ et en injectant ceci dans la première, on trouve $x=6$. Notons que l'on résout ce système "en partant du bas".

Il est tout à fait concevable d'imaginer des systèmes à plus de deux inconnues et à plusieurs équations. Pour faciliter les manipulations dans ce cas général nous allons introduire de nouveaux objets mathématiques qu'on appelle matrices.

3.2 Calcul Matriciel

Définitions : matrice, genre

Une **matrice** est un tableau de nombres. Le **genre** d'une matrice est un couple de nombres naturels (n,m), où n est le nombre de lignes de la matrice et m le nombre de colonnes. Si A est une matrice de genre (n,m), on note A_{ij} l'élément situé à la $i^{\text{ème}}$ ligne et à la $j^{\text{ème}}$ colonne de A. Enfin, l'ensemble des matrices de genre (n,m) est noté $\mathbb{R}^{n\times m}$.

Par exemple, le genre de la matrice

$$A=\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 7 \end{pmatrix}.$$

est $(2,3)$. L'élément A_{12} vaut 2 et l'élément A_{23} vaut 7.

Addition de matrices

Comme pour les nombres réels nous pouvons définir une addition sur les matrices. L'addition matricielle n'est définie que pour des matrices de même genre.

Définition : addition matricielle

Pour $A\in\mathbb{R}^{n\times m}$ et $B\in\mathbb{R}^{n\times m}$ on définit la matrice $A+B\in\mathbb{R}^{n\times m}$ par :

$$(A+B)_{ij}=A_{ij}+B_{ij}$$

où $i\in\{1,\cdots,n\}$ et $j\in\{1,\cdots,m\}$.

Par exemple, soient les matrices :

$$A=\begin{pmatrix} 1 & 3 & 7 \\ 2 & 15 & 4 \\ 2 & 6 & 8 \end{pmatrix} \quad \text{et} \quad B=\begin{pmatrix} 13 & 17 & 2 \\ 4 & 5 & 21 \\ 8 & 3 & 0 \end{pmatrix}.$$

Ces matrices étant de même genre, on peut les additionner. On obtient :

$$A+B=\begin{pmatrix}1+13 & 3+17 & 7+2\\ 2+4 & 15+5 & 4+21\\ 2+8 & 6+3 & 8+0\end{pmatrix}=\begin{pmatrix}14 & 20 & 9\\ 6 & 20 & 25\\ 10 & 9 & 8\end{pmatrix}.$$

Produit de matrices

Il existe aussi un produit entre deux matrices. Cette opération est moins naturelle que l'addition. En effet, elle ne s'effectue pas en général sur des matrices de même genre et ne s'effectue pas élément par élément comme cela se passe pour l'addition.

Définition : produit matriciel

Soient deux matrices $A \in \mathbb{R}^{n\times m}$ et $B \in \mathbb{R}^{m\times p}$. On définit le produit $AB \in \mathbb{R}^{n\times p}$ par

$$(AB)_{ij} = A_{i1}B_{1j} + A_{i2}B_{2j} + \cdots + A_{im}B_{mj}$$

où $i \in \{1, \cdots, n\}$ et $j \in \{1, \cdots p\}$.

Il est très important de remarquer que pour que le produit matriciel de deux matrices soit défini, il faut que le nombre de colonnes de la première matrice soit égale au nombre de lignes de la deuxième.

Par exemple, soient les matrices

$$A=\begin{pmatrix}1 & 2 & 3\\ 4 & 0 & 2\end{pmatrix} \quad \text{et} \quad B=\begin{pmatrix}2 & 1 & 4\\ 0 & 1 & 2\\ 3 & 2 & 0\end{pmatrix}.$$

Le nombre de colonnes de $A \in \mathbb{R}^{2\times 3}$ étant égal au nombre de lignes de $B \in \mathbb{R}^{3\times 3}$, le produit matriciel $AB \in \mathbb{R}^{2\times 3}$ est défini. On a

$$\begin{aligned}AB &= \begin{pmatrix}1\times 2+2\times 0+3\times 3 & 1\times 1+2\times 1+3\times 2 & 1\times 4+2\times 2+3\times 0\\ 4\times 2+0\times 0+2\times 3 & 4\times 1+0\times 1+2\times 2 & 4\times 4+0\times 2+2\times 0\end{pmatrix}\\ &= \begin{pmatrix}2+9 & 1+2+6 & 4+4\\ 8+6 & 4+4 & 16\end{pmatrix}\\ &= \begin{pmatrix}11 & 9 & 8\\ 14 & 8 & 16\end{pmatrix}.\end{aligned}$$

Remarquons que le produit BA n'est pas défini.

Activité 9. Soit le système d'équations

$$\begin{cases} 2x+3y &= 6\\ 3x+2y &= 7\end{cases}$$

En utilisant le produit matriciel, écrire le système sous forme de matrices.

3.3 Opérations élémentaires et échelonnement

Nous allons voir maintenant comment profiter de la notation matricielle pour résoudre des systèmes d'équations.

Définition : matrice échelonnée

Une matrice est dite **échelonnée** si le nombre de zéros précédant la première valeur non nulle d'une ligne augmente strictement d'une ligne à la suivante.

Par exemple, les matrices

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 3 & 4 \\ 0 & 0 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}$$

sont échelonnées. Par contre les matrices

$$\begin{pmatrix} 1 & 2 & 4 \\ 2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 4 \end{pmatrix}$$

ne le sont pas.

Étant donnée une matrice, on notera L_i la $i^{\text{ème}}$ ligne de cette matrice.

Définition : opérations élémentaires

Une **opération élémentaire** sur les lignes d'une matrice est une des opérations suivantes :

(a) Ajouter à une ligne L_i un multiple k d'une autre ligne L_j, ce qu'on note $L_i \rightsquigarrow L_i + kL_j$.

(b) Multiplier une ligne L_i par un réel k non nul, ce qu'on note $L_i \rightsquigarrow kL_i$.

(c) Permuter deux lignes L_i et L_j, ce qu'on note $L_i \leftrightsquigarrow L_j$.

On peut, en n'utilisant que des opérations élémentaires, transformer toute matrice en une matrice échelonnée. C'est ce qu'on appelle l'**échelonnement**.

