

UNIVERSITÉ DU MAINE

Licence BBTE - Première année

Mathématiques

Alexandre POPIER

Année : 2016-2017

Table des matières

1	Étude de fonctions	3
1.1	Rappel : nombres réels	7
1.2	Fonctions	10
1.2.1	Représentation graphique	12
1.2.2	Variations	12
1.2.3	Composition	13
1.2.4	Bijection	13
1.3	Fonctions usuelles et à connaître	14
1.3.1	Polynômes du second degré (***)	14
1.3.2	Fonctions puissance et racine n -ième (*)	17
1.3.3	Fonctions polynomiales et fractions rationnelles	19
1.3.4	La fonction valeur absolue (*)	21
1.3.5	La fonction logarithme (**).	22
1.3.6	La fonction exponentielle (**).	23
1.3.7	Fonctions puissance (suite).	24
1.4	Limite d'une fonction	26
1.4.1	Définitions précises	26
1.4.2	Opérations sur les limites (**).	28
1.4.3	Continuité d'une fonction	29
1.5	Dérivée	30
1.5.1	Dérivée en un point et fonction dérivée	30
1.5.2	Calcul des dérivées (**).	32
1.5.3	Extremum (local) d'une fonction	33
1.6	Étude de fonctions	34
1.7	Modélisation de l'évolution d'un pathogène	37
1.7.1	Premier cas : β sur-linéaire	39
1.7.2	Second cas : β sous-linéaire	41
2	Probabilités	43
2.1	Espace de probabilité	44
2.1.1	L'espace des observables Ω	44
2.1.2	Les événements	44
2.1.3	La probabilité \mathbb{P}	45
2.2	Probabilités discrètes finies, équiprobabilité	47
2.2.1	Équiprobabilité sur les espaces finis	48

2.3	Dénombrement : quelques rappels	49
2.3.1	Arrangements et permutations	49
2.3.2	Indiscernabilité des objets : combinaisons	50
2.4	Indépendance et probabilités conditionnelles	52
2.4.1	Indépendance	52
2.4.2	Probabilités conditionnelles	53
2.5	Évaluation d'un risque de trisomie 21 et tests médicaux	57

Alphabet grec	63
----------------------	-----------

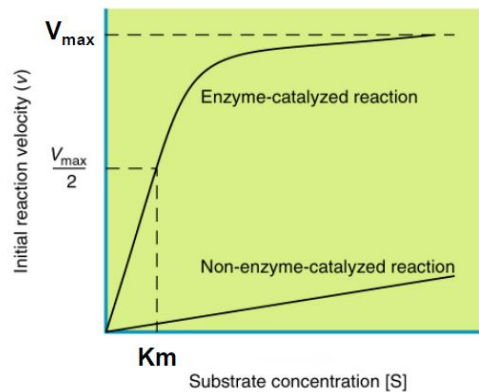
Le texte qui suit constitue un résumé du cours de mathématiques de première année. Il ne saurait se substituer à un exposé complet et commenté et encore moins à la pratique d'exercices d'application. Il servira néanmoins de référence pour tous les groupes et d'aide-mémoire en ce qui concerne les notions et outils de base.

Ces notes de cours sont évidemment une version préliminaire et nous serions reconnaissant à tout lecteur de nous faire part des fautes qu'il y aura détectées à l'adresse suivante : apopier@univ-lemans.fr.

Chapitre 1

Étude de fonctions

Démarrons cette partie en considérant le problème de l'enzymologie (cf. cours de troisième année). Une enzyme est une protéine (sauf rares exceptions), qui augmente la vitesse des réactions chimiques, sans en modifier le résultat et dont la structure se retrouve inchangée à la fin de la réaction. Cette réaction chimique transforme une molécule appelée substrat S en un produit P .



La difficulté est de quantifier cette vitesse v de réaction. Entre un instant t et un instant $t + \delta t$, avec δt petit, la concentration du produit créé par la réaction chimique varie de $C_P(t)$ à $C_P(t + \delta t)$. Classiquement la vitesse se définit alors comme la variation de concentration, soit

$$\frac{C_P(t + \delta t) - C_P(t)}{(t + \delta t) - t}.$$

Si δt tend vers 0, on aura une vitesse instantanée v égale à la dérivée de la concentration :

$$v = C'_P(t) = \frac{dC_P}{dt}(t).$$

On voit apparaître des notions vues au lycée : fonction (du temps), dérivée, etc. Cette vitesse de réaction va varier au cours du temps :

- Elle diminue quand la quantité de substrat diminue.

- Si un équilibre est possible entre S et P , elle tend vers zéro.
- Si le produit inhibe l'enzyme, la concentration de P augmentant, l'activité de l'enzyme diminue et la vitesse de réaction diminue.

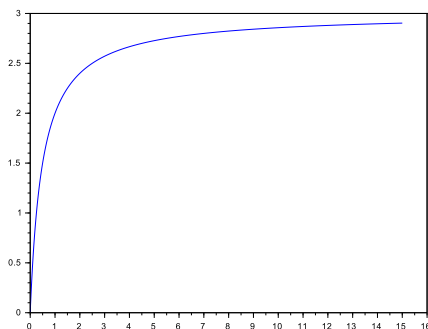
On retrouve les notions mathématiques de fonction croissante, décroissante, etc. Ce qui est étudié par les biologistes n'est donc pas la vitesse de réaction au cours du temps, mais la vitesse initiale de la réaction chimique v_i . Au début du XX^e, les biologistes Leonor Michaelis et Maud Menten ont proposé le modèle simple suivant appelé équation de Michaelis-Menten :

$$(1.1) \quad v_i = \frac{V_{max}C_S}{K_M + C_S},$$

avec

- V_{max} la vitesse maximale de réaction,
- C_S la concentration en substrat,
- K_M la constante de Michaelis spécifique de l'enzyme. Habituellement de l'ordre de 10^{-1} à 10^{-8} .

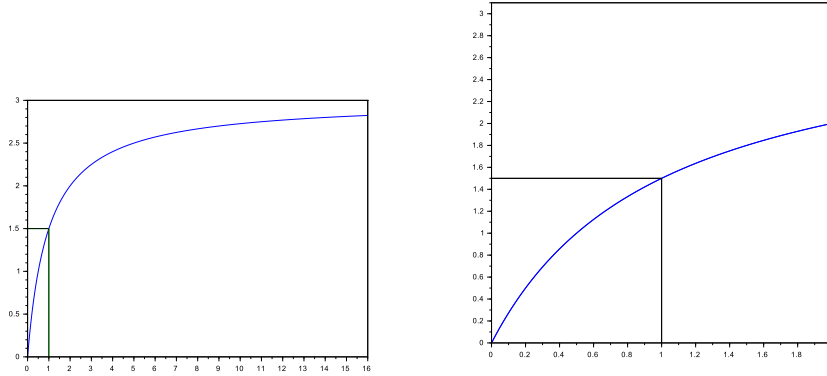
Le pourquoi chimique de cette formule sera vu en troisième année. En tant que fonction de C_S , v_i est une fraction rationnelle (voir le paragraphe 1.3.3) dont le graphe est le suivant pour $V_{max} = 3$:



Expérimentalement si on veut déterminer K_M , on remarque que si $v_i = V_{max}/2$, alors

$$V_{max}/2 = \frac{V_{max}C_S}{K_M + C_S} \Leftrightarrow K_M = C_S.$$

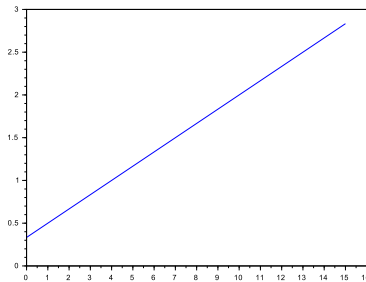
Ceci donne un moyen graphique de trouver K_M .



Néanmoins cette technique demande de connaître précisément V_{max} qui n'est que la limite de la vitesse initiale lorsque la concentration en substrat devient infinie ! Or expérimentalement il est impossible d'augmenter cette concentration infiniment. Pour cela on effectue un changement de variable ; on va exprimer $1/v_i$ comme fonction de l'inverse de la concentration $1/C_S$:

$$\frac{1}{v_i} = \frac{K_M + C_S}{V_{max}C_S} = \frac{K_M}{V_{max}C_S} + \frac{C_S}{V_{max}C_S} = \frac{K_M}{V_{max}} \frac{1}{C_S} + \frac{1}{V_{max}}.$$

Ainsi l'inverse de la vitesse initiale est une fonction affine (ou polynôme de degré 1) de $1/C_S$, dont l'ordonnée à l'origine est $1/V_{max}$ et la pente est K_M/V_{max} . C'est la relation de Lineweaver et Burk (1934).



D'autres représentations sont possibles :

— Eadie et Hofstee (1952) : v_i est linéaire en fonction de v_i/C_S :

$$v_i = V_{max} - K_M \frac{v_i}{C_S}.$$

— Hanes (1932) : C_S/v_i est linéaire en fonction de C_S :

$$\frac{C_S}{v_i} = \frac{K_M}{V_{max}} + \frac{1}{V_{max}} C_S.$$

Si mathématiquement ces formules sont équivalentes, expérimentalement le tracé de Hanes est plus précis que Lineweaver et Burk.

Enfin voici quelques remarques concernant la fonction donnée par la formule (1.1). Dérivons cette fonction :

$$\frac{dv_i}{dC_S} = \frac{V_{max}(K_M + C_S) - V_{max}C_S}{(K_M + C_S)^2} = \frac{V_{max}K_M}{(K_M + C_S)^2} > 0.$$

C'est donc une fonction croissante de C_S . Et V_{max}/K_M est la pente de la tangente à la courbe pour $C_S = 0$. De plus si $C_S < 0,01K_M$, autrement dit si C_S est petite, alors la dérivée est quasiment constante et donc

$$v_i \approx kC_S,$$

où k est une constante. On dit que la vitesse de réaction est d'ordre 1 (proportionnelle à la concentration). Inversement si C_S est très grande, $C_S > 100K_M$, alors $v_i \approx V_{max}$ et la vitesse est d'ordre 0 (constante). Ceci permet de simplifier l'étude de v_i et de donner une idée satisfaisante de v_i .

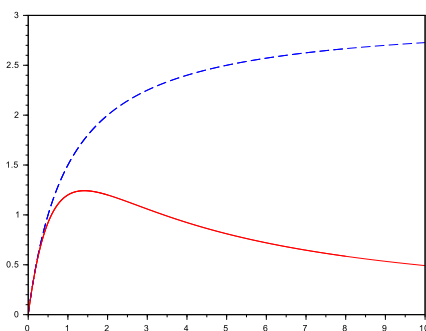
Le modèle de Michaelis-Menten est le plus simple pour décrire la vitesse de réaction initiale v_i . Il peut être ensuite compliqué pour tenir compte de différents aspects de la réaction chimique. Ainsi si un excès de substrat a tendance à inhiber la réaction, on introduit une constante K_i , constante d'inhibition, et l'équation (1.1) devient :

$$v_i = \frac{V_{max}}{1 + \frac{K_M}{C_S} + \frac{C_S}{K_i}},$$

ou en multipliant tout par C_S dans la fraction :

$$(1.2) \quad v_i = \frac{V_{max}C_S}{K_M + C_S + \frac{C_S^2}{K_i}}.$$

C'est toujours une fraction rationnelle dont le graphe est :



Le graphe montre que v_i atteint une valeur maximale pour une concentration donnée, puis diminue jusqu'à tendre vers zéro si la concentration C_S devient

infinie. Cherchons la concentration C_S qui maximise v_i . On dérive v_i par rapport à C_S :

$$\frac{dv_i}{dC_S} = V_{max} \frac{K_M + C_S + C_S^2/K_i - C_S(1 + 2C_S/K_i)}{(K_M + C_S + C_S^2/K_i)^2} = \frac{V_{max}}{K_i} \frac{K_M K_i - C_S^2}{(K_M + C_S + C_S^2/K_i)^2}.$$

Cette dérivée s'annule pour $C_S = \sqrt{K_M K_i}$ et alors la vitesse initiale maximale est

$$v_i = V_{max} \frac{1}{1 + 2\sqrt{K_M/K_i}} < V_{max}.$$

On voit donc sur ce problème d'enzymologie l'importance de savoir manipuler, étudier et tracer les fonctions usuelles. C'est l'objet de cette première partie de cours.

1.1 Rappel : nombres réels

Au collège, puis au lycée vous avez rencontré les nombres réels sous des formes différentes. En voici quelques exemples.

