Chapitre 1

Notations des nombres réels. Proportionnalité

I Identités remarquables

On rappelle que pour tous nombres a et b, on a :

- $(a+b)^2 = a^2 + 2ab + b^2$
- $(a-b)^2 = a^2 - 2ab + b^2$
- $a^2 - b^2 = (a-b)(a+b)$

II Nombre réels

1) Nombres entiers, décimaux, fractions

Un « calcul », au départ c'est un caillou. Un caillou que l'on pose en tas pour compter le bétail qui se rend aux champs. Puis, au retour du troupeau, on enlève un à un les cailloux pour savoir combien de bêtes se sont perdues. On peut ainsi représenter les entiers, par des tas de cailloux. On remarque d'ailleurs qu'avec cette notation l'addition est particulièrement simple (je vous laisse deviner comment faire). Pour le reste, c'est très compliqué. Et comme on a dix doigts, on s'est naturellement mis à compter par tas de 10, puis par tas de 10 tas de 10 etc. La notation décimale était née. Elle permet d'additionner, de multiplier et de soustraire assez rapidement. Je suppose acquis ces algorithmes élémentaires (mais pas si faciles).

Exemple : $999 \times 99 = 1000 \times 99 - 99 = 99000 - 99 = 98901$

Pour compter les pertes, les dettes et pour généraliser la résolution des équations $a + x = 0$ où a est un entier donné et x est l'inconnue, on inventa les nombres entiers relatifs. On distingue alors les entiers positifs dit entiers naturels dont l'ensemble se note $\mathbb{N}$ et les entiers positifs ou négatifs que l'on appelle relatifs et dont l'ensemble se note $\mathbb{Z}$. Les entiers négatifs sont notés avec un signe $-$ (sauf 0 qui est, en France, à la fois positif et négatif).

Exercice 1 : Si a est un entier, $-a$ est-il négatif ?

Pour partager des champs ou des terrains, par exemple lors d'un héritage, les fractions furent inventées. Elles permettent également de résoudre des équations de la forme $ax + b = 0$ avec a et b entiers relatifs. Les mathématiciens grecs pensaient alors qu'on définissait ainsi tous les nombres mais ne parvenaient pas à donner la fraction mesurant la diagonale d'un carré de côté 1 ni celle égale au coefficient de proportionnalité entre la circonférence d'un disque et son diamètre. Les fractions peuvent s'écrire de plusieurs façons comme le quotient de deux entiers. Par exemple $\dfrac{8}{6} = \dfrac{4}{3} = \dfrac{12}{9}$. On choisit en général la fraction avec le dénominateur positif et le plus petit possible. L'ensemble des fractions se note $\mathbb{Q}$.

Remarque : $2 \div 3$ est une opération et non une fraction, le résultat de cette opération est la fraction $\dfrac{2}{3}$.

Multiplier deux fractions est simple : $\frac{a}{b} \times \frac{c}{d} = \frac{ac}{bd}$. En revanche l'addition et la soustraction sont compliquées car il faut trouver, parmi les différents représentants des fractions, des fractions avec le même dénominateur : $\frac{1}{2} - \frac{1}{3} = \frac{3}{6} - \frac{2}{6} = \frac{1}{6}$.

C'est pourquoi les nombres décimaux ont fini par s'imposer. Les mathématiciens arabes ont proposé une première écriture au X^e siècle et ce n'est qu'au XVI^e siècle qu'ils arrivent en Occident. Un nombre décimal peut être considéré comme un entier où on a rajouté, entre deux chiffres, une virgule. Le chiffre situé à gauche de la virgule est celui des unités (choix malheureux car faisant perdre la symétrie entre les dixièmes et les dizaines par exemple).

Exemple : Une dizaine : $10,0 = 10$. Un dixième : $0,1$. Si on mettait un point au dessus du chiffre des unités, cela donnerait : $1\dot{0}$ et $\dot{0}1$.

Tout décimal peut s'écrire comme une fraction mais le contraire n'est pas vrai. L'ensemble des décimaux se note $\mathbb{D}$.

Remarque : Les ordinateurs, les calculatrices calculent avec leurs propres nombres « décimaux » qu'on appelle flottants. Mais pour communiquer avec nous ils utilisent la notation décimale.

Les nombres décimaux, comme les entiers, sont faciles à additionner, à soustraire et à multiplier. De plus on peut avoir une approximation aussi précise que l'on veut d'un quotient de deux décimaux.

Dans ce cours, on utilisera très souvent les pourcentages. Par convention, le signe % est une abréviation de l'écriture fractionnaire avec 100 au dénominateur, par exemple $50\% = \frac{50}{100} = \frac{1}{2} = 0,5$

Exercice 2 : Que vaut 8% de 25 ? Et 25% de 8 ?

2) Puissances de 10

Définition 1.1. *On a la notation suivante pour n entier naturel non nul : $10^n = 10 \times 10 \times \ldots \times 10$ avec n facteurs, ou encore $10^n = 1\underbrace{00\ldots00}_{n \text{ zéros}}$ avec n zéros.*

On définit pour n négatif non nul : $10^{-n} = \frac{1}{10^n}$. On a aussi l'écriture : $10^{-n} = \underbrace{0,000\ldots000}_{n \text{ zéros}}1$, c'est à dire $n-1$ zéros après la virgule.

Par convention $10^0 = 1$

Propriété 1.2. *On a les règles de calcul suivantes, cette fois-ci pour tout entiers relatifs n et m :*

1. $10^n \times 10^m = 10^{n+m}$
2. $\frac{10^n}{10^m} = 10^{n-m}$
3. $(10^n)^m = 10^{nm}$

3) Racines carrées

Définition 1.3. *Soit a un nombre positif. On appelle racine carrée de a, notée $\sqrt{a}$, le nombre positif dont le carré est égal à a : $\sqrt{a}^2 = a$ et $\sqrt{a} \geq 0$.*

Exemples : $\sqrt{1} = 1$, $\sqrt{4} = 2$, $\sqrt{1} \neq -1$, $\sqrt{0,25} = 0,5$.

Propriété 1.4. *Pour tout nombres positifs a et b, on a $\sqrt{ab} = \sqrt{a} \times \sqrt{b}$, et si $b \neq 0$, $\sqrt{\frac{a}{b}} = \frac{\sqrt{a}}{\sqrt{b}}$.*

4) Nombres réels et approximation

Définir rigoureusement les nombres réels n'est pas une chose aisée et largement en dehors des objectifs de ce cours. Pour simplifier, on peut considérer qu'un nombre réel peut être encadré par deux décimaux dont la différence est arbitrairement petite.

Exemple : $3,14 \leq \pi \leq 3,15$.
$3{,}141\,592\,653\,5 \leq \pi \leq 3{,}141\,592\,653\,6$ etc.

Il n'y a pas de notation qui conviennent à tous les réels. Certains sont connus comme π ou e, d'autres s'écrivent à l'aide de racines comme $2-\sqrt{3}$ ou $\sqrt{7-4\sqrt{3}}$ et il est même parfois difficile de dire si deux nombres sont égaux ou non (par exemple les deux derniers cités). En conséquence, pour faire des calculs avec des nombres réels, on utilisera des approximations.

Définition 1.5. *Soit x un réel et $n \in \mathbb{Z}$ (n est souvent négatif). On dit que $a \in \mathbb{R}$ est une approximation de x à 10^n près lorsque*
$a \leq x \leq a + 10^n$ (a est une approximation par défaut de x)
ou bien
$a - 10^n \leq x \leq a$ (a est une approximation par excès de x).

On note : $x \approx a$ à 10^n près.

Remarque : En général a est un nombre décimal, et très souvent on garde trois ou quatre chiffres significatifs.
Exemple : $1234^2 = 1\,522\,756 \approx 1\,520\,000$ à 10^5 près.
Exemple : $0{,}125^2 = 0{,}015\,625 \approx 0{,}0156$ à 10^{-4} près.
Exercice 3 : Donner deux approximations décimales de π à 10^{-4} près.

5) Notation scientifique

Théorème 1.6. *Tout nombre réel non nul peut s'écrire de façon unique : $x = \pm z \times 10^n$ avec $z \in [1, 10[$ et $n \in \mathbb{Z}$.*
Cette écriture est appelée notation scientifique.

Exemple : : $0{,}0012 = 1,2 \times 10^{-3}$.

III Proportionnalité

Définition 1.7. *Une grandeur Y est proportionnelle à une grandeur X s'il existe une constante $k \neq 0$ telle que : Pour chaque valeur x de la grandeur X, la valeur correspondante y de Y vérifie $y = kx$.*

Le nombre k s'appelle le coefficient de proportionnalité. On a alors

Propriété 1.8. *Si Y est proportionnel à X, alors X est proportionnel à Y. Le coefficient de proportionnalité est alors $\frac{1}{k}$.*

On peut utiliser un tableau de proportionnalité :

Y	y_1	y_2	y_3
X	x_1	x_2	x_3

Exemple :

Y= distance en km	10	15	25
X=durée en h	2	3	5

Trouver un coefficient de proportionnalité Dans un tableau de proportionnalité, on a toujours :

$$y_1 = kx_1,\ y_2 = kx_2,\ y_3 = kx_3,\ \ldots$$

k vérifie donc

$$\frac{y_1}{x_1} = \frac{y_2}{x_2} = \frac{y_3}{x_3} = \ldots = k$$

Exemple : Dans l'exemple précédent, on a :

$$\frac{10}{2} = \frac{15}{3} = \frac{25}{5} = 5$$

Résoudre un problème de proportionnalité Un problème de proportionnalité peut s'écrire

Y	y_1	y_2
X	x_1	x_2

où x_1, y_1, y_2 sont connues et x_2 est inconnue. Il suffit donc d'écrire, dans le premier cas par exemple :

$$\frac{y_1}{x_1} = \frac{y_2}{x_2} \iff y_1 \times x_2 = x_1 \times y_1 \iff x_2 = \frac{x_1 \times y_2}{y_1}$$

On appelle cette méthode « les produits en croix ». Il existe d'autres méthodes (passage à l'unité, calcul du coefficient de proportionnalité, combinaison de colonnes ou de lignes ...)
Exercice 4 : Un de mes amis professeur d'EPS m'a demandé récemment à quelle distance installer des piquets de telle façon qu'un coureur, allant à une vitesse constante de 12km h^{-1}, franchisse les piquets toutes les minutes. Et si le coureur arrive à maintenir cette vitesse pendant tout un marathon (42,195km), quel sera son temps final ?

Solution : On a affaire au tableau de proportionnalité suivant :

Y= distance	12km	y_1	42,195km
X=durée	1h	1min	x_2

$$y_1 = \frac{12\text{km} \times 1\text{min}}{1\text{h}} = \frac{12\,000\text{m} \times 1\text{min}}{60\text{min}} = \frac{12\,000\text{m}}{60} = 200\text{m}$$

$$x_2 = \frac{1\text{h} \times 42{,}195\text{km}}{12\text{km}} \approx 3{,}52\text{h} \approx 3\text{h}31\text{min}$$

On remarque la cohérence des résultats en terme d'unités (homogénéité).

IV Résolution des équations du second degré

Un peu de vocabulaire : $ax^2 + bx + c$ avec $a \neq 0$ est un **trinôme** (polynôme de degré 2), et les solutions de l'équation $ax^2 + bx + c = 0$ sont appelés les **racines du polynôme**.
On dit que r est racine d'un polynôme P lorsqu'on peut mettre $(x - r)$ en facteur dans P.
On dit que r est racine double d'un polynôme P lorsqu'on peut mettre $(x - r)^2$ en facteur dans P.

1) Résolution d'équations du second degré

Exemple : Résolution de l'équation : $x^2 + 2x - 35 = 0$. On écrit :

$$\begin{aligned} x^2 + 2x - 35 &= (x^2+1)^2 - 36 = (x^2+1)^2 - 6^2 \\ &= (x+1+6)(x+1-6) = (x+7)(x-5) \end{aligned}$$

On a donc trouvé deux solutions : $x = -7$ et $x = 5$.
Nous donnons maintenant les formules générales : Soit à résoudre l'équation $ax^2 + bx + c = 0$ avec $a \neq 0$

Théorème 1.9. *On pose $\Delta = b^2 - 4ac$. Alors on a :*

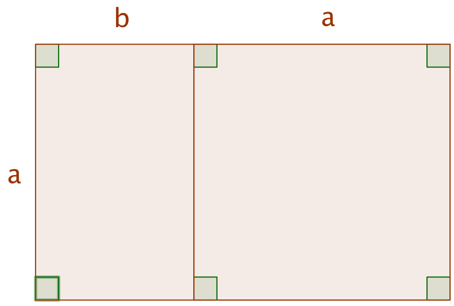
1. *Si $\Delta > 0$, on a deux solutions distinctes :* $x_1 = \frac{-b+\sqrt{\Delta}}{2a}$ et $x_2 = \frac{-b-\sqrt{\Delta}}{2a}$
2. *Si $\Delta = 0$, on a une solution unique $x_1 = -\frac{b}{2a}$*
3. *Si $\Delta < 0$, il n'y a pas de solution réelle.*

La démonstration est semblable à celle donnée dans le cas particulier.
Exemples :

1. Soit $x^2 - x - 1 = 0$. Alors $\Delta = 5$ et on a deux solutions distinctes : $x_1 = \frac{1+\sqrt{5}}{2}$ et $x_2 = \frac{1-\sqrt{5}}{2}$.
2. Soit $x^2 + x + 2 = 0$. Alors $\Delta = -7$ et il n'y a pas de solution.
3. Soit $x^2 - 4x + 4 = 0$. Alors $\Delta = 0$ et on a une unique solution : $x_1 = 2$.