Par exemple, échelonnons la matrice

$$\begin{pmatrix} 1 & 2 & 4 \\ 2 & 5 & 7 \\ 6 & -2 & 8 \end{pmatrix}.$$

Pour cela, commençons par ajouter $-2 \times L_1$ à L_2 :

$$\begin{pmatrix} 1 & 2 & 4 \\ 2 & 5 & 7 \\ 6 & -2 & 8 \end{pmatrix} \overset{L_2 \rightsquigarrow L_2 - 2L_1}{\rightarrow} \begin{pmatrix} 1 & 2 & 4 \\ 0 & 1 & -1 \\ 6 & -2 & 8 \end{pmatrix}.$$

Continuons en ajoutant $-6 \times L_1$ à L_3 :

$$\begin{pmatrix} 1 & 2 & 4 \\ 0 & 1 & -1 \\ 6 & -2 & 8 \end{pmatrix} \overset{L_3 \rightsquigarrow L_3 - 6L_1}{\rightarrow} \begin{pmatrix} 1 & 2 & 4 \\ 0 & 1 & -1 \\ 0 & -14 & -16 \end{pmatrix}.$$

Et pour finir, on peut ajouter $14 \times L_2$ à L_3 :

$$\begin{pmatrix} 1 & 2 & 4 \\ 0 & 1 & -1 \\ 0 & -14 & -16 \end{pmatrix} \overset{L_3 \rightsquigarrow L_3 + 14L_2}{\rightarrow} \begin{pmatrix} 1 & 2 & 4 \\ 0 & 1 & -1 \\ 0 & 0 & -30 \end{pmatrix}.$$

On obtient que $\begin{pmatrix} 1 & 2 & 4 \\ 0 & 1 & -1 \\ 0 & 0 & -30 \end{pmatrix}$ est une matrice échelonnée obtenue à partir de la matrice $\begin{pmatrix} 1 & 2 & 4 \\ 2 & 5 & 7 \\ 6 & -2 & 8 \end{pmatrix}$.

3.4 Résolutions de systèmes d'équations linéaires

3.4.1 Systèmes à solution unique

Notre objectif est de résoudre des systèmes d'équations linéaires. On peut remarquer qu'un système de la forme

$$\begin{cases} 3x + 5y + 2z & = 3 \\ \phantom{3x+{}} 3y + 2z & = 7 \\ \phantom{3x+3y+{}} 5z & = 10 \end{cases}$$

est facile à résoudre. En effet, de la troisième équation, on tire $z = 2$. Dès lors, en sachant que $z = 2$, la deuxième équation nous donne $y = 1$. Finalement en connaissant les valeurs de y et de z on obtient en utilisant la première équation que $x = -2$.

Ce système peut s'écrire sous forme matricielle

$$\begin{pmatrix} 3 & 5 & 2 \\ 0 & 3 & 2 \\ 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ 7 \\ 10 \end{pmatrix}.$$

On appelle matrice complète du système la matrice

$$\left(\begin{array}{ccc|c} 3 & 5 & 2 & 3 \\ 0 & 3 & 2 & 7 \\ 0 & 0 & 5 & 10 \end{array}\right)$$

où on a ajouté, à la matrice des coefficients, une dernière colonne formée par les termes indépendants.

On peut remarquer que dans le cas présent la matrice complète de notre système est échelonnée. C'est ceci qui fait que le système est simple à résoudre. Nous allons utiliser cette idée pour résoudre tout type de système.

Commençons par un exemple. Soit le système

$$\begin{cases} 2x - y + z & = 1 \\ x + y + z & = 1 \\ -x + 2y + z & = -2. \end{cases}$$

Il peut encore s'écrire

$$\begin{pmatrix} 2 & -1 & 1 \\ 1 & 1 & 1 \\ -1 & 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}.$$

La matrice complète du système est

$$\left(\begin{array}{ccc|c} 2 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ -1 & 2 & 1 & -2 \end{array}\right).$$

Cette matrice n'est pas échelonnée. Par contre on peut l'échelonner en faisant, par exemple :

$$\left(\begin{array}{ccc|c} 2 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ -1 & 2 & 1 & -2 \end{array}\right) \overset{L_1 \leftrightsquigarrow L_2}{\rightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 2 & -1 & 1 & 1 \\ -1 & 2 & 1 & -2 \end{array}\right)$$

$$\overset{L_2 \rightsquigarrow L_2 - 2L_1}{\rightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -3 & -1 & -1 \\ -1 & 2 & 1 & -2 \end{array}\right) \overset{L_3 \rightsquigarrow L_3 + L_1}{\rightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -3 & -1 & -1 \\ 0 & 3 & 2 & -1 \end{array}\right)$$

$$\xrightarrow{L_3 \rightsquigarrow L_3+L_2} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -3 & -1 & -1 \\ 0 & 0 & 1 & -2 \end{array}\right).$$

De cette nouvelle matrice complète on peut en retirer un système d'équations :

$$\begin{cases} x+y+z & =1 \\ \quad -3y-z & =-1 \\ \qquad\quad z & =-2. \end{cases}$$

Ce nouveau système est assez facile à résoudre. En effet, on obtient assez rapidement que $(2,1,-2)$ est la solution du système. La partie fondamentale est maintenant de savoir si la solution de ce dernier système facile à résoudre est la même solution que celle de notre système de départ. La réponse est oui, ce que nous verrons dans la suite.

Activité 10. Traduire, au niveau du système, les opérations effectuées sur la matrice complète lors de l'échelonnement. Justifier pourquoi ces opérations ne vont pas changer la solution.

3.4.2 Systèmes avec plusieurs solutions

Jusqu'à présent, les solutions aux systèmes existaient et étaient toujours uniques. Mais cela n'est pas vrai pour tous les systèmes. Illustrons cela avec les systèmes suivant :

$$\begin{cases} 4x+2y+5z & =12 \\ \qquad 2y+2z & =14. \end{cases}$$

Si on fixe une valeur pour z disons $\alpha \in \mathbb{R}$, alors on obtient que $y = \frac{14-2\alpha}{2} = 7-\alpha$ et finalement $x = \frac{-2-3\alpha}{4}$. On peut remarquer que pour n'importe quel α qu'on choisisse, le triplet $(\frac{-2-3\alpha}{4}, 7-\alpha, \alpha)$, est une solution du système. En fait, l'ensemble

$$\left\{ \left(\frac{-2-3\alpha}{4}, 7-\alpha, \alpha \right) \mid \alpha \in \mathbb{R} \right\}$$

est l'ensemble de toutes les solutions du système et contient une infinité d'éléments.