- Les nombres *entiers naturels* ou *relatifs*.
- Les nombres *décimaux* $a \times 10^n$ où a et n sont des entiers relatifs ; par exemple 3×10^5 et -13×10^{-2} sont des nombres décimaux.
- Les nombres *rationnels*, c'est-à-dire de la forme $\frac{a}{b}$ où a est un entier relatif et b un entier positif non nul. Un nombre décimal est un nombre rationnel.
- Les nombres définis par leur *développement décimal*, comme $0,33\dots$, où les points de suspension signifient que toutes les décimales valent 3. Un autre exemple est $-4,8964616161\dots$, où les trois points signifient que la suite des décimales continue comme indiquée, c'est-à-dire en mettant alternativement 6 et 1.

En général on ne voit pas sur les premières décimales de règle permettant de trouver les décimales suivantes. Ainsi le nombre π qui est le rapport de la circonférence d'un cercle à son diamètre est représenté par un développement décimal illimité :

$$\pi = 3,1415926535897932384626433832795028841971693993751\dots$$

Ici les points de suspension indiquent simplement qu'il n'y a pas «d'arrêt», ni de règle simple pour obtenir les décimales suivantes.

En revanche les nombres décimaux ont un développement qui se termine par des zéros : $-13 \times 10^{-2} = -0,13000\dots$. La réciproque est vraie : si un développement décimal se termine par des zéros alors le nombre est décimal.

Un nombre réel peut être représenté par un développement décimal. Mais il existe d'autres «représentations».

- Les nombres définis par des opérations à partir d'autres nombres réels ; ainsi $\frac{1+\sqrt{5}}{2}$, $\sqrt[3]{-7}$ ou $(1 - 3^{-7})^{(2^5)}$.

- Des nombres particuliers, fréquents utilisés en mathématiques (ou en physique, mécanique, etc.), qui ont une notation spéciale : le nombre π ou encore le nombre e , base du logarithme népérien.

Notations : dans tout ce cours,

- \mathbb{N} est l'ensemble des nombres entiers positifs ou nuls,
- \mathbb{Z} l'ensemble des entiers relatifs,
- \mathbb{D} l'ensemble des nombres décimaux,
- \mathbb{Q} l'ensemble des nombres rationnels,
- \mathbb{R} l'ensemble des nombres réels.

Ces ensembles vérifient

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{D} \subset \mathbb{Q} \subset \mathbb{R}.$$

Ceci signifie qu'un nombre entier positif est un nombre entier, que tout entier est décimal, tout décimal est rationnel, tout rationnel est réel. Autrement dit les ensembles sont *inclus* les uns dans les autres. Rappelons quelques notations utiles :

- Si A est un sous-ensemble de \mathbb{R} , $x \in A$ indique que le nombre réel x appartient à l'ensemble A .
- Si A et B sont deux sous-ensembles de \mathbb{R} , $A \cup B$ désigne l'union des ensembles A et B : $A \cup B$ contient tous les éléments de A et tous les éléments de B .
- $A \cap B$ est l'intersection de A et B : $A \cap B$ contient tous les éléments qui sont à la fois dans A et dans B .
- A^c est le complémentaire de A : les éléments qui sont dans A^c sont exactement ceux qui ne sont pas dans A .

On ne donnera pas ici de définition précise de l'ensemble \mathbb{R} . Néanmoins les propriétés suivantes sont à connaître.

Théorème 1.1 \mathbb{R} est un corps commutatif qui contient le corps des nombres rationnels \mathbb{Q} . L'addition et la multiplication de \mathbb{R} prolongent celles de \mathbb{Q} .

Ce théorème résume en deux phrases les règles de calcul suivantes, qui sont celles que vous avez toujours pratiquées :

$$\begin{aligned} a + b &= b + a, & 0 + a &= a, & a + b = 0 &\Leftrightarrow a = -b, & a + (b + c) &= (a + b) + c, \\ ab &= ba, & 1 \times a &= a, & ab = 1 &\Leftrightarrow a = 1/b, & a(bc) &= (ab)c, & a(b+c) &= ab+ac, \\ & & & & ab = 0 &\Leftrightarrow (a = 0 \text{ ou } b = 0). \end{aligned}$$

On peut ajouter, soustraire, multiplier et diviser par un nombre réel différent de zéro, sans se soucier de l'ordre dans lequel sont effectuées les opérations.

L'ordre sur \mathbb{R}

Tout nombre réel non nul est (strictement) positif ou (strictement) négatif. On note $\mathbb{R}_+ = [0, +\infty[$ l'ensemble des nombres réels positifs ou nul et

$\mathbb{R}_- =] - \infty, 0]$ les nombres négatifs ou nul. Si a et b sont deux nombres réels, on dit que

Définition 1.1 a est inférieur ou égal à b si le nombre $b - a$ est positif ou nul. Cette relation se note $a \leq b$ ou $b \geq a$.

La seconde notation correspond en français à : b est supérieur ou égal à a . Si $a \leq b$ et $a \neq b$, on dit que a est strictement plus petit que b , ce qui se note $a < b$ ou $b > a$.

Propriétés 1.1 \mathbb{R} est muni d'une relation d'ordre, notée \leq , qui satisfait par définition :

- $\forall x \in \mathbb{R}, x \leq x$ (réflexivité),
- $\forall (x, y) \in \mathbb{R}^2, x \leq y$ et $y \leq x \Rightarrow x = y$ (anti-symétrie),
- $\forall (x, y, z) \in \mathbb{R}^3, x \leq y$ et $y \leq z \Rightarrow x \leq z$ (transitivité),

De plus cette relation vérifie également

1. l'ordre est total, i.e. $\forall (x, y) \in \mathbb{R}^2, x \leq y$ ou $y \leq x$;
2. la relation d'ordre est compatible avec l'addition et la multiplication :

$$\forall (x, y, z) \in \mathbb{R}^3, x \leq y \Rightarrow x + z \leq y + z,$$

$$\forall (x, y) \in \mathbb{R}^2, \forall z \in \mathbb{R}_+, x \leq y \Rightarrow xz \leq yz.$$

On rappelle que si $z \leq 0$, alors $x \leq y \Rightarrow xz \geq yz$: on « retourne » l'inégalité. Lorsqu'on prend l'inverse d'une inégalité entre deux nombres ayant le même signe, alors on change le sens : si $a \leq b$ et a et b sont non nuls et ont le même signe (sont tous les deux positifs ou tous les deux négatifs), alors $\frac{1}{a} \geq \frac{1}{b}$.

On définit le plus grand nombre des nombres réels a et b en posant

$$\max(a, b) = \begin{cases} b & \text{si } b \geq a \\ a & \text{si } a \geq b \end{cases}$$

On définit de même $\min(a, b)$, le plus petit des nombres a et b .

L'ordre sur \mathbb{R} permet également de définir la notion d'intervalle de \mathbb{R} .

Définition 1.2 Soient a et b des nombres réels.

1. Si $a \leq b$, le segment $[a, b]$ est l'ensemble des nombres réels x tels que $a \leq x \leq b$.
2. On définit les intervalles ouverts :
 - si $a < b$, l'intervalle $]a, b[$ est l'ensemble des nombres réels x tels que $a < x < b$;
 - l'intervalle $]a, +\infty[$ est l'ensemble des nombres réels x tels que $a < x$;
 - l'intervalle $] - \infty, a[$ est l'ensemble des nombres réels x tels que $x < a$;

- l'ensemble vide \emptyset et $] -\infty, +\infty[= \mathbb{R}$.
- 3. Si $a < b$, l'intervalle $]a, b]$ est l'ensemble des nombres réels x tels que $a < x \leq b$ et l'intervalle $[a, b[$ est l'ensemble des nombres réels x tels que $a \leq x < b$.
- 4. Les intervalles $] -\infty, a]$ et $[a, +\infty[$ sont formés respectivement des nombres réels x tels que $x \leq a$ et $a \leq x$.

Il y a donc en tout dix « types » d'intervalles ! Le segment $[a, a]$ est l'ensemble $\{a\}$ et ne comporte qu'un élément a .

Définition 1.3 Si $a < b$, les nombres a et b s'appellent les extrémités des intervalles $[a, b]$, $]a, b]$, $[a, b[$ ou $]a, b[$. Le nombre positif $b - a$ s'appelle la longueur de l'intervalle. Le centre, ou milieu, est le nombre $c = (b + a)/2$.

Le centre c vérifie $c - a = b - c = (b - a)/2 > 0$. Donc c appartient à l'intervalle ouvert $]a, b[$.

Enfin nous dirons qu'un nombre réel x est compris entre a et b si on a ($a \leq b$ et $a \leq x \leq b$) ou bien si on a ($b \leq a$ et $b \leq x \leq a$).

1.2 Fonctions

Les phrases suivantes définissent ce que l'on appelle une fonction :

- Pour chaque temps t , le nombre de larves d'un parasite dans une plantation.
- Pour chaque nombre d , la concentration en sel de mer à la profondeur d (à la verticale d'un point fixé).
- Le nombre d'individus d'un âge donné en France en 2008.
- L'efficacité de transmission d'un agent pathogène ayant une virulence donnée.

Définition 1.4 Une fonction f est un objet mathématique qui à chaque nombre réel x situé dans un ensemble D_f associe un nombre réel noté $f(x)$. L'ensemble D_f s'appelle le domaine de définition de la fonction, x la variable et le nombre réel $f(x)$ la valeur ou l'image de f en x .

Exemple 1.1 Le nombre de larves d'un parasite détermine bien une fonction : à chaque instant t on associe ce nombre que l'on notera $L(t)$. La fonction est naturellement définie par L , son domaine de définition est \mathbb{R} et elle donne le nombre de larves en fonction du temps.

Une fonction f sera souvent connue à travers une formule qui indique comment calculer $f(x)$ à partir de x . Mais il ne faut pas confondre fonction et formule. Dans les exemples précédents, les fonctions ne font référence à aucune formule particulière. Néanmoins en voici deux exemples.

Exemple 1.2 La fonction $x \mapsto x^2$ désigne l'application f telle que $f(x) = x^2$ pour tout $x \in \mathbb{R}$. En revanche la fonction $x \mapsto 1/x$ est l'application qui à un nombre réel non nul associe son inverse.

Le domaine D_f d'une fonction f est formé des réels x pour lesquels, d'une part la valeur $f(x)$ a un sens (mathématique), et d'autre part un intérêt dans la situation considérée. Ainsi dans l'exemple de la salinité de l'eau de mer, comme d désigne la profondeur, d doit être positif. $d = 0$ correspond à la surface, $d = 1$ signifiera un mètre sous la surface par exemple. Dans l'exemple des larves, le nombre $L(t)$ est positif ou nul et donc on éliminera de D_L les valeurs de t telles que $L(t) < 0$. En absence d'un contexte clair de modélisation (ou de toute autre indication), si la fonction f est définie par une formule, on considérera que le domaine de définition est formé de toutes les valeurs de la variable (la plus grande partie de \mathbb{R} au sens de l'inclusion) pour lesquelles la formule a un sens. Dans l'exemple (1.2), on n'a pas le droit de diviser par zéro. Donc la valeur $x = 0$ doit être enlevé de D_g . En revanche on peut toujours élever au carré; donc $D_f = \mathbb{R}$.

Les opérations entre nombres réels s'étendent aux fonctions comme suit.

Définition 1.5 Soient $f : D \rightarrow \mathbb{R}$ et $g : D \rightarrow \mathbb{R}$ des fonctions définies sur le même ensemble de départ D et $\lambda \in \mathbb{R}$. On définit

- la fonction **somme** $f + g : D \rightarrow \mathbb{R}$ qui à $x \in D$ associe $(f + g)(x) = f(x) + g(x)$;
- la fonction $\lambda f : D \rightarrow \mathbb{R}$ en posant $(\lambda f)(x) = \lambda \times f(x)$ pour tout $x \in D$;
- la fonction **produit** $fg : D \rightarrow \mathbb{R}$ en posant $(fg)(x) = f(x) \times g(x)$ pour tout $x \in D$.

Pour les relations d'ordre,

Définition 1.6 Soit $f : D_f \rightarrow \mathbb{R}$ une fonction et soit A une partie de D_f .

- La fonction f est **positive** ou **nulle** sur A si $f(x) \geq 0$ pour tout $x \in A$.
- La fonction f est **strictement positive** sur A si $f(x) > 0$ pour tout $x \in A$.