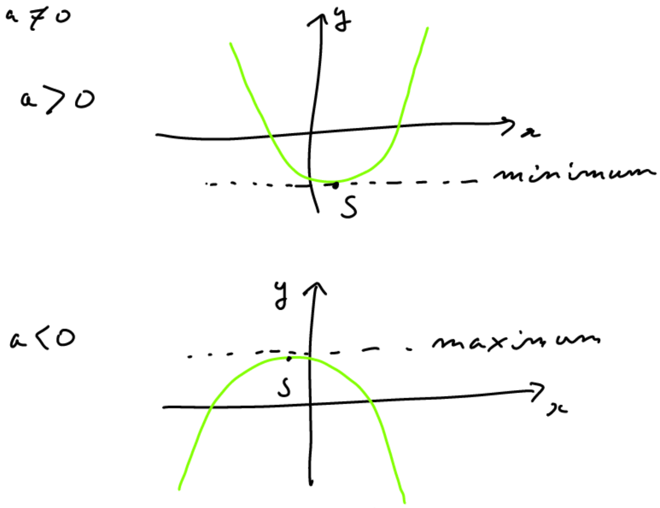
Exercice 5 : Le nombre d'or, ou proportion divine, est le réel Φ donnant le rapport entre le grand côté, a, d'un rectangle et son petit côté, b alors le rectangle construit en ajoutant un carré de côté a au rectangle initial garde les mêmes proportions. Trouvez sa valeur.

Solution :



$\Phi = \frac{a}{b} = \frac{a+b}{a} = 1+\frac{1}{\Phi}$. Donc $\Phi^2 - \Phi - 1 = 0$. L'équation est résolue dans les exemples précédents, on garde bien sûr la racine positive $\boxed{\phi = \dfrac{1+\sqrt{5}}{2}}$.

2) Courbe représentative de $x \to ax^2 + bx + c$



La courbe représentative s'appelle une parabole. Il y a deux cas. Si $a > 0$, le sommet est un minimum. Si $a < 0$, le sommet est un maximum.

3) Inéquations du second degré

On cherche à résoudre l'inéquation : $ax^2 + bx + c \geq 0$.
Cela revient á étudier le signe du polynôme $P(x) = ax^2 + bx + c$.
On doit considérer les cas où :

- $\Delta < 0$, le polynôme P ne s'annule jamais et est donc du signe de a.
- $\Delta = 0$, le polynôme P admet une unique racine r et peut se factoriser sous la forme $P(X) = a(X-r)^2$; P est donc du signe de a et s'annule en r.
- $\Delta > 0$, le polynôme P admet deux racines distinctes r_1 et r_2 ; il peut se factoriser sous la forme $P(x) = a(x-r_1)(x-r_2)$; P est donc du signe de a « à l'extérieur des racines ».

Chapitre 2

Étude de fonctions

I Limites, opérations sur les limites, comparaison

1) Limite infinie en l'infini

Définition 2.1.

1. *La fonction f tend vers $+\infty$ quand $x \to +\infty$ (resp. $x \to -\infty$) si et seulement si tout intervalle $]\lambda, +\infty[$, ($\lambda \in \mathbb{R}$) contient toutes les valeurs de $f(x)$ pour x assez grand (resp. $-x$ assez grand). On note* $\lim_{x\to+\infty} f(x) = +\infty$ *(resp.* $\lim_{x\to-\infty} f(x) = +\infty$*).*
2. *La fonction f tend vers $-\infty$ quand $x \to +\infty$ (resp. $x \to -\infty$) si et seulement si tout intervalle $]-\infty, \lambda[$, ($\lambda \in \mathbb{R}$) contient toutes les valeurs de $f(x)$ pour x assez grand (resp. $-x$ assez grand). On note* $\lim_{x\to+\infty} f(x) = -\infty$ *(resp.* $\lim_{x\to-\infty} f(x) = -\infty$*).*

2) Limite finie en l'infini

Définition 2.2. *Soit ℓ un réel, une fonction f tend vers ℓ quand x tend vers $+\infty$ (resp. $-\infty$) si et seulement si tout intervalle ouvert contenant ℓ contient aussi toutes les valeurs de $f(x)$ pour x assez grand (resp. $-x$ assez grand). On note alors* $\lim_{x\to+\infty} f(x) = \ell$ *(resp.* $\lim_{x\to-\infty} f(x) = \ell$*). Si* $\lim_{x\to+\infty} f(x) = \ell$ *(resp.* $\lim_{x\to-\infty} f(x) = \ell$*), alors la droite d'équation $y = l$ est asymptote à la courbe représentative de f en $+\infty$ (resp. en $-\infty$).*

3) Limite finie ou infinie d'une fonction en a, ($a \in \mathbb{R}$)

Soient a un réel et un intervalle I contenant a ou dont a est une borne, f une fonction définie dans I sauf peut-être en a.

Définition 2.3. *On dit que la fonction f tend vers ℓ quand x tend vers a si et seulement si tout intervalle ouvert contenant ℓ contient aussi toutes les valeurs de $f(x)$ pour tout réel x dans I et assez proche de a.*

Définition 2.4. *On dit que la fonction f tend vers $+\infty$ (resp. $-\infty$) quand x tend vers a si et seulement si tout intervalle $]\lambda, +\infty[$ (resp. $]-\infty, \lambda[$) ($\lambda \in \mathbb{R}$)), contient toutes les valeurs de $f(x)$ pour tout réel x dans I et assez proche de a.*

Définition 2.5. *Lorsque* $\lim_{x\to a} f(x) = +\infty$ *(ou $-\infty$), la droite d'équation $x = a$ est asymptote verticale à la courbe représentative de f.*

4) Asymptote oblique

Définition 2.6. *Si* $\lim_{x\to+\infty}(f(x)-(ax+b))=0$*, alors la droite d'équation* $y=ax+b$ *est asymptote à la courbe représentative de* f *en* $+\infty$*.*
Si $\lim_{x\to-\infty}(f(x)-(ax+b))=0$*, alors la droite d'équation* $y=ax+b$ *est asymptote à la courbe représentative de* f *en* $-\infty$*.*

Méthode
Quand une fonction admet une limite infinie ($\pm\infty$) en l'infini, on peut se poser la question de l'existence d'une asymptote oblique.
On commence par étudier si $x\mapsto\dfrac{f(x)}{x}$ admet une limite finie. Si la réponse est non, il n'y a pas d'asymptote. Si la réponse est oui, on peut noter a cette limite, alors on peut dire que la courbe admet une direction asymptotique donnée par l'équation de droite $y=ax$.
Dans le dernier cas, on étudie alors si $x\mapsto f(x)-ax$ admet une limite finie. Si la réponse est non, il n'y a pas d'asymptote. Si la réponse est oui, on peut noter b cette limite, alors on peut dire que la courbe admet une asymptote oblique d'équation $y=ax+b$.

Exemples
On étudiera les deux fonctions définies pour tout $x\geq 0$ par $f(x)=\dfrac{x(x+\sqrt{x})}{x+1}$ et $g(x)=\dfrac{x^2+\sqrt{x}}{x+1}$.

5) Opérations sur les limites

Table 2.1 – Limite d'une somme

Si f a pour limite	ℓ	ℓ	ℓ	$+\infty$	$-\infty$	$+\infty$
Si g a pour limite	ℓ'	$+\infty$	$-\infty$	$+\infty$	$-\infty$	$-\infty$
Alors $f+g$ a pour limite	$\ell+\ell'$	$+\infty$	$-\infty$	$+\infty$	$-\infty$	?

Table 2.2 – Limite d'un produit

Si f a pour limite	ℓ	$\ell>0$	$\ell>0$	$\ell<0$	$\ell<0$	$+\infty$	$+\infty$	$-\infty$	0
Si g a pour limite	ℓ'	$+\infty$	$-\infty$	$+\infty$	$-\infty$	$+\infty$	$-\infty$	$-\infty$	$\pm\infty$
Alors $f\times g$ a pour limite	$\ell\times\ell'$	$+\infty$	$-\infty$	$-\infty$	$+\infty$	$+\infty$	$-\infty$	$+\infty$	?

Remarque : La limite en ∞ ou en $-\infty$ d'une fonction polynôme est la limite de son terme de plus haut degré.

Table 2.3 – Limite d'un quotient - Cas où le dénominateur a une limite non nulle

Si f a pour limite	ℓ	ℓ	$+\infty$	$+\infty$	$-\infty$	$-\infty$	$\pm\infty$
Si g a pour limite	$\ell'\neq 0$	$\pm\infty$	$\ell'>0$	$\ell'<0$	$\ell'>0$	$\ell'<0$	$\pm\infty$
Alors $\frac{f}{g}$ a pour limite	$\frac{\ell}{\ell'}$	0	$+\infty$	$-\infty$	$-\infty$	$+\infty$	?

Table 2.4 – Limite d'un quotient - Cas où le dénominateur a une limite nulle

Si f a pour limite	$\ell>0$	$\ell>0$	$\ell<0$	$\ell<0$	0
Si g a pour limite	0 en restant positive	0 en restant négative	0 en restant positive	0 en restant négative	0
Alors $\frac{f}{g}$ a pour limite	$+\infty$	$-\infty$	$-\infty$	$+\infty$	?

Remarque : La limite en $+\infty$ ou en $-\infty$ d'une fonction rationnelle est égale à la limite du quotient des termes du plus haut degré.

II Continuité

I désigne un intervalle non vide non réduit à une seule valeur et a un élément de I. On note $f(I)$ l'ensemble des valeurs de $f(x)$ quand x varie dans I.

Définition 2.7. *Soit f une fonction définie sur I. On dit que f est continue en a si f admet une limite en a telle que :* $\lim_{x\to a} f(x) = f(a)$

On dira que f est continue sur I si f est continue en tout point de I.

Exemples : Les fonctions rationnelles (dont les fonctions polynômes et la fonction inverse), la fonction racine carrée et les fonctions trigonométriques (sinus, cosinus et tangente) sont continues sur tout intervalle inclus dans leur ensemble de définition.

Propriétés Soient f deux fonctions définies sur I et soit α un réel.

1. Si f est continue en a (resp. sur I) alors αf est continue en a (resp.sur I).
2. Si f et g sont continues en a (resp. sur I) alors $f+g$ est continue en a (resp. sur I).
3. Si f et g sont continues en a (resp. sur I) alors $f\times g$ est continue en a (resp. sur I).
4. Si f et g sont continues en a et $g(a)\neq 0$ alors $\frac{f}{g}$ est continue en a.

Remarque : La notion intuitive de continuité de *"on ne lève pas le crayon"* n'est vraie que pour des fonctions continues sur un intervalle. Il existe par exemple des fonctions continues sur $\mathbb{R}\setminus\mathbb{Q}$ et discontinue sur $\mathbb{Q}$.

Composée de deux fonctions.
Soient f une fonction définie d'un intervalle I vers un intervalle J et g une fonction définie de J vers $\mathbb{R}$. La composée de f et g notée $g\circ f$ est définie par $(g\circ f)(x) = g(f(x))$.
Si f est continue en a et si g est continue en $f(a)$ alors $g\circ f$ est continue en a. Si f est continue sur I et si g est continue sur $f(I)$ alors $g\circ f$ est continue sur I.

III Dérivabilité

1) Nombre dérivé

Soit f une fonction définie sur un intervalle I et a un élément de I. Si la limite $\lim_{h\to 0,\, h\neq 0} \frac{f(a+h)-f(a)}{h}$ existe, on la note $f'(a)$ et on l'appelle **nombre dérivé de la fonction f en** a. Dans ce cas, on dit que la fonction f est dérivable en a.

Interprétation géométrique
Soit f une fonction définie sur un intervalle I et a un élément de I. Soit C_f la courbe représentative de f dans un repère. Si f est dérivable en a alors C_f admet une tangente au point $(a; f(a))$ et une équation de cette tangente est : $y = f'(a)(x-a) + f(a)$
On en tire l'équation réduite de la tangente : $y = f'(a).x + f(a) - f'(a).a$

2) Dérivabilité, continuité

On a la propriété suivante :

Propriété 2.8. *Soit f une fonction définie sur un intervalle I et a un élément de I. Si f est dérivable en a (sur I) alors f est continue en a (sur I).*

Remarque : la réciproque est fausse (considérer la fonction valeur absolue en 0).

3) Fonction dérivée et sens de variation

Soit f une fonction définie sur un intervalle I. Si f est dérivable pour tout x de I, alors on dit que la fonction f est dérivable sur I et on note f' la fonction définie par :

$$\begin{array}{l} I \to \mathbb{R} \\ x \mapsto f'(x) \end{array}$$

La fonction f' est appelée fonction dérivée de la fonction f.