Par contre, le système suivant :

$$\begin{cases} 2x+2y & =0 \\ \quad x+y & =1. \end{cases}$$

n'admet aucune solution.

3.4.3 Systèmes avec plusieurs solutions

3.4.4 Pour aller plus loin : le cas général

Notre objectif est de montrer que le raisonnement qu'on a utilisé pour les exemples précédents peut facilement se généraliser aux systèmes avec des nombres variés d'équations et d'inconnues. Il faut malheureusement travailler avec des notations beaucoup plus lourdes.

Un **système de m équations linéaires à n inconnues** $(x_1, \dots, x_n)$ est une suite d'équations linéaires

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = b_2 \\ & \ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = b_m. \end{cases}$$

Une **solution** du système ci-dessus est une suite de n nombres $(s_1, \ldots, s_n)$ tels que :

$$\left\{\begin{array}{rl} a_{11}s_1 + a_{12}s_2 + \cdots + a_{1n}s_n & = b_1 \\ a_{21}s_1 + a_{22}s_2 + \cdots + a_{2n}s_n & = b_2 \\ & \vdots \\ a_{m1}s_1 + a_{m2}s_2 + \cdots + a_{mn}s_n & = b_m. \end{array}\right.$$

Par exemple, le système d'équations linéaires :

$$\left\{\begin{array}{rl} 2x_1 + 3x_2 + 4x_3 & = 10 \\ x_1 + x_2 + x_3 & = 4 \end{array}\right.$$

admet $(1, 4, -1)$ comme solution, mais aussi $(2, 2, 0)$.

Quand on nous demande de résoudre un système d'équations, ce qu'on nous demande est de déterminer l'ensemble de toutes les solutions du système. Une définition importante, dans la poursuite de notre objectif, est la suivante.

Définition : systèmes équivalents

Deux systèmes d'équations linéaires sont **équivalents** s'ils ont le même ensemble de solutions.

Avec ce qui précède, nous avons les ingrédients pour développer une méthode pour produire des systèmes équivalents à un système donné.

Proposition

Toute combinaison d'opérations élémentaires sur un système nous donne un système qui lui est équivalent.

De cette propriété, on peut formuler maintenant une méthode générale de résolution d'un système d'équations :

Méthode pour résoudre un système d'équations

Étant donné un système d'équations :

(a) écrire la matrice complète A du système,

(b) échelonner cette matrice, ce qui donne une matrice A',

(c) écrire et résoudre le système d'équations correspondant à la matrice A'. Comme ce nouveau système est équivalent au système initial, les solutions sont les mêmes.

Nous pouvons maintenant appliquer notre méthode de résolution de systèmes à tout système d'équations linéaires. Si on a le système

$$\left\{\begin{array}{rl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = b_2 \\ & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = b_m. \end{array}\right.$$

Il peut se réecrire

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \cdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

La matrice complète du système est

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ & & \cdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right).$$

On peut maintenant échelonner cette matrice complète à l'aide d'opérations élémentaires sur les lignes de la matrice. Il suffit alors d'écrire le système obtenu à l'aide de cette matrice échelonnée. Ce système est un système équivalent au système de départ et son ensemble des solutions est facile à déterminer.

3.5 Exercices

Exercice 3.1. Donner un exemple de matrice $\mathbb{R}^{2\times 2}$, un exemple de matrice $\in \mathbb{R}^{3\times 1}$ et un exemple de matrice $\in \mathbb{R}^{1\times 3}$.

Exercice 3.2. Pour $i \in \{1,2,3\}$ et $j \in \{1,2,3,4\}$, construire la matrice M telle que $M_{ij} = i+j$.

Exercice 3.3. Si possible, calculer $A + B$ pour :

(a) $A = \begin{pmatrix} 4 & 5 \\ 2 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}$,

(b) $A = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$, $B = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$,

(c) $A = \begin{pmatrix} 0 & 0 & 2 & 4 \\ 2 & 0 & 1 & 7 \\ 15 & 0 & 5 & 10 \end{pmatrix}$, $B = \begin{pmatrix} 12 & 21 & 11 & 11 \\ 0 & 2 & 9 & \frac{1}{2} \\ 3 & 2 & 4 & 6 \end{pmatrix}$.

Exercice 3.4. Si possible calculer AB et BA, où A et B sont données par :

(a) $A = \begin{pmatrix} 2 & 3 & 4 \end{pmatrix}$ et $B = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$.

(b) $A = \begin{pmatrix} 2 & 3 & 4 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$.

(c) $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ et $B = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$.

Exercice 3.5. (a) Ecrire sous forme matricielle et résoudre le système.

$$\begin{cases} x + z & = 1 \\ y + z & = 0 \\ x + z & = -1. \end{cases}$$

(b) Faire la même chose pour le système

$$\begin{cases} x - y + z & = -1 \\ 2x - 3y + z & = 2 \\ x + 4z & = -5. \end{cases}$$

Exercice 3.6. (a) Résoudre :

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

(b) Résoudre

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

Exercice 3.7. (a) Résoudre le système

$$\begin{cases} -x_1 + 2x_2 - x_3 &= 1 \\ 2x_1 - x_2 + 5x_3 &= 4 \\ x_1 + x_2 + 4x_3 &= 5. \end{cases}$$

(b) Résoudre

$$\begin{cases} 3x_1 - 9x_2 + x_3 - x_4 &= 3 \\ 2x_1 - 6x_2 - x_3 + 2x_4 &= 2 \\ -x_1 + 3x_2 - 7x_3 + 11x_4 &= -1. \end{cases}$$

Chapitre 4

Introduction aux statistiques

Objectifs du chapitre : *Familiariser les étudiants avec le vocabulaire de base de la statistique descriptive. Leur apprendre à interpréter différentes notions, comme la moyenne, la médiane ou encore l'écart-type, à partir d'exemples simples.*

Introduction

Population – Échantillon – Statistique descriptive et inductive

La **statistique** est l'ensemble des méthodes scientifiques permettant d'extraire de l'information à partir d'une collection de données. En général, les étapes d'une étude statistique sont les suivantes :

- recueil de données : sondages, recensements, etc.,
- organisation des données pour les rendre utilisables : tri, répartition en classes, etc.,
- présentation des données sous la forme de graphiques, de tableaux, etc.,
- analyser les données pour en tirer des conclusions.