On définit de même une fonction **négative** ou **nulle** ou une fonction **strictement négative** sur A .

Si g est une fonction elle aussi définie sur $D = D_f$, f est **inférieure** ou **égale à g** , et on note $f \leq g$, si on a $f(x) \leq g(x)$ pour tout $x \in D$. On définit de même la relation **supérieure** ou **égale**.

Attention : deux nombres réels sont toujours comparables ($x \leq y$ ou $y \leq x$). Mais il n'est pas toujours possible de comparer deux fonctions. Ainsi si $f : \mathbb{R} \rightarrow \mathbb{R}$ est donnée par $f(x) = x$ et $g : \mathbb{R} \rightarrow \mathbb{R}$ est définie par la formule $g(x) = x^2$, alors pour tout $x \geq 1$, $f(x) \leq g(x)$, donc $f \leq g$ sur $[1, +\infty[$ (et aussi sur $] - \infty, 0]$), mais $f \leq g$ n'est pas vrai sur \mathbb{R} car $f(1/2) = 1/2 > 1/4 = g(1/2)$.

S'il existe un nombre a tel que $f(x) = a$ pour tout $x \in A$, f est dite **constante sur A** . Si $a = 0$, f est **nulle sur A** . Si de plus A est égal à l'ensemble D_f sur lequel est définie f , on dit simplement que f est constante de valeur a (l'ensemble D_f est sous-entendu s'il n'y a pas d'ambiguïté possible).

1.2.1 Représentation graphique

La **représentation graphique** ou **graphe** d'une fonction f est la courbe obtenue dans un repère formé de deux axes orthogonaux en positionnant pour chaque x dans D_f un point d'abscisse x sur l'axe horizontal et d'ordonnée $f(x)$ sur l'axe vertical. Le domaine D_f apparaît ainsi comme la zone de l'axe des abscisses au dessus de laquelle la représentation graphique de la fonction est tracée. La représentation graphique permet d'embrasser en un coup d'oeil toutes les valeurs d'une fonction $f(x)$ (plutôt que de les calculer une à une) et surtout de se rendre compte de l'influence de la variable x sur ces valeurs.

1.2.2 Variations

Définition 1.7 Soit $f : D_f \rightarrow \mathbb{R}$ une fonction et soit I un intervalle contenu dans D_f . On dit que

— f est **croissante sur I** si pour tous x et y dans I ,

$$x \leq y \Leftrightarrow f(x) \leq f(y);$$

— f est **strictement croissante sur I** si pour tous x et y dans I ,

$$x < y \Leftrightarrow f(x) < f(y);$$

— f est **décroissante sur I** si pour tous x et y dans I ,

$$x \leq y \Leftrightarrow f(x) \geq f(y);$$

— f est **strictement décroissante sur I** si pour tous x et y dans I ,

$$x < y \Leftrightarrow f(x) > f(y).$$

Une fonction f est **monotone sur I** si elle est croissante ou décroissante sur I , **strictement monotone** si elle est strictement décroissante ou strictement croissante sur I .

On peut remarquer que f est décroissante si et seulement si $-f$ est croissante, et qu'une fonction est constante si et seulement si elle est croissante et décroissante. Il faut prendre soin de toujours préciser l'ensemble sur lequel il y a (ou non) monotonie d'une fonction.

Exemple 1.3 La fonction $f(x) = 3x - 5$ est croissante sur \mathbb{R} . La fonction $f(x) = 1/x$ est décroissante sur $]0, +\infty[$. La fonction $f(x) = x^2$ est décroissante sur $] -\infty, 0]$ et croissante sur $[0, +\infty[$.

1.2.3 Composition

Définition 1.8 (Composition) Soient $f : D_f \rightarrow \mathbb{R}$ et $g : D_g \rightarrow \mathbb{R}$ deux fonctions telles que pour tout $x \in D_f$, $f(x) \in D_g$. La composée de g par f est la fonction $h : D_f \rightarrow \mathbb{R}$ telle que pour tout $x \in D_f$, $h(x) = g(f(x))$. On la note $g \circ f$.

Attention : cette définition n'est pas symétrique. La composée $g \circ f$ peut exister sans que $f \circ g$ existe. De plus, même si $g \circ f$ et $f \circ g$ existent, en général $g \circ f \neq f \circ g$.

1.2.4 Bijection

On dit que f est une bijection de D vers D' si pour tout $y \in D'$, l'équation $f(x) = y$, dont l'inconnue est x , admet une unique solution dans D . Cette unique solution x est une expression en $y \in D'$, on la note $x = f^{-1}(y) \in D$. L'application $g : D' \rightarrow D$ qui au réel y associe $g(y) = f^{-1}(y)$ est appelée la bijection réciproque de f . Elle vérifie :

$$(1.3) \quad \forall x \in D, g(f(x)) = (g \circ f)(x) = x,$$

$$(1.4) \quad \forall y \in D', f(g(y)) = (f \circ g)(y) = y.$$

C'est la seule fonction qui vérifie ces deux relations !

La fonction $x \mapsto x^2$ est une bijection de \mathbb{R}^+ sur \mathbb{R}^+ et sa bijection réciproque est la fonction $t \mapsto \sqrt{t}$. On retiendra que $\boxed{\sqrt{t^2} = |t|}$, pour tout $t \in \mathbb{R}$.

Exemple 1.4

1. Soit $f(x) = 3x - 2$ et $D = D' = \mathbb{R}$. Résolvons l'équation $f(x) = y$.

$$3x - 2 = y \implies x = \frac{1}{3}y + \frac{2}{3}$$

est l'unique réel vérifiant l'équation. Donc f est bijective et sa bijection réciproque est $y \mapsto f^{-1}(y) = \frac{1}{3}y + \frac{2}{3}$.

2. Soit $f(x) = \frac{1}{x^2 + 1}$ et $D = D' =]0, +\infty[$. Résolvons l'équation $f(x) = y$. En remarquant que $x^2 + 1 > 1$ pour tout $x \in D$, alors l'équation $\frac{1}{x^2 + 1} = y$ n'admet pas de solution pour $y \in [1, +\infty[$.

On se place donc dans le cas où $D =]0, +\infty[$ et $D' =]0, 1[$. Alors

$$f(x) = y \iff \frac{1}{x^2 + 1} = y \implies x^2 = \frac{1 - y}{y} \implies x = \sqrt{\frac{1 - y}{y}}$$

est l'unique réel vérifiant l'équation. La bijection réciproque est

$$y \mapsto f^{-1}(y) = \sqrt{\frac{1 - y}{y}}.$$

On pourra vérifier directement que les relations (1.3) et (1.4) ont vraies.

1.3 Fonctions usuelles et à connaître

Dans ce chapitre, nous allons redonner les principales propriétés de fonctions telles que les polynômes, ln ou exp.

1.3.1 Polynômes du second degré (***)

Définition 1.9 Soient a, b et c réels. On appelle fonction polynôme du second degré toute application

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto f(x) = ax^2 + bx + c \quad (a \neq 0) \end{aligned}$$

Forme canonique :

$$f(x) = a \left(x^2 + \frac{b}{a}x + \frac{c}{a} \right).$$

Comme on a

$$\left(x + \frac{b}{2a} \right)^2 = x^2 + \frac{b}{a}x + \frac{b^2}{4a^2},$$

alors

$$\begin{aligned} f(x) &= a \left(\left(x + \frac{b}{2a} \right)^2 - \frac{b^2}{4a^2} + \frac{c}{a} \right) = a \left(\left(x + \frac{b}{2a} \right)^2 - \frac{b^2 - 4ac}{4a^2} \right) \\ &= a \left(\left(x + \frac{b}{2a} \right)^2 - \frac{b^2 - 4ac}{4a^2} \right). \end{aligned}$$

C'est cette expression qui est nommée **forme canonique** du polynôme du second degré.

Variation :

Considérons la fonction $x \mapsto x^2$. Elle est croissante si $x \geq 0$ et décroissante si $x \leq 0$. On en déduit alors le résultat suivant : la fonction $x \mapsto \left(x + \frac{b}{2a} \right)^2$ est croissante si $x + \frac{b}{2a} \geq 0$ et décroissante si $x + \frac{b}{2a} \leq 0$.

Il en est de même de la fonction $g : x \mapsto \left(x + \frac{b}{2a} \right)^2 - \frac{b^2 - 4ac}{4a^2}$.

La fonction f varie dans le même sens que la fonction g si $a > 0$

x	$-\infty$	$-\frac{b}{2a}$	$+\infty$
$a > 0 : g(x)$	$+\infty$	$-\frac{b^2 - 4ac}{4a^2}$	$+\infty$

et elle varie en sens contraire si $a < 0$.

x	$-\infty$	$-\frac{b}{2a}$	$+\infty$
$a < 0 : g(x)$	$-\infty$	$-\frac{b^2 - 4ac}{4a^2}$	$-\infty$

Toute fonction de \mathbb{R} dans \mathbb{R} , définie par $f(x) = ax^2 + bx + c$, ($a \neq 0$) est représentée dans le plan rapporté à un repère orthonormé, par une **parabole** dont l'axe a pour équation $x = -\frac{b}{2a}$. Le point $S = \left(-\frac{b}{2a}, -\frac{b^2 - 4ac}{4a^2}\right)$ est le sommet de la parabole.

Racines du polynôme du second degré

On pose $\Delta = b^2 - 4ac$. L'expression de la fonction polynôme prend la forme

$$f(x) = a \left[\left(x + \frac{b}{2a}\right)^2 - \frac{\Delta}{4a^2} \right].$$

1. Si $\Delta < 0$, alors $\left(x + \frac{b}{2a}\right)^2 - \frac{\Delta}{4a^2} > 0$. Par conséquent la fonction polynôme ne s'annule pas, ce qui revient à dire que le polynôme du second degré n'admet pas de racines réelles.
2. Si $\Delta > 0$, alors

$$\left(x + \frac{b}{2a}\right)^2 - \left(\frac{\sqrt{\Delta}}{2a}\right)^2 = \left(x + \frac{b}{2a} - \frac{\sqrt{\Delta}}{2a}\right) \left(x + \frac{b}{2a} + \frac{\sqrt{\Delta}}{2a}\right)$$

et

$$f(x) = a \left(x - \frac{-b + \sqrt{\Delta}}{2a}\right) \left(x - \frac{-b - \sqrt{\Delta}}{2a}\right).$$

Donc le polynôme du second degré admet deux racines réelles :

$$x_1 = \frac{-b + \sqrt{\Delta}}{2a}, \quad \text{et} \quad x_2 = \frac{-b - \sqrt{\Delta}}{2a}.$$

3. Si $\Delta = 0$ alors $f(x) = a\left(x + \frac{b}{2a}\right)^2$. Le polynôme du second degré admet une racine double $x_1 = x_2 = -\frac{b}{2a}$.

Remarque : Si $\Delta \geq 0$ alors

$$ax^2 + bx + c = a(x - x_1)(x - x_2).$$

La somme et le produit des racines sont donnés par les relations suivantes :

$$s = x_1 + x_2 = -\frac{b}{a}, \quad \text{et} \quad p = x_1 x_2 = \frac{c}{a}.$$

Signe du polynôme du second degré

Rappelons que

$$f(x) = a\left(x + \frac{b}{2a}\right)^2 - \frac{\Delta}{4a^2}.$$

D'après l'étude précédente,

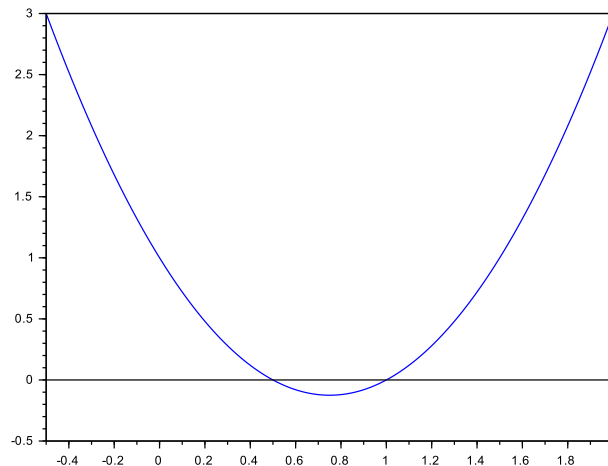
1. Si $\Delta < 0$, alors le signe de $f(x)$ est le signe de a .
2. Si $\Delta > 0$, alors $f(x) = a(x - x_1)(x - x_2)$ et le signe de $f(x)$ est le signe de a , à l'extérieur des racines et le signe contraire de a , à l'intérieur des racines.
3. Si $\Delta = 0$, alors $f(x) = a(x - x_1)^2$ et le signe de $f(x)$ est le signe de a .