On a le théorème important suivant

Théorème 2.9. *Soit f une fonction définie et dérivable sur un intervalle I.*
Si f' est de signe constant sur I alors f est monotone sur I.
Si pour tout réel $x \in I$, $f'(x) \geq 0$ alors f est croissante sur I.
Si pour tout réel $x \in I$, $f'(x) \leq 0$ alors f est décroissante sur I.
Si pour tout réel $x \in I$, $f'(x) > 0$ alors f est strictement croissante sur I.
Si pour tout réel $x \in I$, $f'(x) < 0$ alors f est strictement décroissante sur I.
Si pour tout réel $x \in I$, $f'(x) = 0$ alors f est constante sur I.

Remarque : Attention, dans le théorème précédent il est important d'avoir un signe constant de la dérivée sur tout un intervalle. Il existe des fonctions dérivables sur $\mathbb{R}$, avec en un point x_0 que $f'(x_0) > 0$ mais qui ne sont croissantes sur aucun intervalle contenant x_0.

4) Fonctions dérivées des fonctions usuelles

Les fonctions rationnelles (dont les fonctions polynômes et la fonction inverse), la fonction racine carrée et les fonctions trigonométriques sont dérivables sur tout intervalle inclus dans leur ensemble de définition (ATTENTION ! Ce résultat n'est pas valable pour la fonction racine carrée en 0).

TABLE 2.5 – Dérivées des fonctions usuelles

Fonction	Dérivée	Intervalle I (maximal)
$x \mapsto k, (k \in \mathbb{R})$	$x \mapsto 0$	$\mathbb{R}$
$x \mapsto x$	$x \mapsto 1$	$\mathbb{R}$
$x \mapsto x^2$	$x \mapsto 2x$	$\mathbb{R}$
$x \mapsto \frac{1}{x}$	$x \mapsto -\frac{1}{x^2}$	$\mathbb{R}_+^*$ ou $\mathbb{R}_-^*$
$x \mapsto \sqrt{x}$	$x \mapsto \frac{1}{2\sqrt{x}}$	$\mathbb{R}_+^*$
$x \mapsto x^n$ $n \in \mathbb{Z}^*$	$x \mapsto nx^{n-1}$	$\mathbb{R}$ (si $n > 0$) $\mathbb{R}_+^*$ ou $\mathbb{R}_-^*$ (si $n < 0$)
$x \mapsto \sin x$	$x \mapsto \cos x$	$\mathbb{R}$
$x \mapsto \cos x$	$x \mapsto -\sin x$	$\mathbb{R}$
$x \mapsto \tan x$	$x \mapsto 1 + \tan^2 x = \frac{1}{\cos^2 x}$	$\mathbb{R} \setminus \{(2k+1)\frac{\pi}{2},\, k \in \mathbb{Z}\}$

5) Opérations et dérivation

Si f et g sont deux fonctions dérivables sur un intervalle I, et si k est un réel alors :

1. La fonction $f + g$ est dérivable sur I et on a : $(f + g)' = f' + g'$
2. La fonction kf est dérivable sur I et on a : $(kf)' = kf'$
3. La fonction fg est dérivable sur I et on a : $(fg)' = f'g + fg'$

De plus, si g ne s'annule pas sur I,

4. La fonction $\frac{1}{g}$ est dérivable sur I et on a : $(\frac{1}{g})' = -\frac{g'}{g^2}$
5. La fonction $\frac{f}{g}$ est dérivable sur I et on a : $(\frac{f}{g})' = \frac{f'g-fg'}{g^2}$

6) Dérivée d'une fonction composée

Soit f une fonction dérivable sur un intervalle I et prenant ses valeurs dans un ensemble inclus dans un intervalle J. Soit g une fonction dérivable sur l'intervalle J. Dans ces conditions, la fonction $g \circ f$ est dérivable sur l'intervalle I et on a : $(g \circ f)' = f' \times (g' \circ f)$ Soit : $(g \circ f)'(x) = f'(x) \times (g' \circ f)(x) = f'(x) \times g'(f(x))$

On a en particulier :
Soit a et b deux réels, a étant non nul. Soit f définie et dérivable sur I et soit g la fonction définie par $g : x \mapsto f(ax+b)$ (elle est donc définie pour tout x tel que $ax+b$ appartient é I). Dans ces conditions, la fonction g est dérivable et on a : $g'(x) = af'(ax+b)$ (la fonction g est la composée des fonctions $x \mapsto ax+b$ et f).
Formulaire
Pour toute fonction u définie et dérivable sur un intervalle I (et, éventuellement, ne s'annulant pas sur I), on a :

Fonction	Dérivée
$x \mapsto u^n(x)$	$n.u'(x).u^{n-1}(x)$
$x \mapsto \sqrt{u(x)}$	$x \mapsto \frac{u'(x)}{2\sqrt{u(x)}}$

Remarque : dans le deuxième cas, la fonction u est à valeurs strictement positives.

IV Plan d'étude d'une fonction

On considère la fonction : $x \to f(x)$. Le plan d'étude se déroule de la façon suivante :
1. On détermine l'ensemble de définition.
2. On cherche si la courbe présente de propriétés particulières (parité, périodicité).
3. On détermine les variations de f (à l'aide du signe le la dérivée par exemple).
4. On cherche les limites aux bornes du domaine de définition et on écrit le tableau de variation.
5. On détermine les asymptotes.
6. On trace la courbe représentative de f.

Exemple Etudier en suivant ce plan la fonction définie par $f(x) = \sqrt{x^2+1}$.

V Les fonctions logarithmes et exponentielle

1) Fonction logarithme

Elle est notée $\ln(x)$. Son ensemble de définition est $\mathbb{R}_+^*$. Cette fonction est continue et dérivable sur son ensemble de définition. On a $\ln'(x) = \frac{1}{x}$. Les limites suivantes sont à connaître :

$$\lim_{x\to+\infty} \ln(x) = +\infty,\ \lim_{x\to0^+} \ln(x) = -\infty$$

$$\lim_{x\to+\infty} \frac{\ln(x)}{x^n} = 0,\ (n \in \mathbb{N}^*),\ \lim_{x\to0^+} x^n \ln(x) = 0,\ (n \in \mathbb{N}^*)$$

$$\lim_{x\to0} \frac{\ln(1+x)}{x} = 1$$

On a les propriétés suivantes :

$$\ln(1) = 0,\ \ln(a.b) = \ln(a) + \ln(b),\ \ln(\frac{a}{b}) = \ln(a) - \ln(b),\ \ln(x^n) = n\ln(x),\ (n \in \mathbb{N})$$

2) La fonction exponentielle

Elle est notée $exp(x)$ ou e^x. Son ensemble de définition est $\mathbb{R}$. Cette fonction est continue et dérivable sur son ensemble de définition. On a $exp'(x) = exp(x)$. Les limites suivantes sont à connaître :

$$\lim_{x\to+\infty} \mathrm{e}^x = +\infty, \ \lim_{x\to-\infty} \mathrm{e}^x = 0$$

$$\lim_{x\to+\infty} \frac{\mathrm{e}^x}{x^n} = +\infty, \ (n \in \mathbb{N}^*), \ \lim_{x\to-\infty} x^n\mathrm{e}^x = 0, \ (n \in \mathbb{N}^*)$$

$$\lim_{x\to 0} \frac{\mathrm{e}^x - 1}{x} = 1$$

On a les propriétés suivantes :

$$\mathrm{e}^0 = 1, \ \mathrm{e}^{(a+b)} = \mathrm{e}^a.\mathrm{e}^b, \ \mathrm{e}^{a-b} = \frac{\mathrm{e}^a}{\mathrm{e}^b}, \ \mathrm{e}^{na} = (\mathrm{e}^a)^n$$

3) Exemples

Étude de la fonction définie par $f(x) = \frac{\ln(x)}{x}$

Étude de la fonction définie par $f(x) = \ln(x + \sqrt{x^2+1})$

Étude de la fonction définie par $f(x) = \mathrm{e}^{2x} - \mathrm{e}^x$

VI Calcul de primitives, primitives usuelles, intégrales généralisées

1) Intégrale sur un segment

Soit f une fonction définie sur un segment $[a, b]$. On suppose que f est continue et positive. On considère alors sa courbe $\mathcal{C}_f$ dans un repère orthonormé.

Définition 2.10. *On appelle* $\displaystyle\int_a^b f(t)\,dt$ *l'aire située entre la courbe de f, l'axe des abscisses et les droites d'équations $x = a$ et $x = b$.*

Exemple : $\displaystyle\int_{2,5}^3 2t - 3\,dt$

Interprétation algébrique de $\int_a^b f(t)\,dt$. Le signe $\int$ signifie « somme ».
$\displaystyle\int_a^b$ signifie qu'on va sommer pour t qui va de a vers b.
dt est le pas infinitésimal.
Par exemple, avec un pas égal à $0,01$, on peut écrire que :

$$\int_1^2 t^2\,dt \approx 1^2 \times 0,01 + 1,01^2 \times 0,01 + 1,02^2 \times 0,01 + \cdots + 1,99^2 \times 0,01.$$

Conséquence : On peut généraliser l'intégrale avec des fonctions pas forcément positives et/ou avec $a > b$.

Propriété 2.11. *Linéarité, Chasles, cas où $a = b$.*

2) Primitives

Définition 2.12. *Soit f définie sur un intervalle I. On appelle primitive de f sur I toute fonction F dérivable telle que $F' = f$ sur I.*

Exemple : $f(x) = x$, $F(x) = \dfrac{x^2}{2} + 5$

Propriété 2.13.
1. *Soient F et G deux primitives sur I d'une même fonction f. Il existe une constante k telle que $F = G + k$*
2. *Soit F une primitive de f sur I. Pour tout $k \in \mathbb{R}$, $F + k$ est également une primitive de f*

Conséquence : Si on connait une primitive F de f sur un intervalle I, l'ensemble des primitives de f sur I est l'ensemble contenant les fonctions $F + k$ avec $k \in \mathbb{R}$.

Propriété 2.14. *Soit $a \in I$.*

$$F(x) = \int_a^x f(t)\,dt$$

F est LA primitive de f qui s'annule en a.

Idée de la preuve. On va montrer que le taux d'accroissement de F en x : $\dfrac{F(x+h) - F(x)}{h}$ est proche de $f(x)$ pour h petit. En effet

$$F(x+h) - F(x) = \int_a^{x+h} f(t)\,dt - \int_a^x f(t)\,dt = \int_a^{x+h} f(t)\,dt + \int_x^a f(t)\,dt \overset{Chasles}{=} \int_x^{x+h} f(t)\,dt$$

Or, f est continue et h est « petit », de sorte qu'on peut considérer f constante entre x et $x + h$, et on a donc $\displaystyle\int_x^{x+h} f(t)\,dt \approx h \times f(x)$. On trouve bien ainsi que le taux d'accroissement se rapproche de $f(x)$), d'où $F' = f$.

Propriété 2.15 (fondamentale)**.** *Soit F une primitive de f sur I et a, b dans I.*

$$\int_a^b f(t)\,dt = F(b) - F(a)$$

Notation : On notera $[F(t)]_a^b = F(b) - F(a)$

3) Primitives usuelles

Voir la table 2.6.

4) Outils de calculs de primitives

Linéarité : primitive de la somme et d'un multiple.

Propriété 2.16 (Intégration par parties, IPP)**.** *Soient u et v deux fonctions dérivables sur un intervalle I et $(a, b) \in I^2$. ALors,*

$$\int_a^b u'(t)v(t)\,dt = [u(t) \times v(t)]_a^b - \int_a^b u(t)v'(t)\,dt$$

Exemple : primitive de la fonction logarithme.

TABLE 2.6 – Primitives des fonctions usuelles

$f(x)$	$F(x)$	Intervalle I (maximal)
0	$k,\ k \in \mathbb{R}$	$\mathbb{R}$
1	$x + k,\ k \in \mathbb{R}$	$\mathbb{R}$
$x^n,\ n \in \mathbb{N}$	$\dfrac{x^{n+1}}{n+1} + k,\ k \in \mathbb{R}$	$\mathbb{R}$
$\dfrac{1}{x}$	$\ln\lvert x\rvert + k,\ k \in \mathbb{R}$	$\mathbb{R}_+^*$ ou $\mathbb{R}_-^*$
$\dfrac{1}{x^n},\ n \in \mathbb{N},\ n \geq 2$	$\dfrac{1}{(1-n)(x^{n-1})} + k,\ k \in \mathbb{R}$	$\mathbb{R}_+^*$ ou $\mathbb{R}_-^*$
$x^\alpha = e^{\alpha \ln(x)},\ \alpha \in \mathbb{R},\ \alpha \neq -1$	$\dfrac{x^{\alpha+1}}{\alpha+1} + k,\ k \in \mathbb{R}$	$\mathbb{R}_+^*$
$\sin x$	$-\cos x + k,\ k \in \mathbb{R}$	$\mathbb{R}$
$\cos x$	$\sin x + k,\ k \in \mathbb{R}$	$\mathbb{R}$
e^x	$e^x + k,\ k \in \mathbb{R}$	$\mathbb{R}$

5) Inégalités et intégrales

Propriété 2.17. *Si $a \leq b$ et si pour tout $t \in [a,b]$ on a $f(t) \leq g(t)$ alors $\int_a^b f(t)\,dt \leq \int_a^b g(t)\,dt$.*

Conséquence : Si f est majorée par M sur $[a,b]$ alors $\int_a^b f(t)\,dt \leq M(b-a)$.
De même, si f est minorée par m sur $[a,b]$, $m(b-a) \leq \int_a^b f(t)\,dt$.