Les données concernent certaines caractéristiques d'un ensemble d'individus ou d'objets (appelé **population**). Par exemple, on peut étudier le métier des habitants d'une région, le nombre de pièces défectueuses fabriquées par jour dans une usine, les marques de vêtements des étudiants de l'UCL, les notes obtenues à un examen par les élèves d'une classe, etc.

Pour certaines populations, le nombre d'individus est tellement grand qu'il n'est pas possible de recueillir toutes les données. Dans ce cas, on va sélectionner un sous-ensemble d'individus (appelé **échantillon**) considéré comme représentatif de la population entière. La **statistique inductive** consiste alors à tirer des informations sur une population entière à partir des données recueillies sur un échantillon. Si l'on se contente d'analyser une population donnée sans tirer de conclusions sur une population plus grande, on fait de la **statistique descriptive**.

Ce syllabus est une introduction élémentaire à la statistique descriptive.

Variables

Les caractéristiques d'une population sont souvent assimilées à des **variables** pouvant prendre diverses valeurs. Par exemple, l'âge, le poids, la taille, les goûts musicaux ou le pouvoir d'achat sont des variables possibles pour une population de personnes. La couleur, le prix ou la matière sont des variables possibles pour la population des vêtements d'un magasin.

Une variable peut être **quantitative** ou **qualitative** selon que ses valeurs sont des nombres ou des qualités. La couleur est un exemple de variable qualitative. Dans la suite, nous nous restreindrons aux variables quantitatives.

Notations

Lorsqu'une population est de taille finie, on notera N le nombre d'individus qui la composent et ces individus seront numérotés à l'aide d'un indice i où $i = 1, \ldots, N$. Les variables seront représentées par une lettre majuscule (X, Y, Z, etc.) et X_i désignera la variable X sur le i-ème individu de la population.

4.1 Distribution de fréquences

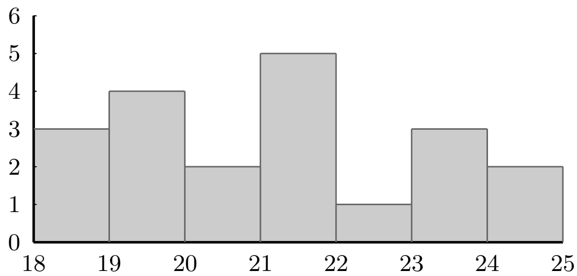
Quand on recueille les données d'une variable X sur les individus d'une population, on obtient d'abord un ensemble de données brutes. Il est souvent pratique de commencer par les ordonner par ordre croissant ou décroissant, ce qui forme une suite ordonnée à partir de laquelle il est facile de calculer certaines quantités comme la médiane ou l'étendue (que nous verrons plus loin).

On peut aussi représenter une suite ordonnée sous la forme d'un diagramme en bâtonnets appelé **histogramme de fréquences**. L'axe des abscisses est celui des valeurs X_i prises par la variable sur les individus, et l'axe des ordonnées est celui des effectifs, c'est-à-dire du nombre d'individus ayant donné une même valeur en abscisses.

Par exemple, considérons le tableau suivant qui liste les âges des 20 participants à un camp de vacances destiné aux 18-24 ans :

18	18	21	22	19	23	24	20	21	23
19	24	21	18	20	19	23	21	19	21

On constate que 3 participants ont 18 ans, 4 ont 19 ans, 2 ont 20 ans, etc. L'histogramme de fréquences associé prend donc la forme suivante :



Données groupées

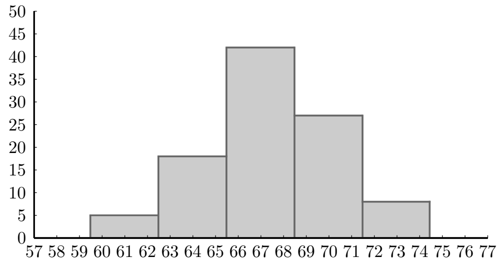
Lorsque la variable d'intérêt peut prendre un nombre infini de valeurs, il est souvent plus pratique de regrouper les données en **classes**, où une classe est un certain intervalle de données. Par exemple, imaginons que l'on mesure le poids des 100 étudiants d'une université. On peut regrouper ces données en classes comme l'illustre le tableau suivant :

Poids (kg)	Nombre d'étudiants
60–62	5
63–65	18
66–68	42
69–71	27
72–74	8

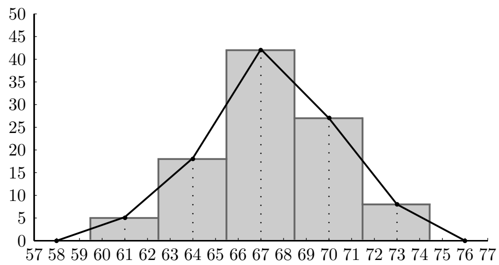
On appelle le **centre d'une classe** la moyenne des bornes de cette classe. Par exemple, le centre de la classe 63–65 est $(63 + 65)/2 = 64$.

Quand on groupe les données de cette façon, on assume que sur les individus d'une classe donnée, la variable prend comme valeur le centre de cette classe. Ainsi, les 27 individus de la classe 69–71 sont considérés comme ayant tous un poids de 70 kg. La répartition en classe provoque donc une certaine perte de données, mais le gain en clarté compense souvent ce défaut. Toutefois, il faut être prudent dans le choix des classes. Une façon commode de procéder est de les choisir de sorte que leurs **amplitudes**, c'est-à-dire les écarts entre leurs bornes, soient toutes égales (ce qui est le cas du tableau ci-dessus où l'amplitude y vaut 2).