Exemple 1 : Soit $f(x) = 2x^2 - 3x + 1$. La courbe de f est la parabole de sommet S d'abscisse $-\frac{b}{2a} = \frac{3}{4}$. De plus $\Delta = 1$; $x_1 = \frac{1}{2}$; $x_2 = 1$. On obtient la factorisation du polynôme :

$$f(x) = 2\left(x - \frac{1}{2}\right)(x - 1).$$

Et enfin

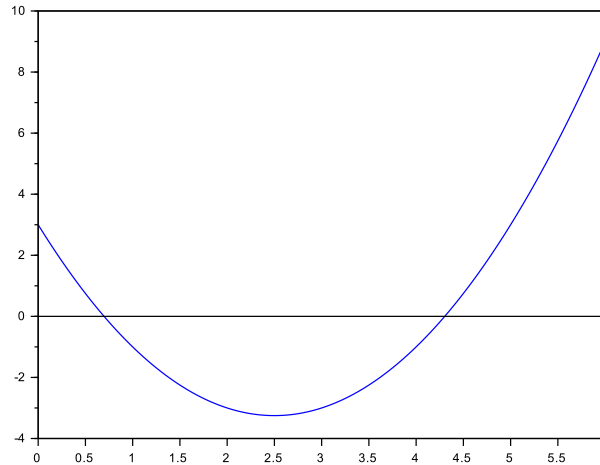
$$f(x) \begin{cases} \geq 0 & \text{si } x \in]-\infty, \frac{1}{2}] \cup [1, +\infty[\\ < 0 & \text{si } x \in]\frac{1}{2}, 1[\end{cases}$$



Exemple 2 : Un mouvement rectiligne de loi horaire $u = f(t)$ est dit uniformément varié (ou à accélération constante) si f est une fonction polynôme du second degré.

Soit $u = t^2 - 5t + 3$. Considérons le déplacement du mobile M d'abscisse $f(t)$ sur l'axe (O, \vec{U}) pour les réels $t \geq 0$. En $t = 0$, le mobile est à la

position M_0 d'abscisse 3, il se déplace ensuite dans le sens contraire de l'axe pour atteindre la position M_1 d'abscisse $\frac{-13}{4}$ (en passant par l'origine en $t = \frac{5 - \sqrt{13}}{2}$), puis change de direction pour aller dans le sens croissant des abscisses (en passant de nouveau par O en $t = \frac{5 + \sqrt{13}}{2}$).



Position d'un réel par rapport aux racines du polynôme.

On considère le cas où la fonction polynôme du second degré admet 2 zéros distincts (soit $\Delta > 0$). Nous avons alors $f(x) = a(x - x_1)(x - x_2)$. Il en résulte que pour un réel x_0 , le signe de $af(x_0)$ est le même que celui de $(x_0 - x_1)(x_0 - x_2)$. Par conséquent :

- si $af(x_0) < 0$ alors $x_0 \in]x_1, x_2[$;
- si $af(x_0) > 0$ et si $x_0 < \frac{s}{2}$, alors $x_0 \in]-\infty, x_1[$.

1.3.2 Fonctions puissance et racine n -ième (★)

Vous connaissez déjà les fonctions puissances définies pour les entiers $n \in \mathbb{N}$. En effet si n est un entier positif

$$\forall x \in \mathbb{R}, \quad x^n = \underbrace{x \times \dots \times x}_{n \text{ fois}}$$

avec par convention $x^0 = 1$ pour tout $x \in \mathbb{R}$.

Soit n un entier supérieur ou égal à 2. La fonction $x \mapsto x^n$ est continue (voir chapitre ??) et strictement croissante sur l'intervalle $[0, +\infty[$. La valeur en 0 est 0 et l'on a $\lim_{x \rightarrow +\infty} x^n = +\infty$. La fonction $x \mapsto x^n$ définit une bijection de $[0, +\infty[$ sur $[0, +\infty[$.

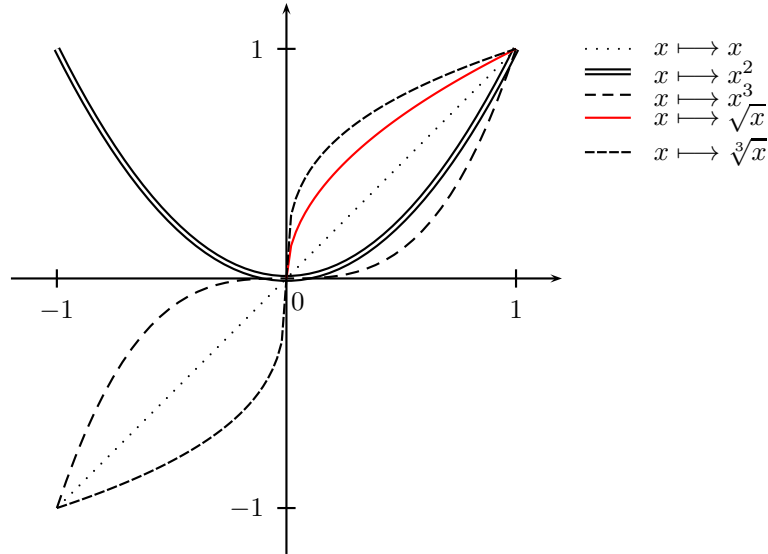


FIGURE 1.1 – Fonctions puissance

Si l'entier n est impair, la fonction $x \mapsto x^n$ est continue et strictement croissante sur \mathbb{R} , et l'on a $\lim_{x \rightarrow -\infty} x^n = -\infty$. Dans ce cas la fonction $x \mapsto x^n$ est une bijection de \mathbb{R} sur \mathbb{R} .

Pour un entier n strictement positif fixé, si $x \geq 0$, il existe un unique nombre réel $y \geq 0$ tel que

$$x = y^n.$$

y est la racine n -ième de x et se note $y = \sqrt[n]{x}$ ou $y = x^{1/n}$. On a alors pour tout $x \geq 0$:

$$x = (\sqrt[n]{x})^n = (x^{1/n})^n$$

Ainsi on a défini la fonction racine n -ième sur le domaine $[0, +\infty[$: c'est la bijection réciproque de $x \mapsto x^n$. On la note $x \mapsto \sqrt[n]{x}$ ou $x \mapsto x^{1/n}$. Si $n = 2$, c'est la fonction racine carrée que l'on note simplement $x \mapsto \sqrt{x}$.

Remarque 1.1 Notons que si n est un entier impair, on peut définir la racine n -ième sur \mathbb{R} tout entier.

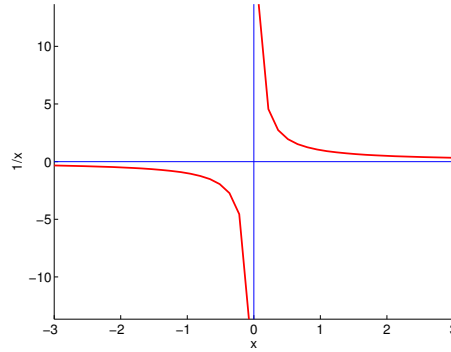
La racine n -ième est une fonction continue et strictement croissante. De plus :

- pour tout x et y positifs ou nuls, $y = x^n \Leftrightarrow x = \sqrt[n]{y} = y^{1/n}$.
- Si $x \in [0, 1]$, alors $x^n \leq x$ et en prenant la racine n -ième, on obtient $x \leq \sqrt[n]{x}$.
- Si $x \geq 1$, alors $x \leq x^n$, donc $\sqrt[n]{x} \leq x$.

Enfin si n est strictement négatif, alors

$$\forall x \neq 0, \quad x^n = \left(\frac{1}{x}\right)^{-n}.$$

Terminons ce paragraphe par les règles de calcul sur les puissances.

FIGURE 1.2 – Fonction inverse $x^{-1} = 1/x$ sur $[-3; 3]$ **Proposition 1.1**

— Si m et n sont deux entiers strictement positifs et si $x \geq 0$:

$$(x^m)^{1/n} = (x^{1/n})^m = x^{m/n}$$

— En général

$$(x^\alpha)^\beta = x^{\alpha\beta} \text{ et } x^\alpha x^\beta = x^{\alpha+\beta}.$$

— $x^{-c} = 1/(x^c) = (1/x)^c$.

$$\lim_{x \rightarrow +\infty} x^\alpha = \begin{cases} 0 & \text{si } \alpha < 0, \\ 1 & \text{si } \alpha = 0, \\ +\infty & \text{si } \alpha > 0. \end{cases}$$

$$\lim_{x \rightarrow 0} x^\alpha = \begin{cases} +\infty & \text{si } \alpha < 0, \\ 1 & \text{si } \alpha = 0, \\ 0 & \text{si } \alpha > 0. \end{cases}$$

1.3.3 Fonctions polynomiales et fractions rationnelles

Les fonctions polynomiales sont les fonctions définies sur \mathbb{R} de la forme

$$x \mapsto a_0 + a_1x + a_2x^2 + \dots + a_nx^n = \sum_{k=0}^n a_kx^k,$$

où n est un entier naturel et les a_i , $i = 0, \dots, n$ sont des nombres réels. Si tous les a_i , $1 \leq i \leq n$, sont nuls, cette fonction constante égale à a_0 ; si $a_n \neq 0$, l'entier n est le **degré du polynôme**. Pour $n = 1$, on parle de fonction affine ou linéaire (le graphe est une droite) et pour $n = 2$, on retrouve les polynômes du second degré (le graphe est une parabole).

Exemple 1.5 Les fonctions $x \mapsto 1 + 2x$, $x \mapsto x^2 - 3x^5$ sont des fonctions polynomiales de degré respectif 1 et 5.

Tous les fonctions polynomiales sont dérivables sur \mathbb{R} avec pour dérivée si $n \geq 1$:

$$x \mapsto a_1 + 2a_2x + \dots + na_nx^{n-1} = \sum_{k=0}^{n-1} (k+1)a_{k+1}x^k,$$

et $x \mapsto 0$ si $n = 0$ ou si la fonction est nulle.

On rappelle qu'en dehors des polynômes constants, les limites en $+\infty$ et $-\infty$ d'une fonction polynomiale sont $\pm\infty$, suivant la parité de n et le signe de $a_n \neq 0$.

Les fractions rationnelles sont des fonctions de la forme :

$$x \mapsto \frac{P(x)}{Q(x)} = \frac{a_0 + a_1x + a_2x^2 + \dots + a_nx^n}{b_0 + b_1x + b_2x^2 + \dots + b_mx^m},$$

c'est-à-dire le quotient des deux fonctions polynomiales, la fonction Q devant être non nulle. En général ces fonctions ne sont pas définies sur \mathbb{R} tout entier, mais uniquement sur \mathbb{R} privé des racines (ou des zéros) de Q , c'est-à-dire des nombres réels x tels que $Q(x) = 0$.

Exemple 1.6 — La fonction $x \mapsto \frac{x+1}{x^2+x+1}$ est une fraction rationnelle définie sur \mathbb{R} tout entier, car $x^2+x+1 > 0$ pour tout $x \in \mathbb{R}$ (discriminant strictement négatif).

— La fonction $x \mapsto \frac{2x^3-7x}{x^3-5x^2+6x}$ est une fraction rationnelle définie sur \mathbb{R} privé des points 0, 2 et 3 car $x^3-5x^2+6x = x(x-2)(x-3)$.

Il faut parfois faire attention toutefois à ce qu'un zéro du dénominateur peut aussi être un zéro du numérateur :

Exemple 1.7 Ainsi la fraction rationnelle définie par $x \mapsto \frac{x^3-x^2-2x}{x^2-4x}$ a pour ensemble de définition \mathbb{R} privé de 4. En effet on a $x^3-x^2-2x = x(x^2-x-2) = x(x-2)(x+1)$ et $x^2-4x = x(x-4)$, d'où une simplification possible par x .