Définition 2.18. *On appelle* ***valeur moyenne*** *de f sur $[a,b]$ le réel $\dfrac{1}{b-a}\int_a^b f(t)\,dt$*

6) Intégrales généralisées

On définit juste $\int_A^{+\infty} f(t)\,dt$, $\int_{-\infty}^{A} f(t)\,dt$, et $\int_{-\infty}^{+\infty} f(t)\,dt$.
Interprétation avec les aires

Exemple : sur $\mathbb{R}_+$ $f(t) = 1/t^2$, $f(t) = \lambda e^{-\lambda t}$, et sur $\mathbb{R}$: $f(t) = \dfrac{e^{-t^2/2}}{\sqrt{2\pi}}$

VII Équations différentielles linéaires à coefficients constants

1) Introduction - Vocabulaire

Exemple : Beaucoup d'êtres vivants ont une croissance proportionnelle à leur taille. Si on note $T(t)$ la taille au temps T, il existe alors une constante K telle que $T'(t) = KT(t)$ pour tout t (t en jours).
Sachant qu'au temps $t = 0$ la taille vaut $2cm$ et que la taille double en une semaine : $T(7) = 4cm$, quelle taille aura-t-on au bout de 4 semaines ? (résolu en fin du 2))

Définition 2.19. *Résoudre une équation d'inconnue x, c'est trouver toutes les valeurs de x pour lesquelles l'équation est vérifiée, c'est à dire pour lesquelles on a l'égalité entre les deux membres de l'équation.*
Une équation ***différentielle*** *est une équation où l'inconnue est une fonction dérivable, et fait intervenir la fonction inconnue et quelques unes de ses dérivées.*

Exemples : $y' = y$, $y''y' + y^2 = \cos x$.

Définition 2.20. *L'ordre d'une équation différentielle est le grand nombre de fois que la fonction inconnue est dérivée dans l'équation.*

Exemples : Dans les exemples précédents, la première équation est d'ordre 1 et la deuxième d'ordre 2.
Remarque : Dans ce cours on étudiera uniquement les équations différentielles linéaires suivantes (d'ordre 1 et 2) :

$$y' + ay = g(x)$$
$$y'' + by + cy = f(x)$$

où a, b, c sont des constantes et f et g des fonctions continues.

2) Equations différentielles linéaires du premier ordre à coefficient constant

Il s'agit de résoudre une équation de la forme :

$$(E) \quad y' + ay = g(x)$$

où a est une constante réelle, b une fonction connue et continue, et y est la fonction inconnue qui doit être dérivable.

Première étape On résout l'équation homogène associée, qu'on appelle aussi l'équation sans second membre (ESSM) :

$$(E_0) \quad y' + ay = 0$$

Propriété 2.21. *Les solutions de l'équation homogène sont les fonctions de la forme $y = \lambda e^{-ax}$ où $\lambda \in \mathbb{R}$. (à savoir par cœur)*

Démonstration : Posons $z(x) = y(x)e^{ax}$. On va montrer que z est une fonction constante ssi y est solution de (E_0).
Comme produit de deux fonctions dérivables, z est dérivable et on a :

$$z'(x) = y'(x)e^{ax} + y(x) \times ae^{ax} = e^{ax}\left(y'(x) + ay(x)\right)$$

Il est alors facile de montrer que y solution de (E_0) entraîne $z' = 0$. La réciproque est juste également car $e^{ax} > 0$ pour tout x.
On a donc y solution de $(E_0) \iff z' = 0 \iff z = \lambda$ avec $\lambda \in \mathbb{R} \iff ye^{ax} = \lambda \iff y = \lambda e^{-ax}$

Deuxième étape On cherche une solution particulière de l'équation de départ (E).
Soit on trouve cette solution rapidement (donnée par l'énoncé ou évidente),
soit on utilise la méthode dite de la « variation de la constante » : on cherche une solution de la forme $y_p(x) = \lambda(x)e^{-ax}$. où λ est une fonction dérivable. y_p est alors solution ssi

$$y_p' + ay_b = g(x) \iff \lambda'(x)e^{ax} + \lambda(x) \times (-ae^{-ax}) + a\lambda(x)e^{-ax} = g(x) \iff \lambda'(x)e^{-ax} = g(x)$$

Il ne reste plus qu'à trouver une primitive de $x \mapsto g(x)e^{ax}$. Appelons cette primitive λ_p, alors $y_p = \lambda_p(x)e^{-ax}$ est une solution particulière.
Exemple : Solution particulière de $(E_1) \quad y' - 3y = 1 - 3x$ et de $(E_2) \quad y' - 3y = e^{2x}$

Propriété 2.22 (Principe de superposition)**.** *Si y_1 est solution particulière de l'équation $(E_1) \quad y' + ay = g_1(x)$ et y_2 est solution particulière de $(E_2) \quad y' + ay = g_2(x)$ alors $y_1 + y_2$ est solution particulière de $(E) \quad y' + ay = g_1(x) + g_2(x)$*

Troisième étape On termine la résolution.

Propriété 2.23. *Soit y_p une solution particulière de (E). Alors y est solution de (E) ssi $y - y_p$ est solution de (E_0) ssi $y = y_0 + y_p$ avec y_0 solution de (E_0).*

Exemple : Résolution de $(E_) \quad y' - 3y = 1 - 3x + e^{2x}$.

Propriété 2.24. *Résolution de l'équation avec une condition initiale $y(x_0) = \mu_0$. On montre qu'il existe une unique solution avec cette condition intitiale.*

Exemple : Résolution de l'exemple donné au début.

3) Equations différentielles linéaires d'ordre 2 à coefficients constants

Il s'agit ici de résoudre des équations de la forme

$$y'' + by' + cy = f(x)$$

où f est une fonction donnée continue et b et c sont des réels.

Nous allons procéder de la même façon que pour les équations différentielles d'ordre 1 :

1. Résolution de l'équation homogène ou sans second membre
2. Recherche d'une solution particulière de l'équation initiale
3. Résolution complète

a) Résolution de l'équation homogène ou sans second membre

On va résoudre ici l'équation différentielle :

$$y'' + by' + cy = 0 \quad (E_0)$$

Propriété 2.25. *$y_0 = e^{\lambda x}$ est solution de (E_0) ssi λ est solution de l'équation $x^2 + bx + c = 0$.*

Démonstration : Posons $y_0 = e^{\lambda x}$. On calcule facilement que $y_0' = \lambda e^{\lambda x}$ et $y_0'' = \lambda^2 e^{\lambda x}$. De sorte que y_0 est solution de l'équation (E_0) ssi $\lambda^2 e^{\lambda x} + b\lambda e^{\lambda x} + e^{\lambda x} = 0$. Ce qui équivaut, en mettant $e^{\lambda x}$ en facteur, à $e^{\lambda x}(\lambda^2 + b\lambda + c) = 0$. Comme $e^{\lambda x}$ n'est jamais nul on en déduit que y_0 est solution de (E_0) ssi λ est solution de $x^2 + bx + c = 0$.

Définition 2.26. *L'équation $x^2 + bx + c = 0$ d'inconnue $x \in \mathbb{R}$ est appelée **équation caractéristique** associée à l'équation différentielle.*

Exemple : L'équation caractéristique de $y'' + y' - 6y = 0$ est $x^2 + x - 6 = 0$.

Propriété 2.27 (Lemme technique à ne pas apprendre)**.** *Soit λ une solution de l'équation caractéristique. Soit y_0 une fonction deux fois dérivable. On pose alors $z = y_0/e^{\lambda x}$, de sorte que $y_0 = ze^{\lambda x}$. Alors, y_0 est solution de (E_0) ssi z' est solution de l'équation différentielle $y' + (2\lambda + b)y = 0$, autrement dit : ssi $z'' + (2\lambda + b)z' = 0$.*

Démonstration : On a $y_0' = z'e^{\lambda x} + \lambda ze^{\lambda x}$ et $y_0'' = z''e^{\lambda x} + \lambda z'e^{\lambda x} + \lambda z'e^{\lambda x} + \lambda^2 ze^{\lambda x} = e^{\lambda x}(z'' + 2\lambda z' + \lambda^2 z)$. Donc y_0 est solution de (E_0) ssi on a :

$$e^{\lambda x}(z'' + 2\lambda z' + \lambda^2 z + bz' + b\lambda z + cz) = 0$$

Comme $e^{\lambda x} \neq 0$ pour tout x, cela équivaut également à :

$$z'' + (2\lambda + b)z' + (\lambda^2 + b\lambda + c)z = 0$$

Or λ est racine de $x^2 + bx + c = 0$, et donc le terme en facteur de z s'annule et on obtient donc bien l'équivalence demandée.

Propriété 2.28 (A connaître !)**.** *Si $\Delta = b^2 - 4c > 0$ alors on appelle λ_1 et λ_2 les solutions distinctes de l'équation caractéristique et les solutions de l'équation (E_0) (équation homogène) sont les fonctions de la forme $y = k_1 e^{\lambda_1 x} + k_2 e^{\lambda_2 x}$ avec k_1 et k_2 constantes réelles.*

Démonstration : On se sert de la propriété précédente, en prenant pour λ la racine λ_1.

Résolvons alors l'équation différentielle $y' + (2\lambda_1 + b)y = 0$. D'après la résolution faite dans la section précédente, on sait que cela équivaut à $y = ke^{(-2\lambda 1 - b)x}$ où $k \in \mathbb{R}$.

Or, $\lambda_1 + \lambda_2 = -b$, on en déduit que $-2\lambda_1 - b = -2\lambda_1 + (\lambda_1 + \lambda_2) = \lambda_2 - \lambda_1$. Ainsi, z' solution de cette équation différentielle équivaut à $z = \frac{k}{\lambda_2 - \lambda_1} e^{(\lambda_2-\lambda_1)x} + k_1$ avec k et k_1 constantes réelles.

Lorsque k décrit $\mathbb{R}$, $\frac{k}{\lambda_2-\lambda_1}$ décrit $\mathbb{R}$ également donc, en posant $k_2 = \frac{k}{\lambda_2-\lambda_1}$, on peut dire que z est de la forme $z = k_2 e^{(\lambda_2-\lambda_1)x} + k_1$ où k_1 et k_2 sont des constantes réelles.

En revenant à y_0, cela donne : y_0 solution de (E_0) ssi $y_0 = (k_2 e^{(\lambda_2-\lambda_1)x} + k_1)e^{\lambda_1 x} = k_2 e^{\lambda_2 x} + k_1 e^{\lambda_1 x}$. C'est ce qu'il fallait démontrer.

Exemple : Résolution de $y'' + y' - 6y = 0$.

L'équation caractéristique est $x^2 + x - 6 = 0$. On a $\Delta = 1^2 - 4 \times (-6) = 1 + 24 = 25 = 5^2 > 0$. Il y a deux solutions distinctes $\lambda_1 = \frac{-1-5}{2} = \frac{-6}{2} = -3$ et $\lambda_2 = \frac{-1+5}{2} = \frac{4}{2} = 2$.

Donc les solutions de l'équation différentielle sont les fonctions de la forme $y = k_1 e^{-3x} + k_2 e^{2x}$ avec k_1 et k_2 dans $\mathbb{R}$.

Propriété 2.29 (A connaître!). *Si $\Delta = b^2 - 4c = 0$ alors on appelle λ_1 l'unique solution de l'équation caractéristique et les solutions de l'équation (E_0) (équation homogène) sont les fonctions de la forme $y = (k_1 x + k_2)e^{\lambda_1 x}$ avec k_1 et k_2 constantes réelles.*

Démonstration : (pour les plus curieux) D'après la propriété 2.27, en prenant pour λ la racine double λ_1, on trouve que z' est solution de $y' + (2\lambda_1 + b)y = 0$. Or, $\lambda_1 = \frac{-b}{2}$, on en déduit que z' est solution de $y' = 0$. Autrement dit, $z'' = 0$ et z est de la forme $z = (k_1 x + k_2)$ avec k_1 et k_2 constantes réelles. On obtient alors facilement le résultat demandé.

Exemple : Résoudre $y'' + 8y' + 16y = 0$.

L'équation caractéristique est $x^2 + 8x + 16 = 0$. Le discriminant vaut alors $\Delta = 8^2 - 4 \times 16 = 64 - 64 = 0$. On a donc une racine double qui est $\lambda_1 = \frac{-8}{2} = -4$.

Les solutions de l'équation différentielle sont dont les fonctions de la forme $y = (k_1 x + k_2)e^{-4x}$

Propriété 2.30 (Hors programme examen, mais au programme de TD). *Si $\Delta = b^2 - 4c < 0$ alors on appelle $\omega = \frac{\sqrt{-\Delta}}{2}$ et les solutions de l'équation (E_0) (équation homogène) sont les fonctions de la forme $y = e^{-\frac{b}{2}x}(k_1 \cos(\omega x) + k_2 \sin(\omega x))$ avec k_1 et k_2 constantes réelles.*

Exemple : Résoudre l'équation différentielle $y'' + 25y = 0$.