On peut dresser un histogramme de fréquences pour des données groupées. La largeur d'un bâtonnet du diagramme couvre une classe et sa hauteur correspond à l'effectif de cette classe :

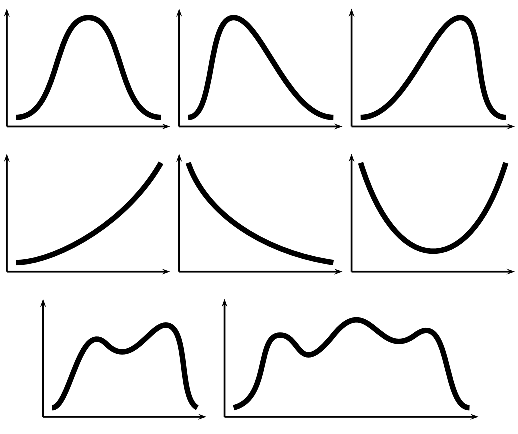


Au lieu d'utiliser des bâtonnets, on pourrait aussi relier entre eux les centres de classes successifs situés aux milieux des sommets des bâtonnets :



La courbe obtenue est appelée **polygone de fréquences** et est formée de segments successifs. Remarquons que l'aire sous ce polygone est égale à l'aire couverte par les bâtonnets.

Lorsque la population est élevée, le nombre de segments augmente avec le nombre de classes et leur longueur diminue. Le polygone se rapproche alors de plus en plus d'une courbe lisse théorique appelée **fonction de densité**. Voici quelques fonctions de densité fréquemment rencontrées en statistiques :



La première courbe est symétrique, c'est-à-dire que les observations situées à égale distance du milieu de la "cloche" ont la même fréquence. L'exemple le plus important est la courbe de **distribution normale** sur laquelle nous reviendrons dans la Section 4.3.

4.2 Mesures de tendance centrale

4.2.1 La moyenne

Activité 11. Considérez le tableau suivant qui présente les notes (sur 20) obtenues à un examen dans une classe de 25 élèves :

20	1,5	8,5	15	6
7,5	19	11,5	8,5	10
5	7	18	18,5	20
8	7	14	11	19,5
6	10	8	8	7,5

- Déterminez la population et la variable qui nous intéressent.
- Quelle est la moyenne des notes obtenues ? [Aide : la somme des notes vaut 275.]
- Que diriez-vous du niveau général de cette classe ?

Activité 12. Pierre, un jeune homme de 20 ans, a réussi ses examens de juin. Pour fêter cela, il s'offre un séjour d'été en Turquie. Après avoir consulté diverses agences de voyages, il finit par hésiter entre deux camps de vacances. Dans la brochure des voyages, il voit que l'âge moyen des femmes célibataires du camp A est de 19 ans, tandis que celui des femmes célibataires du camp B est de 31 ans. Étant célibataire et ouvert à de nouvelles rencontres, il décide d'opter pour le camp A. Que pensez-vous de sa décision ?

Définition : moyenne

Si X est une variable prenant les valeurs $X_1, X_2, \ldots, X_N$ au sein d'une population de taille N, la **moyenne** de X est le nombre suivant, noté μ :

$$\mu = \frac{1}{N} \sum_{i=1}^{N} X_i.$$

Activité 13. Voici maintenant le tableau de l'Activité 11 dans lequel les notes des élèves ont été classées par ordre croissant :

1,5	5	6	6	7
7	7,5	7,5	8	8
8	8,5	8,5	10	10
11	11,5	14	15	18
18,5	19	19,5	20	20

- Quelle est la note la plus souvent obtenue ?
- La moitié des étudiants ont-ils réussi ou raté ?
- Que diriez-vous maintenant du niveau général de la classe ? Si votre réponse diffère de celle de l'exercice 1, comment l'expliquez-vous ?

Activité 14. Marc est un petit agriculteur dans une province A et a un revenu d'environ 700€ par mois, ce qui le met dans une situation difficile. Il lit dans le journal que dans la province B, le revenu moyen des agriculteurs est de 1200€ par mois et décide d'y emménager. Deux ans plus tard, il se rend compte que son revenu mensuel est resté à 700€ par mois, et qu'il est dans la même situation que la majorité des agriculteurs de la province B. Comment peut-on expliquer cela ?

4.2.2 Le mode et la médiane

Définition : mode
Le **mode** d'un ensemble de N données $X_1, \ldots, X_N$ est la valeur la plus souvent rencontrée.

Notons qu'un ensemble de données peut avoir plusieurs modes. Par exemple, l'ensemble de données suivant :

$$3, 1, 2, 4, 5, 8, 7, 4, 5, 6, 8, 8, 5, 6, 9$$

possède deux modes : 5 et 8 (qui apparaissent tous les deux 3 fois). On dit que cet ensemble de données est *bimodal*.

Définition : médiane
Soit $X_1, \ldots, X_N$ un ensemble de N données classées par ordre croissant. La **médiane** de ces données est le nombre défini comme suit :
– si N est impair, alors la médiane est la valeur du milieu,
– si N est pair, alors la médiane est la moyenne des deux valeurs du milieu.
La médiane est donc la valeur qui partage la population étudiée en deux groupes de même effectif.

Par exemple, dans l'ensemble de nombres suivant :

$$1 \quad 3 \quad 3 \quad 6 \quad 7 \quad 9 \quad 20$$

la médiane vaut 6. Dans l'ensemble suivant :

$$2 \quad 4 \quad 5 \quad 8 \quad 10 \quad 12 \quad 13 \quad 14$$

la médiane vaut $(8 + 10)/2 = 9$.

Activité 15. Sur base du tableau de l'Activité 13, déterminez le mode et la médiane des notes obtenues par les élèves de la classe. Que diriez-vous maintenant du niveau général de cette classe ?

Quartiles, déciles et centiles

La médiane est la valeur qui divise la population en deux parts égales. On pourrait chercher les valeurs qui la divisent en 4, 10 ou 100 parts égales, ce qui correspond respectivement aux **quartiles**, aux **déciles** et aux **centiles**.