Sur leur ensemble de définition, les fractions rationnelles sont dérivables. Pour déterminer leurs limites en $-\infty$ et $+\infty$, il y a une indétermination qu'on lève en mettant en facteur les termes de plus haut degré au numérateur et au dénominateur et en effectuant les simplifications adéquates pour se trouver avec une puissance de x multipliée par une fraction dont le comportement à l'infini ne pose pas de problème. Ainsi

$$\begin{aligned} \frac{x+1}{x^2+x+1} &= \frac{x(1+(1/x))}{x^2(1+(1/x)+(1/x^2))} = \frac{1}{x} \times \frac{1+(1/x)}{1+(1/x)+(1/x^2)}, \\ \frac{2x^3-7x}{x^3-5x^2+6x} &= \frac{x^3(2-(7/x^2))}{x^3(1-(5/x)+(6/x^2))} = x^0 \times \frac{2-(7/x^2)}{1-(5/x)+(6/x^2)}, \end{aligned}$$

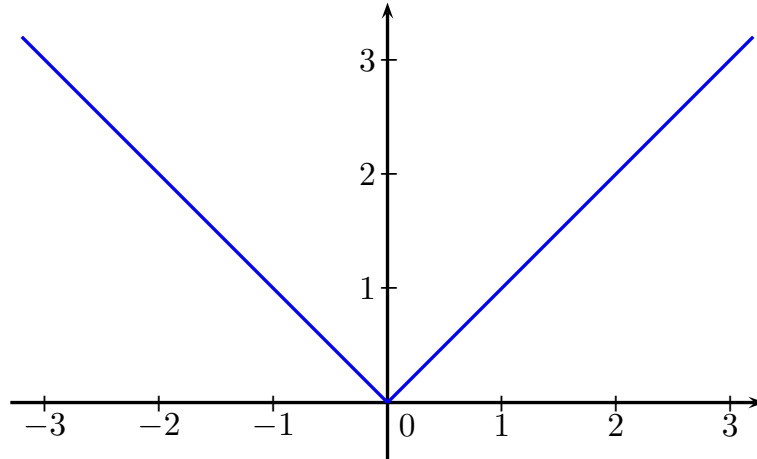


FIGURE 1.3 – Fonction valeur absolue

$$\frac{x^3 - x^2 - 2x}{x^2 - 4x} = \frac{x^3(1 - (1/x) - (2/x^2))}{x^2(1 - (4/x))} = x \times \frac{1 - (1/x) - (2/x^2)}{1 - (4/x)}.$$

On rappelle ensuite

Lemme 1.1

$$\lim_{x \rightarrow +\infty} x^n = \begin{cases} +\infty & \text{si } n \geq 1, \\ 1 & \text{si } n = 0, \\ 0 & \text{si } n \leq -1. \end{cases} \quad \text{et} \quad \lim_{x \rightarrow -\infty} x^n = \begin{cases} +\infty & \text{si } n \geq 1 \text{ et } n \text{ pair}, \\ -\infty & \text{si } n \geq 1 \text{ et } n \text{ impair}, \\ 1 & \text{si } n = 0, \\ 0 & \text{si } n \leq -1. \end{cases}$$

1.3.4 La fonction valeur absolue (★)

Définition 1.10 Soit x un nombre réel. La valeur absolue de x est le nombre réel défini par

$$|x| = \begin{cases} x & \text{si } x \geq 0, \\ -x & \text{si } x \leq 0. \end{cases}$$

La fonction valeur absolue $x \mapsto |x|$ est définie sur \mathbb{R} .

La valeur absolue d'un nombre x peut aussi être définie comme le plus grand des nombres x et $-x$. Rappelons quelques propriétés.

Propriétés 1.2 Pour tous nombres réels x et y ,

1. $|x| \geq 0$, $-|x| \leq x \leq |x|$, $|-x| = |x|$ et $|x| > 0 \Leftrightarrow x \neq 0$.
2. $\sqrt{x^2} = |x|$.
3. $|xy| = |x||y|$ et si $x \neq 0$, $|1/x| = 1/|x|$.
4. $|x + y| \leq |x| + |y|$ (inégalité triangulaire).
5. $||x| - |y|| \leq |x - y|$.

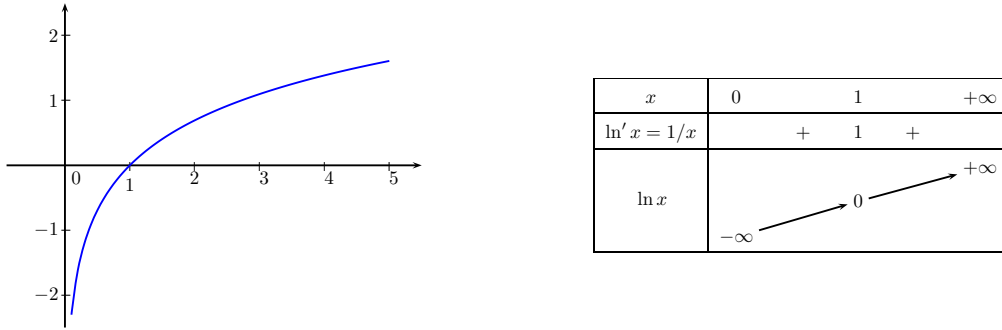


FIGURE 1.4 – Fonction logarithme

Proposition 1.2 Soit r un nombre réel strictement positif. Pour tous les nombres réels a et x

$$|x - a| \leq r \Leftrightarrow a - r \leq x \leq a + r, \quad |x - a| < r \Leftrightarrow a - r < x < a + r.$$

1.3.5 La fonction logarithme (★★).

Définition 1.11 On appelle logarithme népérien, que l'on note \ln , l'unique fonction définie sur $]0, +\infty[$ et à valeurs dans \mathbb{R} telle que

$$\forall x \in]0, +\infty[, \quad \ln'(x) = \frac{1}{x}, \quad \text{et } \ln(1) = 0.$$

Souvent on l'appelle simplement logarithme. Voici ces principales propriétés :

Propriétés 1.3 C'est une fonction continue, dérivable et strictement croissante sur $]0, +\infty[$.

1. Pour tout x et y strictement positifs,
 - (a) $\ln(xy) = \ln(x) + \ln(y)$
 - (b) $\ln(1/x) = -\ln(x)$
 - (c) $\ln(x/y) = \ln(x) - \ln(y)$
 - (d) et pour tout α réel, $\ln(x^\alpha) = \alpha \ln(x)$.
2. Enfin pour tout $x > 0$, $\ln(x) \leq x - 1$.

Rappelons également quelques limites usuelles du logarithme :

Propriétés 1.4

1. $\lim_{x \rightarrow 0} \ln(x) = -\infty$ et $\lim_{x \rightarrow +\infty} \ln(x) = +\infty$.
2. $\lim_{x \rightarrow 0} x \ln x = 0$, $\lim_{x \rightarrow 0} \frac{\ln(1+x)}{x} = 1$.
3. $\lim_{x \rightarrow +\infty} \frac{\ln(x)}{x} = 0$.

Remarque 1.2 La fonction $x \mapsto \ln(|x|)$ est définie sur $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$. On vérifiera que cette fonction est dérivable et a pour dérivée $1/x$ pour tout $x \in \mathbb{R}^*$.

Terminons ce paragraphe par le logarithme de base a .

Définition 1.12 Soit $a \in]0, 1[\cup]1, +\infty[$. On appelle logarithme de base a l'application définie sur $]0, +\infty[$ par

$$\forall x > 0, \quad \log_a(x) = \frac{\ln(x)}{\ln(a)}.$$

Notons que pour tout a , $\log_a(1) = 0$ et par définition $\log_a(a) = 1$. En physique on utilise fréquemment le logarithme de base 10. Il vérifie notamment pour tout $n \in \mathbb{Z}$, $\log_{10}(10^n) = n$, et il est noté \log sans préciser la base.

1.3.6 La fonction exponentielle (★★).

En rassemblant les propriétés de la fonction \ln , cette fonction est donc une bijection de $]0, +\infty[$ sur \mathbb{R} .

Définition 1.13 La bijection réciproque est la fonction exponentielle, notée \exp . Elle est définie sur \mathbb{R} et à valeurs dans $]0, +\infty[$.

Ainsi on a

$$\forall x > 0, \quad \exp(\ln x) = x, \quad \text{et} \quad \forall x \in \mathbb{R}, \quad \ln(\exp x) = x.$$

On en déduit également les propriétés suivantes :

Propriétés 1.5 La fonction \exp est continue et strictement croissante sur \mathbb{R} , dérivable avec $\exp'(x) = \exp(x)$ pour tout $x \in \mathbb{R}$. De plus pour tout x et y dans \mathbb{R} :

- $\exp(x + y) = \exp(x) \exp(y)$, d'où $\exp(nx) = (\exp(x))^n$ pour tout $n \in \mathbb{N}$,
- $\exp(-x) = 1/\exp(x)$,
- $\exp(x - y) = \exp(x)/\exp(y)$,
- pour tout $\alpha \in \mathbb{Q}$, $\exp(\alpha x) = (\exp(x))^\alpha$.

Concernant les limites usuelles concernant cette fonction, on obtient :

Propriétés 1.6

1. $\lim_{x \rightarrow -\infty} \exp(x) = 0$ et $\lim_{x \rightarrow +\infty} \exp(x) = +\infty$.
2. $\lim_{x \rightarrow -\infty} x \exp(x) = 0$, $\lim_{x \rightarrow 0} \frac{\exp(x) - 1}{x} = 1$.
3. $\lim_{x \rightarrow +\infty} \frac{\exp(x)}{x} = +\infty$.

Notation. Le nombre réel $\exp(1)$ se note e ; on a donc $\ln(e) = 1$. Puisque la fonction exponentielle est strictement croissante, il vient $\exp(1) > \exp(0)$, donc $e > 1$.

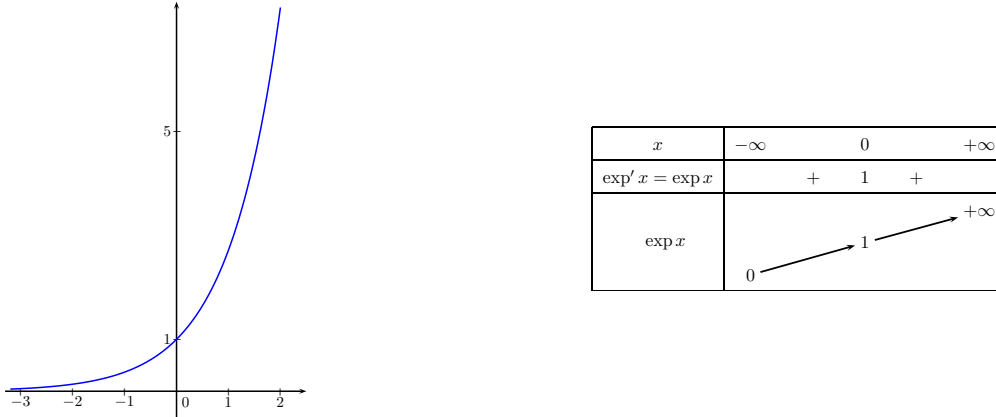


FIGURE 1.5 – Fonction exponentielle

1.3.7 Fonctions puissance (suite).

Soit a un nombre réel strictement positif.

— Pour tout entier $n \in \mathbb{Z}$, $\exp(n \ln a) = (\exp \ln a)^n = a^n$.

— Supposons que n est un entier positif au moins égal à 2 et posons $y = \exp\left(\frac{1}{n} \ln a\right)$. On a $y^n = \exp\left(n\left(\frac{1}{n}\right) \ln a\right) = a$. Puisque y est strictement positif, on en déduit $y = \sqrt[n]{a}$ par définition de la racine n -ième. On a donc

$$\sqrt[n]{a} = \exp\left(\frac{1}{n} \ln a\right), \text{ pour tout } a > 0.$$

Plus généralement

Définition 1.14 Soit a un nombre strictement positif et soit $b \in \mathbb{R}$. On définit le nombre réel a^b , appelé **a puissance b**, en posant

$$a^b = \exp(b \ln a).$$

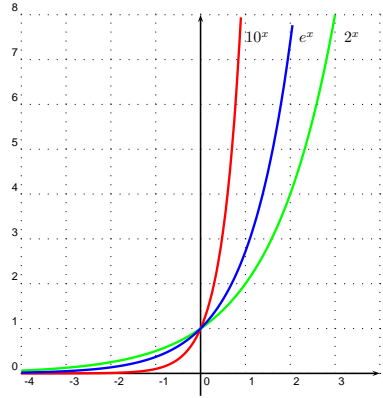
On peut donc élever un nombre **strictement positif** à une puissance réelle quelconque. Les règles de calcul sont ensuite celles dont on a l'habitude.