L'équation caractéristique est $x^2 + 25 = 0$. Le discriminant vaut $\Delta = 0^2 - 4 \times 25 = -100 < 0$. On pose alors $\omega = \frac{\sqrt{100}}{2} = \frac{10}{2} = 5$ et les solutions de l'équation différentielle sont les fonctions de la forme $y = e^{-\frac{0}{2}x}(k_1 \cos(5x) + k_2 \sin(5x)) = k_1 \cos(5x) + k_2 \sin(5x)$ avec k_1 et k_2 réels.

b) Solution particulière

On ne va s'intéresser ici qu'à deux cas particuliers de fonctions : les polynômes et les polynômes fois une exponentielle.

Propriété 2.31.

- *Si f est un polynôme alors il existe une solution particulière qui est un polynôme, souvent du même degré que f.*
- *Si f est de la forme $f(x) = e^{\lambda x}P(x)$ où λ est un réel non nul et P est un polynôme, alors il existe une solution particulière de la forme $y_p = e^{\lambda x}Q(x)$ où Q est également un polynôme, souvent du même degré que P.*

Exemple : Trouvez une solution particulière de $y'' + y' - 6y = 31 - 28x - 6x^2$.

Le second membre étant un polynôme de degré 2, on va chercher une solution de la forme $y_p = \alpha + \beta x + \gamma x^2$. On a alors $y_p' = \beta + 2\gamma x$ et $y_p'' = 2\gamma$. Donc y_p est solution de l'équation différentielle ssi $2\gamma + \beta + 2\gamma x - 6\alpha - 6\beta x - 6\gamma x^2 = 31 - 28x - 6x^2$, ce qui donne :

$$2\gamma + \beta - 6\alpha + (2\gamma - 6\beta)x - 6\gamma x^2 = 31 - 28x - 6x^2$$

Il suffit d'avoir $-6\gamma = -6$, $2\gamma - 6\beta = -28$ et $2\gamma + \beta - 6\alpha = 31$, ce qui donne $\gamma = 1$, $\beta = 5$ et $\alpha = -4$.

Donc $y_p = -4 + 5x + x^2$ est une solution particulière de l'équation différentielle.

c) Résolution complète de l'équation différentielle

Propriété 2.32. *Les solutions de l'équation différentielle* $y'' + by' + cy = f(x)$ *s'écrivent sous la forme* $y = y_0 + y_p$ *où* y_0 *est solution de l'équation homogène et* y_p *est une solution particulière.*

Exemple : Résoudre $y'' + y' - 6y = 31 - 28x - 6x^2$.

D'après ce qui précède, les solutions sont les fonctions $y = k_1e^{-3x} + k_2e^{2x} + -4 + 5x + x^2$ avec k_1 et k_2 constantes réelles.

d) Solution avec conditions initiales

Chapitre 3

Probabilités

I Dénombrement, cardinal d'un ensemble fini, combinaisons

1) Introduction

Les problèmes de dénombrement font souvent partie des problèmes les plus difficiles en mathématiques. Ils font intervenir un nombre d'outils assez impressionnant que ce soit en algèbre ou en analyse.

Dans ce cours nous allons étudier trois situations élémentaires de dénombrement, qui se ramènent chacune à un tirage de boules dans une urne.

On considère donc une urne avec n boules numérotées. On tire successivement k boules de l'urne, soit en les remettant dans l'urne après chaque tirage (avec remise), soit en ne les remettant pas (sans remise). On veut alors dénombrer le nombre de possibilités de tirage, en considérant à nouveau deux options : soit l'ordre du tirage est important, soit l'ordre n'est pas important.

Exemple : On choisit $n = 3$ et $k = 2$.

- Sans remise et avec ordre, on trouve les tirages suivants : 1-2, 1-3, 2-1, 2-3. Il y a donc 4 possibilités.
- Sans remise et sans ordre, on trouve les tirages suivants : 1-2, 1-3, 2-3, il y a donc 3 possibilités.
- Avec remise et avec ordre, on trouve les tirages suivants : 1-1, 1-2, 1-3, 2-1, 2-2, 2-3, 3-1, 3-2, 3-3. Il y a donc 9 possibilités.
- Avec remise et sans ordre on trouve les tirages suivants : 1-1, 1-2, 1-3, 2-2, 2-3, 3-3. Il y a donc 6 possibilités.

Remarque : Pour ceux qui sont attentifs, nous avons ici quatre situations et nous donnerons des formules seulement pour trois d'entre elles.

Voici des exemples concrets :

Exemples :

1. L'ADN est une succession de paires de 4 nucléotides notés A, C, G et T. A va toujours avec T et C avec G, de sorte que si on ne regarde qu'un côté de cette chaîne, on peut reconstituer l'autre partie. Combien y a-t-il de possibilités pour former 10 paires consécutives ?
2. Un cadenas comporte 4 chiffres, combien de codes différents possibles y a-t-il ?
3. Un biologiste, pour une étude, doit choisir un échantillon de 3 individus parmi 20. Combien y a-t-il d'échantillons possibles ?
4. Le tiercé est constitué des trois premiers chevaux arrivés dans l'ordre pour une course comportant 18 chevaux. Combien y a-t-il de tiercé possibles ?

Chacun de ses problèmes peut se ramener à un tirage dans une urne. Vous pouvez réfléchir dans chacun des cas à la situation correspondante dans le modèle de l'urne.

1. Ici l'ordre est important (et même fondamental). Nous pouvons dire que l'urne contient nos 4 nucléotides, et on va en tirer 10 avec remise. Nous avons 4 choix différents pour le premier, 4 pour le second,

etc. jusqu'au 10ème, donc en tout $\underbrace{4 \times 4 \times 4 \times ... \times 4}_{10 \text{ facteurs}} = 4^{10}$ possibilités.
Ici donc $n = 4$, $k = 10$, l'ordre compte et il y a remise.

2. On a la même situation que l'exemple précédent : on va tirer 4 chiffres avec remise dans une urne qui en contient 10 (les chiffres de 0 à 9). On trouve qu'il y a $10^4 = 10000$ possibilités. Ce qui est normal puisque nous avons tous les nombres de 0 à 9999.
Ici donc $n = 10$, $k = 4$, l'ordre compte et il y a remise.
3. On peut supposer que tous les 20 individus sont dans une urne et on va en choisir 3. Bien qu'on choisisse les individus les uns après les autres, l'ordre n'est pas important puisqu'au final nous nous intéressons qu'à l'échantillon. Supposons que l'on ait choisi les individus A, B et C. Cela correspond à 6 tirages distincts : ABC, ACB, BAC, BCA, CAB et CBA. Donc les $20 \times 19 \times 18 = 6840$ résultats possibles du tirage correspondent à $\frac{6840}{6} = 1140$ groupes de 3 personnes.
Ici $n = 20$, $k = 3$, l'ordre ne compte pas et il n'y a pas de remise.
4. Dans ce dernier exemple, on peut imaginer que les 18 chevaux sont dans une urne et qu'on tire 3 chevaux, le premier correspondant au premier cheval arrivé et ainsi de suite. On a donc $18 \times 17 \times 16 = 4896$ possibilités.
Ici $n = 18$, $k = 3$, l'ordre compte et il n'y a pas de remise.

2) Formules de dénombrement

Définition 3.1. *Soit $n \in \mathbb{N}$ on appelle factorielle de n, que l'on note $n!$ le nombre $n! = n \times (n-1) \times (n-2) \times ... \times 2 \times 1$ lorsque $n \geq 1$ et $0! = 1$ par convention.*

	Avec remise	Sans remise
L'ordre compte	n^k	$A_n^k = \dfrac{n!}{(n-k)!}$
L'ordre ne compte pas	hors programme	$\binom{n}{k} = \dfrac{n!}{k!(n-k)!}$

Remarque : En pratique, pour calculer A_{20}^5 et $\binom{20}{5}$, on écrit $A_{20}^5 = 20 \times 19 \times 18 \times 17 \times 16$ (5 facteurs en décroissant) et $\binom{20}{5} = \dfrac{20 \times 19 \times 18 \times 17 \times 16}{5 \times 4 \times 3 \times 2 \times 1}$, avec également 5 facteurs en décroissant au dénominateur.

3) Cardinal d'un ensemble

Définition 3.2. *Le cardinal d'un ensemble Ω est le nombre d'éléments de cet ensemble. On le note $\text{card}\,\Omega$.*

Exemple : $\text{card}\,\{1; 2; 3; 4; 5; 6\} = 6$.

4) Combinaisons

Propriété 3.3. *Soit $n \in \mathbb{N}$ et $k \in \mathbb{N}$ avec $k \leq n$. $\binom{n}{k}$ représente le nombre de façon de choisir un groupe de k éléments parmi n.*

1. $\binom{n}{0} = \binom{n}{n} = 1$, $\binom{n}{1} = n$.
2. $\binom{n}{k} = \binom{n}{n-k}$.
3. *Si* $k \geq 1$ *alors* $\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}$.
4. *Formule du binôme de Newton : pour tout a et b réels et tout $n \in \mathbb{N}$,*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

Démonstration :

1. Évident.
2. Choisir k élément parmi n revient à en rejeter $n-k$ parmi n donc finalement à en choisir $n-k$ parmi n.
3. Preuve 1 : $\binom{n}{k}+\binom{n}{k-1}=\frac{n!}{(n-k)!k!}+\frac{n!}{(k-1)!(n-k+1)!}=\frac{n!\times(n-k+1)+n!\times k}{k!(n+1-k)!}=\frac{(n+1)!}{k!(n+1-k)!}=\binom{n+1}{k}$
Preuve 2 : choisir k élément parmi $n+1$ revient à choisir k éléments parmi les n premiers ou bien choisir le dernier et choisir $n-1$ éléments parmi les n premiers.
4. (admis)

Remarque : La troisième propriété donne une autre façon de calculer avec le triangle de Pascal :

	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$
$n=0$	1							
$n=1$	1	1						
$n=2$	1	2	1					
$n=3$	1	3	3	1				
$n=4$	1	4	6	4	1			
$n=5$	1	5	10	$\boxed{10}$	$\boxed{5}$	1		
$n=6$	1	6	15	20	$\boxed{15}$	6	1	
$n=7$	1	7	21	35	35	21	7	1

Pour calculer la ligne suivante on ajoute la somme des deux termes au dessus et à gauche, comme sur l'exemple.

5) Principe multiplicatif

On peut parfois décomposer une expérience dont on veut compter le nombre d'issues en une succession de deux expériences. Si pour chacune des m issues de la première expérience, il y a le même nombre d'issues pour la seconde expérience, disons n alors le nombre de résultats de l'expérience est $m\times n$.

Exemple : Les plaques minéralogiques en Italie s'écrivent avec deux lettres suivies de 3 chiffres suivis de deux lettres. Les lettres ne peuvent pas être la lettre O. Combien y a-t-il de plaques possibles ?

On décompose le choix des plaques en 3 expériences indépendantes : choix de deux lettres puis choix de 3 chiffres puis choix de deux lettres. On a donc $25^2\times 10^3\times 25^2$ plaques différentes.

II Univers, événements aléatoires, probabilité

1) Introduction

Le but de la théorie des probabilités est de dégager des méthodes permettant de faire des prédictions qualitatives ou quantifiées sur le déroulement des phénomènes qui sont régis pas le hasard.

Dans chaque cas, le phénomène étudié peut être considéré comme une "expérience" dont le résultat est aléatoire : lorsqu'on reproduit l'expérience dans des conditions identiques, le résultat de l'expérience varie et semble dépendre du hasard. On dit qu'il s'agit d'une expérience aléatoire.

Définition 3.4. *Une expérience aléatoire est une expérience dont le résultat dépend du hasard.*

Exemples :

1. On jette deux dés de couleurs différentes. Le résultat de l'expérience est la donnée des nombres affichés par chacun des dés.
2. On tire au hasard successivement et sans remise deux boules d'une urne contenant 10 boules numérotées de 1 à 10. Le résultat peut être la succession des deux numéros de boules tirées.

2) Univers

Pour travailler sur une expérience aléatoire, on fait ce que l'on appelle une **modélisation** pour représenter cette expérience réelle. La première étape consiste à décrire les issues possibles

Définition 3.5. *On appelle univers associé à l'expérience $\mathcal{E}$, tout ensemble Ω dont les éléments représentent toutes les issues (ou résultats) possibles de $\mathcal{E}$.*

Pour les exemples cités plus haut, on retient habituellement les choix suivants

1. $\Omega = \{(a,b),\ a,b \in \mathbb{N},\ 1 \leq a,b \leq 6\}$
2. $\Omega = \{(a,b),\ a,b \in \mathbb{N},\ \text{entiers } 1 \leq a,b \leq 10, a \neq b\}$

Ce choix de Ω est bien une modélisation car on pourrait très bien retenir pour ces deux cas $\Omega = \mathbb{N}$ ou encore pour l'expérience du lancé de dé que l'on obtienne un "dé cassé".