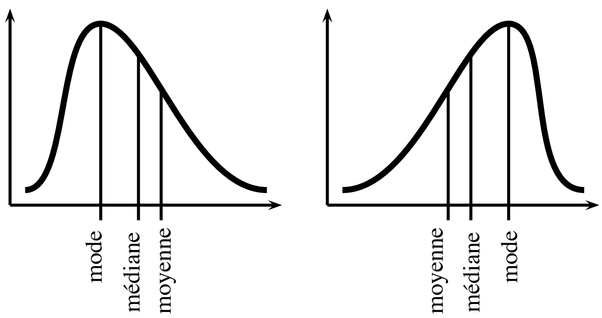
Il y a trois quartiles souvent notés Q_1, Q_2 et Q_3. Le deuxième quartile Q_2 est la médiane. De façon analogue, il y a 9 déciles $D_1, \ldots, D_9$ (D_5 est médiane) et 99 centiles $P_1, \ldots, P_{99}$ ($P_{25} = Q_1$, P_{50} est la médiane et $P_{75} = Q_3$).

De façon générale, on appelle **quantiles** les valeurs qui divisent la population en un nombre donné de parts égales.

4.2.3 Interprétation de la moyenne, du mode et de la médiane

La moyenne, le mode et la médiane sont trois façons d'évaluer, chacune avec ses avantages et ses défauts, la tendance centrale ou dominante d'un ensemble de données $X_1, \ldots, X_N$. Selon la situation, ce sera l'une ou l'autre de ces notions qui sera la plus significative. En présence de valeurs extrêmes, la moyenne peut cesser d'être représentative puisqu'elle est fortement influencée par ces valeurs.

On peut illustrer ces remarques en voyant comment la moyenne, le mode et la médiane sont répartis suivant la forme des courbes de fréquences en "cloche" :



La médiane est toujours située entre la moyenne et le mode. La première courbe correspond à la situation présentée dans l'Activité 11. Si la distribution de fréquences est symétrique, alors les trois mesures ont la même valeur.

Activité 16. Dans les situations suivantes, quelle mesure parmi la moyenne, le mode et la médiane est la plus significative ?

(a) Le patron d'un magasin de vêtements qui analyse le nombre de chemises vendues suivant la taille.

(b) Un professeur qui regarde les points d'un élève à tous les examens pour déterminer sa réussite.

(c) Une personne souhaitant déménager et possédant un budget limité, qui analyse la distribution des prix des maisons dans une certaine région pour déterminer si elle est adaptée à son budget.

4.2.4 Moyenne de données groupées

Soit un ensemble de N données groupées en K classes. On peut le voir comme un ensemble de données non groupées $X_1, \ldots, X_N$ parmi lesquelles certaines valeurs sont égales. Si f_i désigne l'effectif de la $i^{\text{ème}}$ classe, la moyenne vaut alors :

$$\mu = \frac{1}{N} \sum_{i=1}^{K} f_i X_i.$$

L'usage de cette formule est la façon "brutale" de calculer la moyenne. Nous allons voir maintenant une méthode alternative qui s'avère plus pratique et qui a le mérite de s'appliquer à d'autres calculs (comme celui de l'écart-type de données groupées, voir plus loin).

Soit un ensemble de données $X_1, \ldots, X_N$. Commençons par remarquer que si A est un nombre quelconque, on a que :

$$\begin{aligned}
\mu = \frac{1}{N} \sum_{i=1}^{N} X_i &= A + \frac{1}{N} \sum_{i=1}^{N} X_i - A \\
&= A + \frac{1}{N} \left(\sum_{i=1}^{N} X_i - NA \right) = A + \frac{1}{N} \left(\sum_{i=1}^{N} X_i - \sum_{i=1}^{N} A \right) \\
&= A + \frac{1}{N} \sum_{i=1}^{N} (X_i - A).
\end{aligned}$$

Lorsque ces données sont regroupées en K classes, assumons que X_i désigne le centre de la $i^{\text{ème}}$ classe et notons f_i son effectif. La formule ci-dessus devient alors :

$$\mu = A + \frac{1}{N} \sum_{i=1}^{K} f_i (X_i - A)$$

et $N = \sum_{i=1}^{K} f_i$ est la somme des effectifs de toutes les classes.

Maintenant, si les classes ont toutes la même amplitude c, alors $X_i - A$ est un multiple de c. Notons u_i ce multiple, c'est-à-dire que :

$$X_i - A = u_i c.$$

On prendra souvent pour A un des centres de classe. Avec ces notations, la formule précédente devient alors :

Moyenne de N données groupées en K classes de même amplitude c

$$\mu = A + \frac{c}{N} \sum_{i=1}^{K} f_i u_i$$

où A est un des centres de classe (choisi arbitrairement), f_i est l'effectif de la $i^{\text{ème}}$ classe et $u_i = \frac{X_i - A}{c}$.

Mais est-ce que tout ceci facilite réellement le calcul de la moyenne ? L'activité suivante permet de s'en convaincre.

Activité 17. Une compagnie vient de lancer sur le marché une nouvelle variété de boisson énergisante. Elle fournit 50 revendeurs dans une certaine région et a recueilli le nombre de litres vendus par chaque revendeur en deux semaines. Voici les résultats :

Nombre ℓ de litres vendus	Centre de classe X_i	Nombre de vendeurs ayant vendu ℓ litres
$80 \leq \ell < 90$	85	2
$90 \leq \ell < 100$	95	6
$100 \leq \ell < 110$	105	10
$110 \leq \ell < 120$	115	14
$120 \leq \ell < 130$	125	9
$130 \leq \ell < 140$	135	7
$140 \leq \ell < 150$	145	2
		Total : 50

(a) Essayez de calculer "brutalement" la moyenne en utilisant la formule $\mu = \frac{1}{N} \sum_{i=1}^{K} f_i X_i$.

(b) Complétez le tableau suivant et utilisez-le pour calculer la moyenne en utilisant la méthode encadrée ci-dessus. Pour fixer les idées, choisissons de prendre $A = 115$ (le centre de la $4^{\text{ème}}$ classe).