Proposition 1.3 Pour tous nombres réels b et c :

- $1^b = 1$.
- $x^{b+c} = x^b x^c$ et $(x^b)^c = x^{(bc)}$ pour tout $x > 0$;
- si $x > 0$ et $y > 0$, alors $(xy)^c = x^c y^c$;
- si $x > 0$, alors $x^{-c} = 1/(x^c)$.

Méthode. Pour étudier une expression de la forme a^b où b n'est pas un entier, revenez à la définition : $a^b = \exp(b \ln a)$.

Définition 1.15 Soit α un nombre réel. La fonction $f :]0, +\infty[\rightarrow \mathbb{R}$ définie par $f(x) = x^\alpha$ s'appelle la **fonction puissance d'exposant α** .

FIGURE 1.6 – Fonctions exponentielles de base a

Propriétés 1.7 Pour $\alpha \in \mathbb{R}^*$, la fonction puissance d'exposant α

1. est une application continue sur $]0, +\infty[$, strictement monotone (croissante si $\alpha > 0$ et décroissante si $\alpha < 0$),
2. est une bijection de $]0, +\infty[$ sur $]0, +\infty[$.
3. Enfin elle est dérivable sur $]0, +\infty[$ avec pour dérivée la fonction $x \mapsto \alpha x^{\alpha-1}$.

Concernant les limites on a

Propriétés 1.8

1. $\lim_{x \rightarrow +\infty} x^\alpha = \begin{cases} 0 & \text{si } \alpha < 0, \\ 1 & \text{si } \alpha = 0, \\ +\infty & \text{si } \alpha > 0. \end{cases}$
2. $\lim_{x \rightarrow 0} x^\alpha = \begin{cases} +\infty & \text{si } \alpha < 0, \\ 1 & \text{si } \alpha = 0, \\ 0 & \text{si } \alpha > 0. \end{cases}$

Fonction exponentielle de base a .

Définition 1.16 Soit a un nombre réel strictement positif. La fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ définie par $f(x) = a^x$ s'appelle la fonction exponentielle de base a .

En voici quelques propriétés.

Propriétés 1.9 Si $a \neq 1$, la fonction $x \mapsto a^x$ est une bijection continue de \mathbb{R} sur $]0, +\infty[$. Si $a > 1$ cette bijection est strictement croissante ; si $a < 1$, elle est strictement décroissante. De plus elle est dérivable sur \mathbb{R} avec pour dérivée $x \mapsto (\ln a)a^x$.

Remarquons que si $a = e = \exp(1)$, alors pour tout $x \in \mathbb{R}$, $\exp(x) = e^x$. La fonction exponentielle de base e est donc l'exponentielle ordinaire. On

utilisera par la suite indifféremment les deux notations. Enfin pour tout $x > 0$:

$$\forall x \in \mathbb{R}, \log_a(a^x) = x \quad \text{et} \quad \forall x > 0, a^{\log_a(x)} = x.$$

Autrement dit l'exponentielle en base a est la bijection réciproque du logarithme en base a .

1.4 Limite d'une fonction

1.4.1 Définitions précises

Soient I un intervalle et $f : I \rightarrow \mathbb{R}$ une fonction. Soit x_0 un nombre réel qui appartient à I ou bien est une extrémité de I .

Définition 1.17 (Limite) *Soit l un nombre réel. On dit que f a pour limite l en x_0 , ou encore que $f(x)$ tend vers l quand x tend vers x_0 , si pour tout nombre $\varepsilon > 0$, il existe un nombre $\eta > 0$ ayant la propriété suivante :*

$$(x \in I, \quad x \neq x_0, \quad \text{et} \quad |x - x_0| \leq \eta) \implies |f(x) - l| < \varepsilon.$$

Cette propriété se note $\boxed{\lim_{x \rightarrow x_0} f(x) = l.}$

Intuitivement cette définition signifie que $f(x)$ est aussi près que l'on veut de l à condition de choisir x assez près de x_0 , mais différent de x_0 .

Par définition il revient au même de dire que $f(x)$ tend vers l ou que $f(x) - l$ tend vers 0 quand x tend vers x_0 . On a ainsi les équivalences très utiles

$$\lim_{x \rightarrow x_0} f(x) = l \iff \lim_{x \rightarrow x_0} (f(x) - l) = 0 \iff \lim_{x \rightarrow x_0} |f(x) - l| = 0.$$

Remarque 1.3 *Dans la définition de la limite en x_0 , la fonction f n'a pas besoin d'être définie en x_0 et si elle l'est, la valeur $f(x_0)$ n'a aucune influence sur l'existence ou la valeur de la limite.*

Ainsi il est possible de chercher si $\frac{\sqrt{x} - 1}{x - 1}$ a une limite quand x tend vers 1.

Par ailleurs en ce qui concerne la limite en x_0 , seules comptent les valeurs que prend la fonction aux points x assez proches de x_0 , mais différents de x_0 . Aussi si on crée une fonction g en modifiant la fonction f au point x_0 et en dehors d'un intervalle $]a, b[$ tel que $x_0 \in]a, b[$, alors f a pour limite l en x_0 si et seulement si g a pour limite l en x_0 .

Définition 1.18 (Limite à l'infini) *Soit I l'un des intervalles $]-\infty, +\infty[$, $]a, +\infty[$ ou $]a, +\infty[$ où a est un nombre réel, et soit $f : I \rightarrow \mathbb{R}$ une fonction. Si l est un nombre réel, on dit que f a pour limite l en $+\infty$, ou encore que*

$f(x)$ tend vers l quand x tend vers $+\infty$, si pour tout nombre $\varepsilon > 0$, il existe un nombre $r > 0$ ayant la propriété suivante :

$$(x \in I, \text{ et } x > r) \implies |f(x) - l| < \varepsilon.$$

Cette propriété se note $\lim_{x \rightarrow +\infty} f(x) = l$.

Soit I l'un des intervalles $] -\infty, +\infty[$, $] -\infty, a]$ ou $] -\infty, a[$ où a est un nombre réel, et soit $f : I \rightarrow \mathbb{R}$ une fonction. Si l est un nombre réel, on dit que f a pour limite l en $-\infty$, ou encore que $f(x)$ tend vers l quand x tend vers $-\infty$, si pour tout nombre $\varepsilon > 0$, il existe un nombre $r < 0$ ayant la propriété suivante :

$$(x \in I, \text{ et } x < r) \implies |f(x) - l| < \varepsilon.$$

Cette propriété se note $\lim_{x \rightarrow -\infty} f(x) = l$.

Corollaire 1.1 Si une fonction a une limite, cette limite est unique.

Nous parlerons désormais de la limite d'une fonction en x_0 , en $+\infty$ ou $-\infty$. Mais attention la limite d'une fonction en un point n'existe pas toujours.

Définition 1.19 (Limite infinie)

- On dit que $f(x)$ tend vers $+\infty$ quand x tend vers x_0 , et l'on note $\lim_{x \rightarrow x_0} f(x) = +\infty$, si pour tout nombre $A > 0$, il existe un nombre $\eta > 0$ ayant la propriété suivante :

$$(x \in I, \quad x \neq x_0, \quad \text{et } |x - x_0| < \eta) \implies f(x) > A.$$

- Si I est l'un des intervalles $] -\infty, +\infty[$, $[a, +\infty[$ ou $]a, +\infty[$ où a est un nombre réel, on dit que $f(x)$ tend vers $+\infty$ quand x tend vers $+\infty$, et l'on note $\lim_{x \rightarrow +\infty} f(x) = +\infty$, si pour tout nombre $A > 0$, il existe un nombre $r > 0$ tel que

$$x > r \implies f(x) > A.$$

- Si I est l'un des intervalles $] -\infty, +\infty[$, $] -\infty, a]$ ou $] -\infty, a[$ où a est un nombre réel, on dit que $f(x)$ tend vers $+\infty$ quand x tend vers $-\infty$, et l'on note $\lim_{x \rightarrow -\infty} f(x) = +\infty$, si pour tout nombre $A > 0$, il existe un nombre $r < 0$ tel que

$$x < r \implies f(x) > A.$$

- On dit que $f(x)$ tend vers $-\infty$ quand x tend vers x_0 (ou bien quand x tend vers $+\infty$, ou bien quand x tend vers $-\infty$), si $-f(x)$ tend vers $+\infty$. Cette propriété se note $\lim_{x \rightarrow x_0} f(x) = -\infty$ dans le cas, par exemple, de la limite en x_0 .

1.4.2 Opérations sur les limites (★★)

Les **formes indéterminées** correspondent à des cas où il n'y a pas de règle générale permettant de déterminer la limite. Par exemple il n'y a pas de résultat général pour le produit d'une fonction qui tend vers zéro par une fonction qui tend vers $+\infty$: selon les cas, le résultat peut d'ailleurs être 0 ou $\pm\infty$ ou une limite finie non nulle, ou bien il n'y a pas de limite ; on dit que $0 \times \infty$ est une forme indéterminée. Il existe d'autres formes indéterminées comme $\frac{\infty}{\infty}$, $\frac{0}{0}$, $(+\infty - \infty)$, 1^∞ , ou ∞^0 . Pour lever les indéterminations, c'est-à-dire pour voir si de telles expressions ont une limite et éventuellement calculer cette limite, il suffit parfois de transformer convenablement l'expression (ce qui n'est pas toujours simple) et de se ramener aux énoncés précédents. Souvent il faudra faire appel à des techniques plus sophistiquées (dérivabilité, développements limités, etc.).

Les tableaux ci-dessous résument les cas usuels. Soient f et g deux fonctions définies sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} et a dans I ou un bord de I (peut donc être $+\infty$ ou $-\infty$). l et l' sont deux réels finis. **Pour la somme** :

$\lim_{x \rightarrow a} f(x)$	$\lim_{x \rightarrow a} g(x)$	$\lim_{x \rightarrow a} (f(x) + g(x))$
l	l'	$l + l'$
l	$+\infty$	$+\infty$
l	$-\infty$	$-\infty$
$+\infty$	$+\infty$	$+\infty$
$-\infty$	$-\infty$	$-\infty$
$+\infty$	$-\infty$	forme indéterminée

Pour le produit :

$\lim_{x \rightarrow a} f(x)$	$\lim_{x \rightarrow a} g(x)$	$\lim_{x \rightarrow a} (f(x)g(x))$
l	l'	ll'
l avec $l > 0$	$+\infty$	$+\infty$
l avec $l > 0$	$-\infty$	$-\infty$
l avec $l < 0$	$+\infty$	$-\infty$
l avec $l < 0$	$-\infty$	$+\infty$
0	$+\infty$	forme indéterminée
0	$-\infty$	forme indéterminée

Si une fonction g tend vers 0 quand x tend vers a avec $g(x) > 0$ (resp. $g(x) < 0$) pour tout x proche de a , la limite est notée 0^+ (resp. 0^-) :

$$\lim_{x \rightarrow a} g(x) = 0^+ \quad \text{ou} \quad \lim_{x \rightarrow a} g(x) = 0^-.$$

Pour le quotient :

$\lim_{x \rightarrow a} f(x)$	$\lim_{x \rightarrow a} g(x)$	$\lim_{x \rightarrow a} \frac{f(x)}{g(x)}$
l	$l', \text{ avec } l' \neq 0$	l/l'
l	$+\infty$	0
l	$-\infty$	0
$l \text{ avec } l > 0$	0^+	$+\infty$
$l \text{ avec } l > 0$	0^-	$-\infty$
$l \text{ avec } l < 0$	0^+	$-\infty$
$l \text{ avec } l < 0$	0^-	$+\infty$
0	0	forme indéterminée
$\pm\infty$	$\pm\infty$	forme indéterminée

Pour les fonctions usuelles (puissance, logarithme et exponentielle), on peut supprimer une indétermination grâce au théorème suivant.

Théorème 1.2 (Croissances comparées)

1. Si α et β sont deux nombres réels et si $\alpha < \beta$, alors

$$\lim_{x \rightarrow 0^+} \frac{x^\beta}{x^\alpha} = 0, \quad \lim_{x \rightarrow +\infty} \frac{x^\beta}{x^\alpha} = +\infty.$$

2. Si $\alpha > 0$ et $\beta \in \mathbb{R}$,

$$\lim_{x \rightarrow 0^+} x^\alpha |\ln x|^\beta = 0.$$

3. Si $\alpha > 0$ et $\beta \in \mathbb{R}$, au voisinage de $+\infty$,

$$\lim_{x \rightarrow +\infty} x^\beta \exp(-\alpha x) = 0, \quad \lim_{x \rightarrow +\infty} x^\beta \exp(\alpha x) = +\infty.$$

4. Si $\alpha > 0$ et $\beta \in \mathbb{R}$,

$$\lim_{x \rightarrow +\infty} \frac{(\ln x)^\beta}{x^\alpha} = 0.$$

1.4.3 Continuité d'une fonction

Soit I un intervalle et soit $f : I \rightarrow \mathbb{R}$ une fonction.