Considérons l'expérience suivante : on joue deux fois à pile ou face. Alors on peut prendre $\Omega = \{PP, PF, FP, FF\}$, on aurait pu retenir "la pièce tombe sur la tranche". On a un modèle presque identique pour l'expérience suivante : on choisit une famille parmi les familles de deux enfants et on note le sexe des enfants, celui du plus jeune puis celui de l'aîné. On obtient : $\Omega = \{FF, FG, GF, GG\}$. Si on note seulement le sexe des enfants sans se préoccuper du rang de naissance, on a $\Omega = \{FF, FG, GG\}$.

3) Événements

Considérons une propriété (ou qualité) E liée au résultat de l'expérience. A chaque répétition de l'expérience, E est réalisée ou non. Par exemple, dans l'exemple 1, on peut avoir E="les deux dés affichent des nombres pairs" et dans l'exemple 2, on peut considérer E= "on a tiré au moins une boule de numéro impair". Une telle propriété permet de partager l'univers Ω en deux parties : d'une part, l'ensemble A_E des points $\omega \in \Omega$ qui représentent une issue de $\mathcal{E}$ pour laquelle E a lieu, et d'autre part, la partie $\bar{A}_E$ formée des points $\omega \in \Omega$ associés à une issue ne réalisant pas E. On obtient ainsi une représentation de la propriété E par une partie A_E de Ω. On dit que A_E est attaché à la propriété E ou encore que A_E est l'événement "E est réalisé". On voit alors que $\bar{A}_E$ est aussi un événement ; c'est l'événement "E n'est pas réalisé".

Dans l'exemple 1, l'événement : "la somme des nombres obtenus est 6" est l'ensemble :

$$A = \{(1,5), (2,4), (3,3), (4,2), (5,1)\}$$

a) Evénement certain

C'est l'événement lié à une propriété toujours réalisée lorsqu'on effectue l'expérience. Cet événement est l'ensemble Ω tout entier.

b) Evénement impossible

C'est l'événement associé à une propriété qui n'est jamais réalisée. Cet événement est la partie vide notée $\emptyset$.

c) Evénement complémentaire

Si A est un événement, on dit que $\bar{A}$ est l'événement complémentaire de A. $\bar{A}$ est aussi la partie complémentaire de A dans Ω. C'est l'événement "non A".

d) Evénement élémentaire

On appelle ainsi les événements liés à l'expérience E qui ne sont réalisés que pour une seule issue de E. Ce sont les parties de Ω réduites à un point. Un tel élément $\omega \in \Omega$ est appelée aussi issue possible.

e) Opérations sur les événements

En plus du complémentaire, il existe deux autres opérations sur les événements :

1. La réunion : si A et B sont deux événements, l'événement A ou B est la partie de Ω réunion des parties A et B, notée $A \cup B$.
2. L'intersection : si A et B sont deux événements, l'événement A et B est la partie de Ω intersection des parties A et B, notée $A \cap B$.

Par exemple, dans le jeu de pile ou face, l'événement élémentaire "2 fois pile" est l'intersection de l'événement $A =$"d'abord pile" $=\{PP, PF\}$ et de l'événement $B =$"pile au deuxième lancé" $=\{FP, PP\}$. L'événement réunion est l'événement "au moins une fois pile". L'ensemble $\mathcal{A}$ des événements est donc muni des opérations de complémentation, réunion et intersection.

f) Événements incompatibles

On dit que les événements A et B (relatifs à une même expérience) sont incompatibles s'ils ne peuvent être réalisés simultanément, i.e., si $A \cap B = \emptyset$.

g) Système complet d'événements

On appelle ainsi toute suite $A_1, \ldots, A_n$ tels que

1. $A_i \cap A_j = \emptyset$ si $i \neq j$
2. $\cup_{i=1}^{n} A_i = \Omega$

Exemple. L'expérience considérée est le lancé de dés, on pose pour $i = 1, \ldots 6$, $A_i =$ "le premier dé affiche i.

h) Lois de De Morgan

Propriété 3.6. *Soient A et B deux événements.*

1. $\overline{A \cap B} = \overline{A} \cup \overline{B}$,
2. $\overline{A \cup B} = \overline{A} \cap \overline{B}$.

4) Probabilité

Après avoir fixé l'univers, la deuxième étape de la modélisation consiste à associer une probabilité à tout événement. Le choix de cette fonction probabilité est fait de façon à décrire la réalité observée. Quand cette description transcrit bien la réalité , c'est que la modélisation est bien satisfaisante.

Définition 3.7. *Une expérience aléatoire est décrite par un couple $(\Omega, \mathbb{P})$ appelé* **espace de probabilité**, *où*

1. *Ω est l'univers (ensemble des issues possibles).*
2. *$\mathbb{P}$ est une fonction définie sur l'ensemble des événements $\mathcal{P}(\Omega)$ (ensemble des parties de Ω) et vérifiant :*
   (a) $\forall A \subset \Omega,\ 0 \leq \mathbb{P}(A) \leq 1$,
   (b) $\mathbb{P}(\Omega) = 1$,
   (c) *Si $A \cap B = \emptyset$ alors $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$.*

Remarque : Si $A \subset \Omega$, $\mathbb{P}(A)$ mesure la probabilité de réalisation de A. Intuitivement, $\mathbb{P}(A)$ est la fréquence de réalisation de A au cours d'un très grand nombre de mise en oeuvre de l'expérience considérée.
Exemples :

1. On lance un dé équilibré. Alors $\Omega = \{1, 2, 3, 4, 5, 6\}$ et pour tout $\omega \in \Omega$, $\mathbb{P}(\omega) = \frac{1}{6}$.
2. On considère un jeu de 32 cartes. $\Omega = \{7\ trèfle, \ldots, as\ pique\}$. Si toutes les cartes ont la même probabilité d'être tirées, alors $\mathbb{P}(\omega) = \frac{1}{32}$.

Remarque. Si $A = \{\omega_1, \ldots, \omega_n\}$ alors

$$\mathbb{P}(A) = \mathbb{P}(\omega_1) + \ldots + \mathbb{P}(\omega_n)$$

Définition 3.8. *Si $|\Omega| = card(\Omega) < +\infty$ et si pour tout $\omega \in \Omega, \mathbb{P}(\omega) = C = constante$, on dit que $\mathbb{P}$ est uniforme et que toutes les issues sont équiprobables.*

Propriété 3.9. *Si $|\Omega| < +\infty$ et si $\mathbb{P}$ est uniforme, alors*

1. *$C = \frac{1}{|\Omega|}$,*
2. *Pour tout événement A, $\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{\text{nombre d'issues favorables}}{\text{nombre d'issues possibles}}$.*

Démonstration

1. $1 = \mathbb{P}(\Omega) = \sum_{\omega \in \Omega} \mathbb{P}(\omega) = \sum_{\omega \in \Omega} C = |\Omega| \times C$. D'où $C = \frac{1}{|\Omega|}$.
2. $\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega) = \sum_{\omega \in A} \frac{1}{|\Omega|} = \frac{|A|}{|\Omega|}$.

Propriété 3.10. *1. $\mathbb{P}(\emptyset) = 0$.*

2. *$\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$.*
3. *$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$*

Démonstration

1. On a $\emptyset \cup \emptyset = \emptyset$ et $\emptyset \cap \emptyset = \emptyset$. D'après la propriété ($c$), on a $\mathbb{P}(\emptyset) = \mathbb{P}(\emptyset) + \mathbb{P}(\emptyset)$. Donc $\mathbb{P}(\emptyset) = 0$.
2. Soit $A \subset \Omega$. On a $A \cap \overline{A} = \emptyset$ et $A \cup \overline{A} = \Omega$. D'où $1 = \mathbb{P}(\Omega) = \mathbb{P}(A \cup \overline{A}) = \mathbb{P}(A) + \mathbb{P}(\overline{A})$. D'où $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$.
3. Soient A et B deux événements. On a

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

Ces trois événements sont incompatibles. En utilisant une généralisation de la propriété (c), on obtient :

$$\mathbb{P}(A \cup B) = \mathbb{P}(A \cap \overline{B}) + \mathbb{P}(B \cap \overline{A}) + \mathbb{P}(A \cap B)$$

On a aussi $A = (A \cap B) \cup (A \cap \overline{B})$. Donc $\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap \overline{B})$. On obtient

$$\mathbb{P}(A \cap \overline{B}) = \mathbb{P}(A) - \mathbb{P}(A \cap B)$$

On a aussi

$$\mathbb{P}(B \cap \overline{A}) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

Finalement, on a

$$\mathbb{P}(A \cup B) = (\mathbb{P}(A) - \mathbb{P}(A \cap B)) + (\mathbb{P}(B) - \mathbb{P}(A \cap B)) + \mathbb{P}(A \cap B)$$

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

III Probabilités conditionnelles et indépendance

Ces deux notions ont un rôle très important pour faire les calculs. Considérons une expérience aléatoire. On donne l'information que B a été réalisé. Généralement, la réalisation de A sera différente de ce qu'elle était avant de savoir que B a été réalisé.

1) Définition

Commençons par un exemple.
On jette un dé et on s'intéresse à l'événement A : le résultat est supérieur ou égal à 4. On a $\mathbb{P}(A) = \frac{3}{6} = \frac{1}{2}$. Supposons que l'on sache que le résultat est pair (événement B). Si on calcule la probabilité que le résultat soit supérieur à 4 (c'est-à-dire le probabilité que A se réalise sachant que B est déjà réalisé, on obtient $\frac{2}{3}$. Cette probabilité est notée $\mathbb{P}_B(A)$ et s'appelle la probabilité conditionnelle de A, B étant donné ou encore la probabilité de A sachant B.

Définition 3.11. *Soient A et B deux événements, tels que $\mathbb{P}(B) \neq 0$. La probabilité conditionnelle de A , B étant donné est le nombre :*

$$\mathbb{P}_B(A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

On dit aussi que $\mathbb{P}_B(A)$ est la probabilité de A sachant B. Elle vérifie les propriétés d'une probabilité.

Remarque : Dans la littérature on trouve parfois la notation $P(A|B)$ pour probabilité d'avoir A sachant B. Elle est fortement déconseillée car elle a tendance à faire confondre les notions ; en effet dans cette notation $A|B$ n'est pas un évènement et $\mathbb{P}$ n'est pas la fonction probabilité considérée, c'est $\mathbb{P}(\ |B)$ qui est maintenant une fonction probabilité et A est bien l'évènement.

Comme on l'a vu dans l'exemple précédent, en général $\mathbb{P}(A) \neq \mathbb{P}_B(A)$.

Remarque : On peut aussi utiliser les probabilités conditionnelles pour trouver $\mathbb{P}_B(A)$.

Exemple : Dans un magasin d'électroménager, on s'intéresse au comportement d'un acheteur potentiel d'un téléviseur et d'un lecteur DVD. La probabilité qu'il achète un téléviseur est 0,6. La probabilité qu'il achète un lecteur DVD quand il n'a pas acheté un téléviseur est 0,2. La probabilité qu'il achète un lecteur DVD quand il a acheté un téléviseur est 0,4. La probabilité que la personne achète un lecteur DVD et un téléviseur est donnée par $\mathbb{P}$(achète un téléviseur et un lecteur DVD)= $\mathbb{P}_{\{\text{a acheté un téléviseur}\}}$(achète un lecteur DVD)$\times\mathbb{P}$(achète un téléviseur)$= 0,4 \times 0,6 = 0,24$.

2) Formule des probabilités totales

Reprenons l'exemple précédent. On cherche maintenant la probabilité pour que le client achète un lecteur DVD. Le lecteur peut acheter un lecteur DVD soit en ayant acheté un téléviseur, soit sans avoir acheté un téléviseur et ces deux événements sont incompatibles. On a donc

$$\begin{aligned}
\mathbb{P}(\text{acheter un lecteur DVD}) &= \mathbb{P}(\text{acheter un lecteur DVD et un téléviseur}) \\
&\quad +\mathbb{P}(\text{acheter un lecteur DVD sans un téléviseur}) \\
&= \mathbb{P}_{\{\text{ayant acheté un téléviseur}\}}(\text{acheter un lecteur DVD}) \cdot \mathbb{P}(\text{acheter un téléviseur}) \\
&\quad +\mathbb{P}_{\{\text{sans achat de téléviseur}\}}(\text{acheter un lecteur DVD})\mathbb{P}(\text{ne pas acheter un télévis}
\end{aligned}$$

$\mathbb{P}(\text{acheter un lecteur DVD}) = 0,4 \times 0,6 + 0,2 \times 0,4 = 0,32$

Propriété 3.12. *Soient A et B deux événements. Alors*

$$\mathbb{P}(A) = \mathbb{P}(A \cap \bar{B}) + \mathbb{P}(A \cap B) = \mathbb{P}_{\bar{B}}(A) \cdot \mathbb{P}(\bar{B}) + \mathbb{P}_B(A) \cdot \mathbb{P}(B)$$

Démonstration. Il suffit de vérifier que les événements $A \cap B$ et $A \cap \bar{B}$ sont incompatibles.