ℓ	X_i	f_i	$X_i - A$	$\frac{X_i - A}{c} = u_i$	$f_i u_i$
$80 \leq \ell < 90$	85	2			
$90 \leq \ell < 100$	95	6			
$100 \leq \ell < 110$	105	10			
$110 \leq \ell < 120$	$115 = A$	14			
$120 \leq \ell < 130$	125	9			
$130 \leq \ell < 140$	135	7			
$140 \leq \ell < 150$	145	2			
		Total : 50			Total :

Mentionnons enfin que l'on peut également définir le mode et la médiane de données groupées. Mais puisque le groupement de données provoque des pertes de données, ces quantités ne peuvent être déterminées qu'approximativement en faisant une interpolation. Toutefois, nous n'aborderons pas ce sujet.

4.3 Mesures de la dispersion

Activité 18. Revenons à l'histoire de Pierre qui voulait partir en vacances et avait opté pour le camp où les femmes célibataires avaient en moyenne 19 ans. Ce qu'il ignore, c'est qu'en cherchant des informations plus soigneusement, il serait tombé sur le tableau suivant qui liste les âges des femmes célibataires inscrites aux camps :

															Total
Camp A :	2	2	2	4	5	7	10	11	11	34	35	35	50	58	**266**
Camp B :	18	19	19	19	19	19	20	20	45	45	46	47	48	50	**434**

Il ignore donc la mauvaise surprise qui l'attend à son arrivée dans le camp A ! Sur base du tableau, expliquez ce qui a induit Pierre en erreur. En particulier, comment les âges dans le camp A varient-ils autour de la moyenne ? Et dans le camp B ?

4.3.1 Étendue, écart moyen et écart-type

Définition : étendue
L'**étendue** d'un ensemble de données $X_1, \ldots, X_N$ est la différence entre la valeur maximale et la valeur minimale.

Définition : écart moyen
L'**écart moyen** d'un ensemble de données $X_1, \ldots, X_N$ est le nombre :

$$E.M. = \frac{1}{N}\sum_{i=1}^{N} |X_i - \mu|$$

où μ est la moyenne des X_i.

L'écart moyen est donc la moyenne des écarts (en valeur absolue) entre chaque valeur X_i et la moyenne μ. Autrement dit, c'est la moyenne de l'ensemble de données suivantes :

$$|X_1 - \mu|, |X_2 - \mu|, \ldots, |X_N - \mu|.$$

Remarquons que sans les valeurs absolues, la formule ci-dessus donnerait toujours un résultat nul (voir exercice 4.5).

Définition : écart-type
L'**écart-type**, souvent noté σ, d'un ensemble de données $X_1, \ldots, X_N$ est le nombre :

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_i - \mu)^2}.$$

L'écart-type est donc obtenu en plusieurs étapes :

(a) D'abord, on calcule la moyenne du carré des écarts $X_i - \mu$, c'est-à-dire la moyenne de l'ensemble de données suivantes :

$$(X_1 - \mu)^2, (X_2 - \mu)^2, \ldots, (X_N - \mu)^2.$$

Le résultat est appelé la **variance** de l'ensemble $X_1, \ldots, X_N$.

(b) Ensuite, on prend la racine carrée de la variance précédente.

Cette seconde étape se justifie par le fait que cela permet à l'écart-type de s'exprimer dans les mêmes unités que les données X_i. Par exemple, considérons la population des camions d'une société de transport et regardons le nombre de kilomètres parcourus par chaque camion. Notons $X_1, \ldots, X_N$ l'ensemble de ces données, chaque X_i étant une distance en kilomètres. Alors la variance des X_i s'exprime en kilomètres carrés. Sa racine carrée, qui est l'écart-type, s'exprime bien en kilomètres.

Propriété : autre formule pour l'écart-type
Si le nombre de données est élevé, une formule permet de calculer plus rapidement l'écart-type :

$$\sigma = \sqrt{\left(\frac{1}{N}\sum_{i=1}^{N} X_i^2\right) - \mu^2}.$$

Il s'obtient donc en prenant d'abord la différence entre la moyenne des carrés des X_i et le carré de la moyenne μ, puis en extrayant la racine carrée du résultat.

4.3.2 Interprétations de l'étendue, de l'écart moyen et de l'écart-type

La signification de l'étendue est claire. L'écart moyen, lui, exprime à quel point les données s'écartent, en moyenne, de μ.

L'écart-type possède la même signification que l'écart moyen, à la différence près que dans l'écart-type, les valeurs extrêmes acquièrent plus d'importance puisqu'elles sont élevées au carré. Ainsi, l'écart-type est d'autant plus grand que les valeurs extrêmes sont extrêmes.

Activité 19. Voici le tableau de l'Activité 18 auquel des données ont été rajoutées. La variable X désigne l'âge des femmes du camp A, et Y est celui des femmes du camp B. Les moyennes sont $\mu_X = 19$ et $\mu_Y = 31$.

	X_i	$X_i - \mu_X$	$\lvert X_i - \mu_X\rvert$	$(X_i - \mu_X)^2$	Y_i	$Y_i - \mu_Y$	$\lvert Y_i - \mu_Y\rvert$	$(Y_i - \mu_Y)^2$
	2	−17	17	289	18	−13	13	169
	2	−17	17	289	19	−12	12	144
	2	−17	17	289	19	−12	12	144
	4	−15	15	225	19	−12	12	144
	5	−14	14	196	19	−12	12	144
	7	−12	12	144	19	−12	12	144
	10	−9	9	81	20	−11	11	121
	11	−8	8	64	20	−11	11	121
	11	−8	8	64	45	14	14	196
	34	15	15	225	45	14	14	196
	35	16	16	256	46	15	15	225
	35	16	16	256	47	16	16	256
	50	31	31	961	48	17	17	289
	58	39	39	1521	50	19	19	361
Total :	**266**	**0**	**234**	**4860**	**434**	**0**	**190**	**2654**

Calculez l'étendue, l'écart moyen et l'écart-type de l'âge des femmes dans chaque camp. Ces informations auraient-elles pu permettre à Pierre de choisir le camp B plutôt que le camp A ?