Définition 1.20 Si $x_0 \in I$, on dit que la fonction f est continue en x_0 si $\lim_{x \rightarrow x_0} f(x) = f(x_0)$. La fonction f est continue sur I , si quel que soit $x_0 \in I$, f est continue en x_0 .

En utilisant la définition de la limite en un point x_0 et en remarquant que si $x = x_0$, alors $f(x) - f(x_0) = 0$, on obtient que f est continue en x_0 si et seulement si pour tout $\varepsilon > 0$ il existe $\eta > 0$ tel que

$$(x \in I, \quad \text{et } |x - x_0| < \eta) \implies |f(x) - f(x_0)| < \varepsilon.$$

On peut enlever dans l'expression précédente $x \neq x_0$.

Exemple 1.8

- Une fonction constante sur I est continue sur I .
- La fonction racine carrée est continue sur $[0, +\infty[$.
- La fonction valeur absolue (voir paragraphe 1.3.4) est continue sur \mathbb{R} .

La proposition sur la limite d'une somme, d'un produit et d'un inverse permet d'énoncer :

Proposition 1.4 *Soient f et g des fonctions de I dans \mathbb{R} . Supposons que f et g sont continues en $x_0 \in I$.*

- Les fonctions $f + g$, fg et λf pour tout $\lambda \in \mathbb{R}$, sont continues en x_0 .
- Si $g(x_0) \neq 0$, alors la fonction f/g est continue en x_0 .
- La composée de deux fonctions continues est continue.

Une fonction polynomiale est continue sur \mathbb{R} . Si f est une fonction rationnelle (quotient de deux fonctions polynômes) définie sur un intervalle I , alors f est continue sur I .

1.5 Dérivée

La dérivée est l'outil principal pour étudier une fonction. Pour bien utiliser cette notion, il faut connaître parfaitement la définition et s'entraîner à calculer des dérivées rapidement et sans erreur.

1.5.1 Dérivée en un point et fonction dérivée

Soient I un intervalle, $f : I \rightarrow \mathbb{R}$ une fonction et x_0 un élément de I .

Définition 1.21 *On dit que f est dérivable en x_0 si $\frac{f(x) - f(x_0)}{x - x_0}$ a une limite finie quand x tend vers x_0 . La limite $\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$ est notée $f'(x_0)$ et s'appelle le nombre dérivé de f en x_0 .*

On dit que f est dérivable sur I si quel que soit $x_0 \in I$, f est dérivable en x_0 . Dans ce cas, la fonction $f' : I \rightarrow \mathbb{R}$ qui à x associe $f'(x)$, s'appelle la dérivée de f .

On note $D(I)$ l'ensemble des fonctions dérivables sur I .

Supposons f dérivable en x_0 et définissons une fonction ε en posant

$$\varepsilon(x) = \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \quad \text{si } x \neq x_0, \quad \text{et } \varepsilon(x_0) = 0.$$

Pour tout nombre $x \neq x_0$, on a

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + (x - x_0)\varepsilon(x)$$

et cette égalité est encore vraie si $x = x_0$ car dans ce cas les deux membres sont égaux à $f(x_0)$. Par ailleurs

$$\lim_{x \rightarrow x_0} \varepsilon(x) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) = f'(x_0) - f'(x_0) = 0 = \varepsilon(x_0),$$

donc la fonction ε est continue en x_0 .

Finalement si f est dérivable en x_0 , il existe une fonction ε continue en x_0 telle que $\varepsilon(x_0) = 0$ et

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + (x - x_0)\varepsilon(x) \quad \text{quel que soit } x \in I.$$

Cette propriété caractérise les fonctions dérivables en x_0 .

Proposition 1.5 *La fonction f est dérivable en x_0 si et seulement s'il existe un nombre réel a et une fonction ε ayant les propriétés suivantes :*

- ε est continue en x_0 et $\varepsilon(x_0) = 0$,
- $f(x) = f(x_0) + a(x - x_0) + (x - x_0)\varepsilon(x)$ pour tout $x \in I$.

Dans ce cas, le nombre a est égal à $f'(x_0)$.

Corollaire 1.2 *Si f est dérivable en x_0 , alors elle est continue en x_0 .*

Monotonie et dérivée.

Lemme 1.2 *Soient I un intervalle ouvert et $f : I \rightarrow \mathbb{R}$ une fonction dérivable sur I .*

- f est croissante sur I si et seulement si pour tout $x \in I$, $f'(x) \geq 0$.
- Si f est décroissante sur I si et seulement si pour tout $x \in I$, $f'(x) \leq 0$.

Tangente au graphe de f .

Soit \mathcal{C} le graphe de la fonction f dans le plan. Notons M_0 le point $(x_0, f(x_0))$ et si $x \in I$, $x \neq x_0$, notons M le point $(x, f(x))$; par définition les points M_0 et M appartiennent à \mathcal{C} .

Le rapport $\frac{f(x) - f(x_0)}{x - x_0}$ est la *pente* de la droite passant par M_0 et M .

Supposons que f est dérivable en x_0 . Alors intuitivement, quand x tend vers x_0 , la droite (M_0M) a pour position limite la droite passant par M_0 et de pente $f'(x_0)$. Par définition cette droite s'appelle la *tangente* à \mathcal{C} au point M_0 . Ainsi :

Propriétés 1.10 *Si f est dérivable en $x_0 \in I$, la courbe \mathcal{C} a pour tangente au point M_0 la droite d'équation $y = (x - x_0)f'(x_0) + f(x_0)$.*

Dérivées successives

Soient I un intervalle ouvert et $f : I \rightarrow \mathbb{R}$ une fonction dérivable; par définition cela signifie que f est dérivable en tout point de I . Nous avons alors défini la fonction dérivée $f' : I \rightarrow \mathbb{R}$ qui à tout x appartenant à I associe le nombre dérivé $f'(x)$.

Si la fonction f' est à son tour dérivable en tout point de I , alors la fonction (f') dérivée de f' est définie sur I ; cette fonction se note f'' et s'appelle la **dérivée seconde** de f . Plus généralement si n est un entier positif ou nul, on définit, si elle existe, la **dérivée n -ième** de f en posant $f^{(0)} = f$ par convention et

$$f^{(p)} = (f^{(p-1)})' \text{ pour tout entier } p \text{ tel que } 1 \leq p \leq n.$$

Si la dérivée n -ième de f existe, on dit que f est n fois dérivable.

1.5.2 Calcul des dérivées (★★)

Dérivée d'une somme et du produit par une constante. Si f et g sont deux fonctions dérivables en x_0 , alors les fonctions $f + g$ et λf pour tout $\lambda \in \mathbb{R}$, sont dérivables en x_0 et

$$\boxed{(f + g)'(x_0) = f'(x_0) + g'(x_0), \quad (\lambda f)'(x_0) = \lambda f'(x_0)}.$$

Dérivée d'un produit. Si f et g sont deux fonctions dérivables en x_0 , alors la fonction fg est dérivable en x_0 et

$$\boxed{(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)}.$$

Dérivée d'une fonction constante. Soit $u : I \rightarrow \mathbb{R}$ une fonction constante et soit x_0 un élément de I . Quel que soit $x \in I$, nous avons $u(x) = u(x_0)$. Le rapport $\frac{u(x) - u(x_0)}{x - x_0}$ est égal à 0, par conséquent $u'(x_0) = 0$. Ainsi une fonction constante a une dérivée nulle en tout point.

Dérivée d'une composée. Soient f et g deux fonctions telles que la composée $g \circ f$ est définie. Si f est dérivable en x_0 et si g est dérivable en $f(x_0)$, alors la fonction $g \circ f$ est dérivable en x_0 et

$$\boxed{(g \circ f)'(x_0) = g'(f(x_0))f'(x_0)}.$$

Dérivée de l'inverse. Soit f dérivable en x_0 . Si $f(x_0) \neq 0$, alors la fonction $1/f$ est dérivable en x_0 et

$$\boxed{\left(\frac{1}{f}\right)'(x_0) = -\frac{f'(x_0)}{(f(x_0))^2}}.$$

En utilisant les formules donnant la dérivée d'un produit et d'un inverse, on obtient :

Corollaire 1.3 *Si f et g sont des fonctions dérivables en x_0 et si $g(x_0) \neq 0$, alors la fonction f/g est dérivable en x_0 et*

$$\boxed{\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{(g(x_0))^2}}.$$

En conséquence on peut montrer que

- Une fonction polynôme est dérivable sur \mathbb{R} et la fonction dérivée est une fonction polynôme.
- Une fonction rationnelle est dérivable sur son domaine de définition et la fonction dérivée est une fonction rationnelle.

Dérivée d'une fonction réciproque. Soient I un intervalle ouvert et f une fonction dérivable et strictement monotone sur I . Posons $J = f(I)$ et notons $f^{-1} : J \rightarrow I$ la bijection réciproque de $f : I \rightarrow J$. Si l'on a $f'(x) \neq 0$ pour tout $x \in I$, alors f^{-1} est dérivable sur J et l'on a pour tout $x \in J$:

$$(f^{-1})'(x) = \frac{1}{f'(f^{-1}(x))}.$$

1.5.3 Extremum (local) d'une fonction

Définition 1.22 (Extremum) Soient I un intervalle et $f : I \rightarrow \mathbb{R}$ une fonction. Si $x_0 \in I$, on dit que

- f a un maximum en x_0 si pour tout $x \in I$, $f(x) \leq f(x_0)$;
- f a un minimum en x_0 si pour tout $x \in I$, $f(x) \geq f(x_0)$;
- f a un extremum en x_0 si f a un maximum ou un minimum en x_0 .

Définition 1.23 (Extremum local) Soient I un intervalle et $f : I \rightarrow \mathbb{R}$ une fonction. Si $x_0 \in I$, on dit que

- f a un maximum local en x_0 s'il existe un intervalle ouvert J de centre x_0 et contenu dans I , tel que $f(x) \leq f(x_0)$ pour tout nombre x appartenant à J ;
- f a un minimum local en x_0 s'il existe un intervalle ouvert J de centre x_0 et contenu dans I , tel que $f(x) \geq f(x_0)$ pour tout nombre x appartenant à J ;
- f a un extremum local en x_0 si f a un maximum local ou un minimum local en x_0 .

Remarque 1.4 Si une fonction $f :]a, b[\rightarrow \mathbb{R}$ a un maximum (ou un minimum) en un point $x_0 \in]a, b[$, alors f a aussi un maximum local (ou un minimum local) en x_0 . Mais l'inverse n'est pas vrai.

Exemple 1.9 1. Considérons la fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ définie par $f(x) = |1 - x^2|$. Si $x \in]-1, 1[$, on a $x^2 < 1$, donc $f(x) = 1 - x^2$. On en déduit que si $x \in]-1, 1[$, alors $f(x) \leq 1$, c'est-à-dire $f(x) \leq f(0)$. La fonction f a donc un maximum local en 0. Hors de cet intervalle, la fonction f peut prendre des valeurs supérieures à 1 : ainsi $f(4) = 15$.

Pour tout $x \in \mathbb{R}$, $f(x) \geq 0$ et $f(1) = f(-1) = 0$. Donc la fonction f atteint son minimum global (par opposition à local) en 1 et -1. En ces points f a aussi un minimum local.

2. La fonction $g : [0, +\infty[\rightarrow \mathbb{R}$ définie par $g(x) = |1 - x^2|$ a encore un minimum local (et global) en 1, mais n'a pas de maximum local en 0.

Lorsqu'une fonction f est dérivable, le théorème suivant donne une condition nécessaire pour que f ait un extremum local en un point.