On généralise ce résultat de la façon suivante :

Théorème 3.13. *Soit $C_1, C_2, \ldots, C_n$ un système complet d'événements. Soit A un événement. Alors, on a :*

$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A \cap C_i) = \sum_{i=1}^{n} \mathbb{P}_{C_i}(A) \cdot \mathbb{P}(C_i)$$

3) Formule de Bayes

Dans l'exemple précédent, on cherche à connaître la probabilité qu'un client ayant acheté un lecteur DVD ait aussi acheté un téléviseur. Il s'agit ici de la probabilité conditionnelle $\mathbb{P}$ (acheter un téléviseur | avoir acheté un lecteur DVD). Notons T l'événement : acheter un téléviseur et D l'événement : acheter un lecteur DVD. On voit donc que l'on cherche

$$\mathbb{P}_D(T) = \frac{\mathbb{P}(D \cap T)}{\mathbb{P}(D)} = \frac{\mathbb{P}_T(D) \times \mathbb{P}(T)}{\mathbb{P}(D)}$$

On obtient $\mathbb{P}_D(T) = 0,75$

Théorème 3.14. *Soient A et B deux événements. On a*

$$\mathbb{P}_A(B) \cdot \mathbb{P}(A) = \mathbb{P}_B(A) \cdot \mathbb{P}(B) = \mathbb{P}(A \cap B)$$

ou encore

$$\mathbb{P}_A(B) = \frac{\mathbb{P}_B(A) \cdot \mathbb{P}(B)}{\mathbb{P}(A)}$$

Remarque. Il se peut que parfois l'on ne connaisse pas $\mathbb{P}(A)$. dans ce cas, on utilise la formule des probabilités totales et l'on obtient :

$$\mathbb{P}_A(B) = \frac{\mathbb{P}_B(A) \cdot \mathbb{P}(B)}{\mathbb{P}_{\bar{B}}(A) \cdot \mathbb{P}(\bar{B}) + \mathbb{P}_B(A) \cdot \mathbb{P}(B)}$$

Propriété 3.15. *Soient $C_1, C_2, \cdots, C_n$ un système complet d'événements et A un événement. Alors*

$$\mathbb{P}_A(C_i) = \frac{\mathbb{P}(C_i) \cdot \mathbb{P}_{C_i}(A)}{\mathbb{P}(C_1) \cdot \mathbb{P}_{C_1}(A) + \mathbb{P}(C_2) \cdot \mathbb{P}_{C_2}(A) + \cdots + \mathbb{P}(C_n) \cdot \mathbb{P}_{C_n}(A)}$$

4) Événements indépendants

Définition 3.16. *On dit que B est indépendant de A si $\mathbb{P}(A \cap B) = \P(A)\mathbb{P}(B)$.*

Remarque : Si B est indépendant de A, alors A est indépendant de B.

Propriété 3.17. *Deux événements A et B de probabilités non nulles sont indépendants si, et seulement si, $\mathbb{P}_A(B) = \mathbb{P}(B)$ (ou si $\mathbb{P}_B(A) = \mathbb{P}(A)$).*

Remarque : Attention, il ne faut pas confondre indépendants et incompatibles. Si A et B sont tels que $A \cap B = \emptyset$, $\mathbb{P}(A \cap B) = 0$. A et B ne pourront être indépendants que si $\mathbb{P}(A) = 0$ ou $\mathbb{P}(B) = 0$. Si on tire au hasard une carte dans un jeu de 52 cartes, les événements "A : pique" et "B : valet" sont indépendants mais non incompatibles. Les événements "A : pique" et "C : trêfle" sont incompatibles et non indépendants.

5) Exercices

Exercice 6 : Une compagnie d'assurance répartit ses clients en 3 classes : les bons risques (classe R1), les risques moyens (classe R2) et les mauvais risques (classe R3). Les effectifs de ces trois classes représentent 20% de la population totale pour la classe R1, 50% pour la classe R2 et 30% pour la classe R3. Les statistiques indiquent que les probabilités d'avoir un accident au cours de l'année sont de 0,05 pour la classe R1, 0,15 pour la classe R2 et 0,3 pour la classe R3.

1. Quelle est la probabilité pour qu'une personne choisie au hasard dans la population ait un accident au cours de l'année ?
2. Si une personne n'a pas eu d'accident au cours de l'année, quelle est la probabilité qu'elle soit dans la classe R1 ?

Exercice 7 :

Dans une population donnée, 89 % des victimes d'une infection virale présente un symptôme qui n'atteint que 23 % de la population non infectée. On sait de plus que 29 % de la population présente ce symptôme.

1. Traduire en langage mathématique les données de l'énoncé
   Réponse : Si on note V le fait d'être atteint par le virus et si on note S le fait de présenter le symptôme alors l'énoncé nous dit :
   $$P_V(S) = 0,89, \quad P_{\bar{V}}(S) = 0,23, \quad P(S) = 0,29.$$
2. Quelle est la probabilité qu'un individu choisi au hasard dans cette population ne soit pas infecté ?
   Réponse : On cherche $P(\bar{V})$.
   On sait que
   $$\begin{aligned} P(S) &= P(V \cap S) + P(\bar{V} \cap S) \\ &= P(V).P_V(S) + P(\bar{V}).P_{\bar{V}}(S) \\ &= (1 - P(\bar{V})).P_V(S) + P(\bar{V}).P_{\bar{V}}(S) \end{aligned}$$
   On a donc une équation à une seule inconnue $0,29 = 0,89 - 0,66P(\bar{V})$.
3. Quelle est la probabilité qu'un individu présentant le symptôme soit infecté ?
4. Quelle est la probabilité qu'un individu ne présentant pas le symptôme, ne soit pas infecté ?

Chapitre 4

Statistiques descriptives

I Introduction - Vocabulaire de la statistiques

L'objet de la statistique descriptive est de décrire une famille de chiffres appelée **série statistique**. On cherche à concentrer l'information d'un grand nombre de valeurs en une information facilement apréhendable. Bien sûr ces chiffres proviennent de mesures expérimentales, de simulations numériques, d'enquêtes, etc. En les étudiant, on cherche à comprendre certains phénomènes, des relations, une explication.

Les ensembles sont appelées **populations**. Les éléments de cette population sont les **individus**. La population peut être étudiée selon un ou plusieurs **caractères** : ce sont les "propriétés" que présentent les individus de la population. Un caractère permet de découper la population selon différentes **modalités**. Lorsque les modalités prises par le caractère sont de nombres, le caractère est dit **quantitatif**. On lui donne alors le nom de **variable statistique**. Une variable statistique est **discrète** si elle ne prend que des valeurs isolées (par exemple le nombre d'enfants dans une famille) ou **continue** si elle peut prendre toutes le valeurs comprises entre deux valeurs données (par exemple la taille d'un individu). Lorsque les modalités du caractères ne sont pas mesurables, le caractère est dit **qualitatif** (par exemple pour les automobiles, la couleur est un caractère qualitatif).

II Représentation graphique

Une première façon de se faire une idée d'une série de chiffres est de les représenter graphiquement, il y a différentes façons de faire,

1) Caractères quantitatifs discrets

Il s'agit de données associées à une population chiffrées et dont les valeurs sont isolées.

a) Diagrammes en bâtons

Définition 4.1. *Un diagramme en bâtons est un moyen de représenter une série statistique dont le caractère est quantitatif discret. Si $x_1, \dots, x_p$ sont les valeurs possibles prises par le caractère et si les effectifs correspondants sont $n_1 \dots, n_p$, il est constitué par les segments qui relient le point $(x_k, 0)$ au point (x_k, n_k).*

Il permet de faire apparaître sur l'axe des abscisses les valeurs prises (ou non prises) par le caractère, les échelles des deux axes sont importantes.

Dans une classe, les notes obtenues à un devoir sont :

Notes	1	2	3	4	5	6	7	8	9	10
Effectif	2	1	6	5	2	9	7	1	0	1

Le diagramme en bâtons correspondant se trouve dans le diaporama "Statistiques descriptives".
Remarque : dans les diagrammes en bâtons, les longueurs sont alors proportionnelles aux effectifs.

2) Caractères quantitatifs continus

Il s'agit de données associées à une population chiffrées et dont les valeurs sont des intervalles de valeurs.

a) Histogrammes

Définition 4.2. *Un histogramme est un moyen de représenter une série statistique dont le caractère est quantitatif continu. Si la série statistique est donnée par les classes* $([a_i, a_{i+1}[)$*, il est constitué par des rectangles dont la base est le segment* $[a_i, a_{i+1}[$ *(sur l'axe des réels) et l'aire est proportionnelle à l'effectif de la classe.*

Remarque : C'est l'aire qui doit être proportionnelle à l'effectif de la classe et non la hauteur elle-même. Si toutes les classes ont la même étendue, il n'y a pas de problème. Sinon, on note n_i l'effectif de la classe $[a_i, a_{i+1}[$. On choisit un rapport de proportionnalité k. La hauteur du rectangle de base $[a_i, a_{i+1}[$ sera alors $k \times \frac{n_i}{a_{i+1}-a_i}$.
Il permet de faire apparaître sur l'axe des abscisses les intervalles de valeurs prises (ou non prises) par le caractère, seule l'échelle de cet axe est important.
Exemple : On a demandé la taille des élèves dans une classe de 33 élèves. On obtient les résultats suivants :

Taille (en cm)	150-160	160-170	170-175	175-180	180-190	190-200
Effectif	3	12	9	6	2	1

L'histogramme correspondant se trouve dans le diaporama "Statistiques descriptives".

3) Caractères qualitatifs

Il s'agit de données associées à une population non numériques.

a) Camemberts

Définition 4.3. *Un diagramme circulaire est un moyen de représenter une série statistique dont le caractère est qualitatif. Il est obtenu en découpant un disque en secteurs dont les mesures d'angle sont proportionnelles à l'effectif.*

Exemple : Dans une entreprise, on a demandé aux employés leur moyen de transport pour venir au travail. Les résultats sont les suivants :

Moyen utilisé	à pied	en voiture	en métro
Effectif	50	110	200

Le diagramme circulaire correspondant se trouve dans le diaporama "Statistiques descriptives".

4) Fréquence cumulée et courbe de fréquence cumulée

Définition 4.4. *On appelle fréquence le rapport entre l'effectif d'une valeur et l'effectif total.*

Considérons l'exemple suivant des notes obtenues dans une classe :

Notes	1	2	3	4	5	6	7	8	9	10
Effectif	2	1	6	5	2	9	7	1	0	1
Fréquence	0,059	0,029	0,176	0,147	0,059	0,265	0,206	0,029	0	0,029
Fréquence cumulée	0,059	0,088	0,264	0,411	0,47	0,735	0,941	0,970	0,970	1

Définition 4.5. *On note g_k la fréquence cumulée du caractère x_k, c'est-é-dire $g_k = f_1 + ... + f_k$ oé $f_i = \frac{n_i}{N}$ est la fréquence du caractère x_i. La courbe des fréquences cumulées est celle obtenue en joignant les points (x_i, g_i).*

Vous pouvez trouver cette représentation dans le diaporama "Statistiques descriptives".

III Paramètres

Un tableau ou un graphique ne permettent pas toujours d'avoir une idée précise de la série de chiffres étudiés. On va donc définir des paramètres qui donneront une meilleure caractéristique de ce qui est observé.

1) Les paramètres de position

Considérons les deux séries de chiffres suivantes :

1. 0 ; 0 ; 10 ; 15 ; 20
2. 2000 ; 2000 ; 2010 ; 2015 ; 2020

Dans les deux cas, nous avons 5 nombres. Ces deux séries se ressemblent mais leur position est différente. On va donc regarder certains paramètres.

a) La moyenne arithmétique

Elle est égale à la somme des valeurs divisées pas leur nombre. Elle donne une idée sur la localisation.

Cas de données énumérées. Dans ce cas, le calcul de la moyenne est très simple. Pour les exemples précédents, on obtient des moyennes respectives de 9 et 2009.

Cas d'une variable discrète. Si la variable est discrète, on emploie la formule de la moyenne pondérée. Si on a t classes d'effectifs n_i ou de fréquences f_i, la moyenne $\bar{x}$ s'écrit pour les valeurs $x_1, x_2, \ldots, x_t$ de la variable

$$\bar{x} = \frac{n_1x_1 + n_2x_2 + \ldots + n_tx_t}{n_1 + n_2 + \ldots + n_t} = \frac{f_1x_1 + f_2x_2 + \ldots + f_tx_t}{f_1 + f_2 + \ldots + f_t}$$

Reprenons l'exemple des notes obtenues à un devoir :

Notes	1	2	3	4	5	6	7	8	9	10
Effectif	2	1	6	5	2	9	7	1	0	1

La moyenne est : $\frac{2\times1+1\times2+6\times3+5\times4+2\times5+9\times6+7\times7+1\times8+0\times9+1\times10}{34} = 5,09$

b) La médiane

La médiane d'une série statistique, généralement notée $x_{1/2}$, est le nombre qui sépare la série (ordonnée en valeurs croissantes) en deux groupes de même effectif. Pour trouver cette médiane, quand la série est discrète, on écrit la liste de toutes les valeurs de la série par ordre croissant, chacune d'entre elles étant répétée autant de fois que son effectif. Si l'effectif total n est un nombre impair, la médiane est le terme de rang $\frac{n+1}{2}$ Si l'effectif total n est un nombre pair, la médiane est le centre de l'intervalle formé par les termes de rang $\frac{n}{2}$ et $\frac{n}{2} + 1$. Dans le cas de la série statistique 0 ; 0 ; 10 ; 15 ; 20, $x_{1/2} = 10$.

c) Le mode

Le mode d'une série statistique est la valeur la plus fréquente. Dans le cas de la série statistique 0 ; 0 ; 10 ; 15 ; 20, le mode est 0.