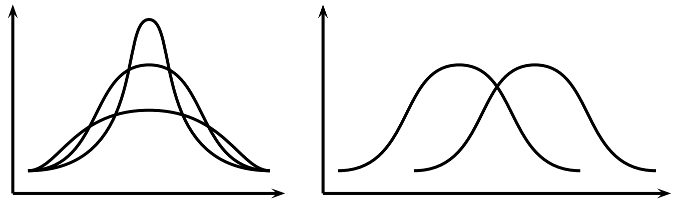
4.3.3 Écart-type et distribution normale

Il y a un lien particulier entre l'écart-type et la distribution normale. En fait, la courbe de cette distribution possède l'équation suivante :

$$f(X) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2}.$$

C'est une formule dans laquelle la moyenne μ et l'écart-type σ sont des paramètres ayant l'impact suivant sur la forme de la courbe :
- la moyenne μ est la valeur en abscisse au-dessus de laquelle se trouve le sommet de la "cloche",
- l'écart-type σ détermine le degré d'aplatissement de la courbe ; plus σ est élevé, plus la courbe est aplatie.

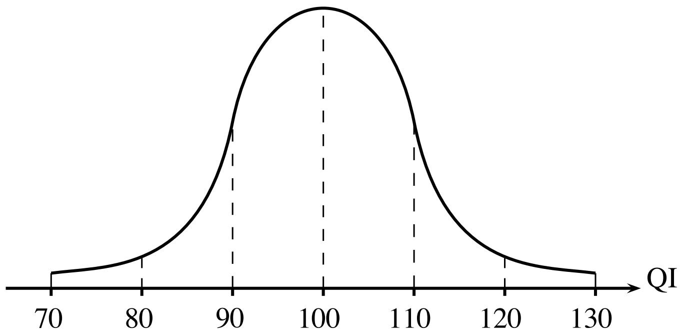
La figure suivante montre différentes courbes normales pour différentes valeurs μ et σ :



Dans la première figure, les courbes ont la même moyenne mais des écarts-types différents. Dans la seconde, c'est l'inverse.

L'écart-type joue aussi un autre rôle dans la distribution normale. Il se trouve que :
- environ $68,27\,\%$ des effectifs sont compris entre $\mu - \sigma$ et $\mu + \sigma$,
- environ $95,45\,\%$ des effectifs sont compris entre $\mu - 2\sigma$ et $\mu + 2\sigma$,
- environ $99,73\,\%$ des effectifs sont compris entre $\mu - 3\sigma$ et $\mu + 3\sigma$.

Un phénomène dont la distribution est habituellement considérée comme normale est celui de la répartition du QI au sein d'une population. La courbe de fréquences a l'allure suivante :



La moyenne mesurée μ vaut environ 100 et l'écart-type σ vaut environ 10. Cela signifie que :
- environ $68,27\,\%$ des gens ont un QI entre 90 et 110,
- environ $95,45\,\%$ des gens ont un QI entre 80 et 120,
- environ $99,73\,\%$ des gens ont un QI entre 70 et 130.

4.3.4 Écart-type de données groupées

En reprenant les notations introduites au point 4.2.4, on a la formule suivante :

Écart-type de N données groupées en K classes de même amplitude c

$$\sigma = c\sqrt{\frac{1}{N}\sum_{i=1}^{K} f_i u_i^2 - \left(\frac{1}{N}\sum_{i=1}^{K} f_i u_i\right)^2}$$

où f_i est l'effectif de la $i^{\text{ème}}$ classe et $u_i = \frac{X_i - A}{c}$ avec A un centre de classe arbitrairement choisi.

Mentionnons que l'on peut aussi définir un écart moyen de données groupées, mais on ne s'attardera pas sur le sujet. Cela dit, l'écart-type est la mesure de dispersion la plus utilisée dans une situation de groupement.

Activité 20. Reprenez le second tableau de l'Activité 17 et ajoutez-y une colonne $f_i u_i^2$. Utilisez ensuite ces données pour calculez l'écart-type.

4.4 Exercices

Exercice 4.1. Écrire les sommes suivantes en développant tous les termes.

(a) $\sum_{i=1}^{6} X_i$

(b) $\sum_{i=1}^{4} (Y_i - 3)^2$

(c) $\sum_{i=1}^{7} A$ où A est un nombre fixé.

(d) $\sum_{i=1}^{5} (X_k - A)$ où A est un nombre fixé.

Exercice 4.2. Condensez les expressions suivantes avec un symbole de sommation $\sum$.

(a) $X_1^2 + X_2^2 + \ldots + X_{10}^2$

(b) $(X_1 + Y_1) + (X_2 + Y_2) + \ldots + (X_8 + Y_8)$

(c) $f_1 X_1^3 + f_2 X_2^3 + \ldots + f_{20} X_{20}^3$

(d) $a_1 b_1 + a_2 b_2 + \ldots + a_N b_N$

Exercice 4.3. Montrez que $\sum_{i=1}^{N} (aX_i + bY_i) = a \sum_{i=1}^{N} X_i + b \sum_{i=1}^{N} Y_i$.

Exercice 4.4. Les salaires annuels de cinq employés sont de 19 200€, 20 000€, 57 600€, 20 200€et 18 600€. Calculez la moyenne et la médiane des salaires puis discutez des résultats.

Exercice 4.5. Montrez que la somme des écarts des $X_1, \ldots, X_N$ à leur moyenne μ est nulle :

$$\sum_{i=1}^{N} (X_i - \mu) = 0.$$

Exercice 4.6. On a recensé les âges des spectateurs d'une salle de cinéma durant la projection du dernier *Harry Potter*. Les résultats sont :

Tranche d'âge	Spectateurs
de 10 à 14 ans	3
de 15 à 19 ans	7
de 20 à 24 ans	16
de 25 à 29 ans	12
de 30 à 34 ans	9
de 35 à 39 ans	5
de 40 à 44 ans	2

Calculez l'âge moyen des spectateurs de cette salle.

Exercice 4.7. Considérons la suite de nombres 3, 6, 2, 1, 7, 5.

(a) Calculez sa moyenne et son écart-type.

(b) Ajoutez 5 à chaque nombre puis calculez la moyenne et l'écart-type de la suite qui en résulte. Comparez avec les résultats du (a).

(c) Multipliez chaque nombre par 2 puis calculez la moyenne et l'écart-type de la suite qui en résulte. Comparez avec les résultats précédents.