2) Les paramètres de dispersion

Considérons deux séries ayant même moyenne et pourtant très différentes :
0 ; 0 ; 10 ; 15 ; 20 et -1000 ; -1000 ; 10 ; 1015 ; 1020.
On remarque que dans le premier cas les valeurs sont beaucoup plus rapprochées que dans le second, on peut mesurer cet éloignement de différentes façons.

a) L'étendue

L'étendue d'une d'une série statistique est la différence entre les deux valeurs extrêmes. Pour les séries précédentes, on obtient respectivement 20 et 2020.

b) La variance

Soit une série statistique (x_k, n_k) telle que

1. N est l'effectif total de la série.
2. Les valeurs x_k sont les valeurs prises par la série.
3. n_k est le nombre de fois où la valeur x_k est prise.
4. $\bar{x}$ représente la moyenne de la série : $\bar{x} = \frac{n_1x_1+n_2x_2+\ldots+n_kx_k}{N}$

Définition 4.6. *Afin de savoir si la population est éloignée de sa moyenne on peut regarder des éléments évaluant la distance à la moyenne.*
On appelle alors variance de la série statistique (x_k, n_k) le nombre :

$$s^2 = \frac{n_1(x_1-\bar{x})^2 + n_2(x_2-\bar{x})^2 + \ldots + n_k(x_k-\bar{x})^2}{N}$$

$$s^2 = \frac{n_1x_1^2 + n_2x_2^2 + \ldots + n_kx_k^2}{N} - \bar{x}^2$$

Dans les séries précédentes, on a respectivement $s^2 = 145$ et $s^2 = 4070725$.

c) L'écart-type

Définition 4.7. *L'écart-type s est la racine carrée de la variance.*

Dans les séries précédentes, on a respectivement $s = 12,04$ et $s = 2017,58$.

d) L'écart moyen absolu et l'écart médian absolu

Pour information, on peut se donner d'autres paramètres de dispersion. Ceux qui suivent ne sont pas au programme mais donnés juste pour information.

Définition 4.8. *L'écart moyen absolu est défini par $e_m = \frac{1}{N}\sum_{k=1}^{N}|x_i-\bar{x}|$.*
L'écart médiant absolu est défini par $e_m^* = \frac{1}{N}\sum_{k=1}^{N}|x_i-x_{1/2}|$.

Considérons la série statistique suivante : 1 ; 2 ; 3 ; 4 ; 5 ; 10 ; 11 ; 12 ; 15. On a $\bar{x} = 7$ et $x_{1/2} = 5$.

x_i	1	2	3	4	5	10	11	12	15	Somme
$\|x_i-\bar{x}\|$	6	5	4	3	2	3	4	5	8	40
$\|x_i-x_{1/2}\|$	4	3	2	1	0	5	6	7	10	38

On obtient $e_m = \frac{40}{9} = 4,44$ et $e_m^* = \frac{38}{9} = 4,22$

e) Les quartiles et l'intervalle interquartile

Les valeurs ont été rangées dans l'ordre croissant, de la plus petite à la plus grande. Médiane

Définition 4.9. *Les quartiles permettent de séparer une série statistique en quatre groupes de même effectif (à une unité près) :*

1. *Un quart des valeurs sont inférieures au premier quartile Q_1.*
2. *Un quart des valeurs sont supérieures au troisème quartile Q_3.*

Définition 4.10. *On appelle intervalle interquartile l'intervalle $]Q_1; Q_3[$. On appelle écart interquartile la différence $Q_3 - Q_1$.*

Pour déterminer les quartiles Q_1 et Q_3 d'une série de N valeurs, on procède de la façon suivante :
On calcule la quantité $\frac{N}{4}$. Deux cas sont possibles : soit le résultat est entier (la division tombe juste), soit le résultat n'est pas entier.

1. Cas 1 : le résultat est entier.
   - Q_1 est la valeur n-ème valeur de la série où $n = \frac{N}{4}$.
   - Q_3 est la valeur n'-ème valeur de la série où $n' = \frac{3N}{4}$.
   Exemple :
   Prenons les valeurs rangées dans l'ordre croissant :
   1-3-3-3-5-5-6-7-7-8-8-8-9-9-10-10-10-10-11-11-12-13-13-13-14-15-16-19. Il y a $N = 28$ valeurs et $\frac{28}{4} = 7$. Donc le résultat est entier. On obtient donc $Q_1 = 6$ (la 7-ème valeur de la série rangée dans l'ordre croissant). Puis on a $Q_3 = 13$ (la 21-ème valeur de la série rangée dans l'ordre croissant).
2. Cas 2 : le résultat n'est pas entier.
   - On arrondit $\frac{N}{4}$ à l'entier supérieur n et Q_1 est la valeur n-ème valeur de la série.
   - On arrondit $\frac{3N}{4}$ à l'entier supérieur n' et Q_3 est la valeur n'-ème valeur de la série.
   Exemple :
   Prenons les valeurs rangées dans l'ordre croissant : 3-5-5-6-7-8-8-9-9-10-10-10-10-11-11-12-13-13-13-14-15-16-19. Il y a $N = 23$ valeurs et $\frac{N}{4} = 5,75$. Donc Q_1 est la 6-ème valeur de la série rangée dans l'ordre croissant c'est-à-dire $Q_1 = 8$. Puisque $\frac{3n}{4} = 17,75$, Q_3 est la 18-ème valeur de la série rangée dans l'ordre croissant donc $Q_3 = 13$.

Interprétation des quartiles.
Si on connait les quartiles Q_1 et Q_3 d'une série, on peut en déduire les renseignements suivants sur la série statistique :

1. Au moins un quart (25%) des valeurs sont inférieures ou égales à Q_1.
2. Au moins trois quarts (75%) des valeurs sont inférieures ou égales à Q_3.
3. Environ la moitié des valeurs se trouvent dans l'intervalle interquartile $[Q_1; Q_3]$.

Exemple : dans une classe, les notes présentent un premier quartile Q_1 égal à 10 et un troisième quartile égal à 14. On peut dire que :

- environ un quart des élèves a moins de 10
- environ trois quarts des élèves a moins de 14
- environ la moitié des élèves a une note entre 10 et 14

3) Les paramètres de forme

Pour information, on peut se donner d'autres paramètres permettant d'étudier les distributions des valeurs. Les notions qui suivent ne sont pas au programme mais donnés juste pour information.
On peut aussi étudier le côté symétrique de la distribution. On note m_r le moment centré d'ordre r N, $r \in \mathbb{N}$ défini par $m_r = \frac{1}{N}\sum_{i=1}^{N}|x_i - \bar{x}|^r$, et f_1 le coefficient de Fisher défini par $f_1 = \frac{m_3}{\sqrt{m_2}^3}$. Alors, le coefficient de Fischer est proche de 0 pour une distribution symétrique. Plus il est grand en valeur absolue, plus la

distribution est étalée, avec un coefficient négatif pour une distribution étalée vers la gauche et un coefficient positif pour une distribution étalée vers la droite.

1. Exemple 1 : 0 ; 0 ;0 ; 0 ; 1 ; 2 ; 3 ; 5 ; 9 est étalée vers la droite $f_1 = 1,31$.
2. Exemple 2 : 0 ; 1 ; 3 ; 4 ; 6 ; 8 ; 8 ; 9 ; 9 est étalée vers la gauche $f_1 = -0,36$.
3. Exemple 3 : 0 ; 1 ; 2 ; 3 ; 5 ; 6 ; 7 ; 8 ; 9 est très faiblement étalée vers la gauche $f1 = -0,05$

4) Boîte à moustaches

Les boîtes à moustaches permettent visuellement d'avoir une idée assez rapide de la distribution, les bords du rectangle représentent le premier et le troisième quartile, la barre à l'intérieur représente la médiane (deuxième quartile), les extrémités des moustaches représentent la plus petite et la plus grande valeur, qui ne sont pas considérées comme exceptionnelles. On fixe par convention comme exceptionnel un nombre qui n'appartient pas à l'intervalle $[Q_1 - 1,5(Q_3 - Q_1) ; Q_3 + 1,5(Q_3 - Q_1)]$. On trace les points exceptionnels à l'aide de points en dehors des moustaches.

Vous trouverez des exemples dans le diaporama "Statistiques descriptives".

Les deux dernières sections qui suivent sont données à titre d'information, elles ne sont pas au programme.

IV Statistique à deux dimensions

1) introduction

Dans les premières sections de ce chapitre, les séries statistiques étudiées étaient des séries simples : on étudiait un seul caractère dans une population. Il peut être utile de considérer en même temps plusieurs caractères d'une même population. Par exemple, la température et la pression d'un milieu à différentes heures d'une journée. Nous nous limitons à l'étude simultanée de deux caractères X et Y.
Supposons que nous ayons n individus ayant pour caractères (x_i, y_i) avec $1 \leq i \leq n$.

On cherche à savoir s'il existe un lien entre les caractères X et Y. Un outil qui va nous servir pour décider si les caractères X et Y sont liés est la covariance :

Définition 4.11. *La covariance est définie ainsi*

$$Cov(x, y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\bar{y}$$

Remarque : $Cov(x, x) = Var(x)$, $Cov(x, -x) = -Var(x)$.
Réciproquement, une covariance nulle ou proche de 0 va indiquer une dispersion de l'ensemble des points $M_i(x_i, y_i)$, nous allons interpréter plus rigoureusement ce résultat au paragraphe suivant.

2) Régression linéaire

Supposons que l'on ait pour chaque individu deux grandeurs (x_i, y_i), $1 \leq i \leq n$ correspondant aux deux caractères. L'ensemble des points $M_i(x_i, y_i)$ dessinés dans un repère est appelé le nuage de points. Pour étudier une éventuelle relation linéaire entre les coordonnées des points, on va chercher une droite d'équation $y = ax + b$ qui passe le plus près possible de ces points. En général, il n'existe pas de droite passant par tous les points. Le problème mathématique n'est pas facile à définir, mais une façon simple de le poser est de considérer, non pas les distances des points à la droite, mais les carrés des distances « verticales » des points à la droite. Autrement dit, on va construire les points de la droite H_i de même abscisse que M_i, et calculer la somme des distance au carrés : $\Delta = \sum_{i=1}^{n} M_i H_i^2$.
On cherche donc a et b tels que la quantité $\Delta = \sum_{i=1}^{n}(y_i - ax_i - b)^2$ soit la plus petite possible.

En développant on trouve $\Delta = nb^2 - 2b\sum_{i=1}^{n}(y_i - ax_i) + \sum_{i=1}^{n}(y_i - ax_i)^2$. On considère d'abord Δ comme un trinôme en b (remarque, si on commence par considérer que le trinôme est par rapport à la variable a, le calcul est beaucoup plus compliqué). On trouve alors :

$$\Delta = nb^2 - 2nb(\bar{y} - a\bar{x}) + \sum_{i=1}^{n}(y_i - ax_i)^2$$

Cette quantité est minimum lorsque $b = \bar{y} - a\bar{x}$, autrement dit lorsque la droite passe par le point moyen de coordonnées $(\bar{x}, \bar{y})$.
En remplaçant b par cette valeur dans Δ, et en cherchant le minimum par rapport à a, on finit par trouver : $a = \frac{Cov(x,y)}{V(x)}$
On remarque que dans le calcul précédent, on fait jouer un rôle dissymétrique aux variables x et y. Or rien, au niveau de la statistique ne permet de dire si l'une des variables dépend de l'autre. Il est donc logique de recommencer les calculs précédents en inversant les rôles de x et y. On obtient alors une droite d'équation $x = \bar{x} + a'(y - \bar{y})$ avec $a' = \frac{Cov(x,y)}{V(y)}$.

Les deux droites ainsi trouvées sont différentes. On pose

$$r^2 = aa'$$

Si les deux droites trouvées se correspondaient, on aurait $a = \frac{1}{a'}$ et donc $r^2 = 1$ et $|r| = 1$. Si $|r|$ est voisin de 1, on a un ajustement valide. Par contre si $|r|$ est proche de 0, les droites sont différentes et les points sont loin d'être alignés.

Définition 4.12. $r = \frac{Cov(x,y)}{\sqrt{V(x)V(y)}}$ *s'appelle le coefficient de corrélation.*

Propriété 4.13. $-1 \leq r \leq 1$. *Le coefficient de corrélation est égal à 1 dans le cas où l'une des variables est fonction affine croissante de l'autre, et à -1 dans le cas où la fonction est affine décroissante.*

Remarque : Plus r est proche de 1 ou -1, plus les variables sont fortement corrélées et le nuage est « plat ». Lorsque l'on a un nuage flou de points, r est proche de 0 et les variables sont peu corrélées.

Table des matières