Théorème 1.3 *Soient I un intervalle ouvert et $f : I \rightarrow \mathbb{R}$ une fonction. Supposons que f a un extremum local en un point $x_0 \in I$ et que f est dérivable en x_0 . Alors $f'(x_0) = 0$.*

Dans l'exemple 1.9, cas 2, la fonction f a un maximum local en 0, et pour tout $x \in]-1, 1[$, $f(x) = 1 - x^2$. Donc $f'(x) = -2x$ et ainsi $f'(0) = 0$, conformément au théorème. En revanche au point 1, f a un minimum local mais n'est pas dérivable en ce point. En effet si $x > 0$ et $x \neq 1$, on a

$$\frac{f(x) - f(1)}{x - 1} = \frac{|1+x||1-x|}{x-1} = (1+x) \frac{|1-x|}{x-1} = \begin{cases} x+1 & \text{si } x > 1 \\ -(x+1) & \text{si } 0 < x < 1. \end{cases}$$

$$\text{Ainsi } \lim_{x \rightarrow 1^-} \frac{f(x) - f(1)}{x - 1} = -2 \text{ et } \lim_{x \rightarrow 1^+} \frac{f(x) - f(1)}{x - 1} = 2.$$

1.6 Étude de fonctions

À partir de maintenant, l'étude d'une fonction f suivra toujours la trame suivante.

1. Comprendre comment la fonction est construite à partir des fonctions usuelles.
2. Déterminer l'ensemble de définition D_f de f .
3. Justifier que f est continue et/ou dérivable sur D_f (ou un sous-ensemble de D_f).
4. Calculer la dérivée f' et déterminer le signe de la dérivée.
5. Trouver les limites au bord du domaine de définition.
6. Établir le tableau de variations.

Concernant le graphe de f , on rappelle que doivent figurer :

- les points où la fonction n'est pas définie ;
- les points où la dérivée s'annule (tangente horizontale) ;
- et plus généralement toutes les valeurs remarquables du tableau de variations.

En particulier

- si f admet une limite infinie en un point $x_0 \in \mathbb{R}$, alors une asymptote verticale doit figurer.
- Si f admet une limite finie en $+\infty$ ou en $-\infty$, alors une asymptote horizontale doit être représenté.

Voici un exemple détaillé pour illustrer notre propos. On souhaite étudier la fonction

$$f(x) = \ln\left(\frac{x}{x+1}\right) + 1 + \frac{1}{2x}.$$

1. Il s'agit de la somme de trois fonctions : $x \mapsto \ln\left(\frac{x}{x+1}\right)$, la fonction constante égale à 1, et la fonction $x \mapsto \frac{1}{2x}$. La première est la composition du logarithme avec $x \mapsto \frac{x}{x+1}$ qui est une fraction rationnelle et la troisième est l'inverse d'une fonction affine.
2. Pour que la première soit correctement définie, il faut d'abord ne pas diviser par zéro dans la fraction rationnelle, soit $x + 1 \neq 0$ ou encore $x \neq -1$. Ensuite il faut ne pas prendre le logarithme d'un nombre négatif ou nul, donc

$$\frac{x}{x+1} > 0.$$

Pour qu'une fraction soit positive, il faut que les deux termes soient ou positifs ou négatifs, donc

$$x > 0 \text{ et } x + 1 > 0, \quad \text{ou} \quad x < 0 \text{ et } x + 1 < 0,$$

ce qui revient à

$$x > 0 \quad \text{ou} \quad x < -1.$$

Donc la première fonction est définie sur $] -\infty, -1[\cup] 0, +\infty[$. La fonction constante est toujours bien définie. Quant à la troisième, il faut éviter la division par zéro, donc $x \neq 0$. Si on rassemble toutes ces conditions, on obtient :

$$D_f =] -\infty, -1[\cup] 0, +\infty[.$$

3. Les fractions rationnelles $x \mapsto \frac{x}{x+1}$ et $x \mapsto \frac{1}{2x}$ sont continues et dérivables sur leur ensemble de définition. Idem pour le logarithme. Une fonction constante est toujours dérivable. Donc par composition et somme, f est continue et dérivable sur D_f .
4. On applique les règles de dérivation pour obtenir pour $x \in D_f$

$$\begin{aligned} f'(x) &= \frac{1}{\frac{x}{x+1}} \left(\frac{1 \times (x+1) - x \times 1}{(x+1)^2} \right) + 0 - \frac{2}{(2x)^2} \\ &= \frac{1}{x(x+1)} - \frac{1}{2x^2}. \end{aligned}$$

Cherchons les $x \in D_f$ tels que $f'(x) > 0$, soit

$$\frac{1}{x(x+1)} > \frac{1}{2x^2}.$$

Séparons les cas $x > 0$ et $x < -1$.

- Si $x > 0$, on multiplie tout par x , soit $\frac{1}{x+1} > \frac{1}{2x}$, d'où : $2x > x+1$. Ainsi si $x > 0$, $f'(x) > 0$ si et seulement si $x > 1$. Et $f'(1) = 0$.

— Si $x < -1$, on multiplie par x , mais en faisant attention que l'inégalité va se « retourner ».

$$f'(x) > 0 \Leftrightarrow \frac{1}{x+1} < \frac{1}{2x}.$$

Et donc $2x < x+1$, soit $x < 1$, ce qui est toujours vrai car $x < -1$.

5. Passons à la recherche des limites. Commençons par la limite en $+\infty$. Quand x tend vers $+\infty$, $\frac{x}{x+1}$ tend vers 1, donc le logarithme vers 0. $\frac{1}{2x}$ tend vers 0, donc $f(x)$ tend vers 1.

Quand x tend vers 0 avec $x > 0$, c'est un peu plus compliqué. En effet $\frac{x}{x+1}$ tend vers 0, donc le logarithme vers $-\infty$. Mais comme $\frac{1}{2x}$ tend vers $+\infty$, on a une forme indéterminée $\infty - \infty$. Pour la lever écrivons :

$$f(x) = \frac{1}{2x} \left[2x \ln \left(\frac{x}{x+1} \right) + 2x + 1 \right] = \frac{1}{2x} [2x \ln(x) - 2x \ln(x+1) + 2x + 1].$$

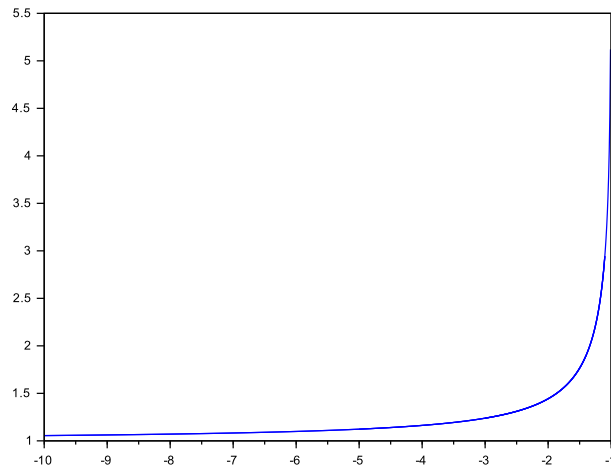
Ce qui est à l'intérieur des crochets tend vers 1 (voir le théorème 1.2). Donc $f(x)$ tend vers $+\infty$ quand x tend vers 0 avec $x > 0$.

Quand x tend vers -1 avec $x < -1$, la fraction rationnelle $\frac{x}{x+1}$ tend vers $+\infty$, donc le logarithme et $f(x)$ aussi. Enfin pour la limite en $-\infty$, on obtient encore 1.

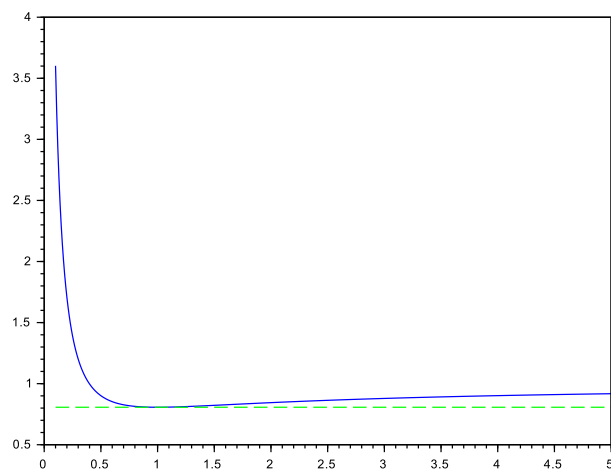
6. Voici le tableau de variations de f .

x	$-\infty$	-1	0	1	$+\infty$
$f'(x)$	+		-	0	+
$f(x)$	1	\nearrow $+\infty$	$+\infty$	\searrow $-\ln(2) + \frac{3}{2}$	\nearrow 1

Voici le graphe de f sur $] -\infty, -1[$



puis sur $]0, +\infty[$.



1.7 Modélisation de l'évolution d'un pathogène

Dans le monde des micro-pathogènes il y a quelques virus et bactéries extrêmement virulents qui conduisent généralement à la mort de leur hôte. C'est le cas par exemple de la fièvre hémorragique Ebola, du virus de la variole ou encore de certaines formes d'anthrax. Ces pathologies quoique fortement médiatisées, restent néanmoins exceptionnelles. La grande majorité des agents infectieux induisent fort heureusement des maladies aux conséquences moins dramatiques.

D'après la théorie de la sélection naturelle de Darwin (XIX^e siècle) l'évolution d'une espèce se fait dans le sens d'une plus grande compétitivité : un

individu sera d'autant plus favorisé qu'il engendrera davantage de descendants que ses compétiteurs. Dans le cas des micro-pathogènes cette « règle » affirme que si plusieurs souches d'une même bactérie ou d'un même virus sont en compétition, la souche favorisée sera la plus adaptée à se développer et à diffuser au sein de la population. Mais alors comment expliquer que la plupart des microbes soient aussi peu agressifs et virulents? Le biologiste peut fournir un modèle exprimant (avec une formule mathématique) l'efficacité de transmission d'un agent pathogène en fonction de sa virulence intrinsèque. La sélection naturelle tend alors à maximiser cette efficacité et là réside l'explication de la grande diversité du monde des micro-pathogènes.

Il faut montrer comment des hypothèses de nature biologique (sur la nature du pathogène, la façon dont il se propage, etc.) peuvent conduire à un modèle mathématique (c'est-à-dire concrètement à une formule, donc une fonction) décrivant la transmission du pathogène. Le premier travail d'une bonne modélisation consiste à bien définir les quantités que l'on souhaite calculer et les informations dont on dispose. Pour cela il importe de faire le tri, au sein des modèles biologiques (qui peuvent être extrêmement complexes), entre les phénomènes pertinents et ceux qui ne joueront pas (ou peu) de rôle dans l'étude.

Ici l'objet d'intérêt est l'efficacité de la transmission d'un pathogène. Nous pouvons mesurer cette efficacité par le *nombre moyen d'individus infectés par un porteur de pathogène au cours de sa vie*. Il s'agit d'un nombre que nous noterons R . L'étude de R ou plus exactement celle de son comportement en fonction des autres paramètres biologiques pertinents (en bref l'étude de la *fonction* R) devient donc à partir de maintenant notre objectif principal.

Établissons la liste des différents paramètres qui peuvent influencer sur la valeur de R . Un pathogène se transmet d'autant plus efficacement qu'il réussit à partir d'un individu à en infecter le plus grand nombre possible. Les infections seront d'autant plus nombreuses qu'il y a beaucoup d'individus sains dans le milieu concerné. Dans un modèle simple, on considère que, chaque jour¹, le nombre d'infections provoquées par le porteur est proportionnel au nombre d'individus sains disponibles dans le milieu (ce nombre sera noté S). La constante de proportionnalité notée β dépend directement des caractéristiques du pathogène autrement dit de sa contagiosité intrinsèque (la probabilité que l'infection soit réussie). L'efficacité de la transmission est donc le produit de S par ce nombre β , qu'il faut ensuite multiplier par la durée de vie en jours d'un individu infecté. On admet que la durée de vie moyenne d'un individu infecté est l'inverse du taux de mortalité de la population à laquelle il appartient, sous réserve que cette dernière suive un modèle dit malthusien : elle vaut donc $1/(\mu + \alpha)$ où μ est le taux de mortalité moyen de la population totale (individus infectés et non infectés) et α est une surmortalité provoquée par la maladie. C'est cette surmortalité qui est appelée *virulence* du parasite. Le nombre moyen d'individus infectés par un porteur

1. si l'unité de temps est la journée

de pathogène au cours de sa vie est donné par l'expression :

$$R = \frac{\beta S}{\mu + \alpha}.$$

Les caractéristiques du parasite sont α (virulence) et β (contagiosité). Elles ne sont pas indépendantes : un pathogène très virulent (α grand) a souvent une contagiosité β élevée. Pour une infection virale, la charge virale (nombre de virus dans l'organisme) détermine la sévérité des symptômes et la probabilité de transmission. Les modèles les plus courants font l'hypothèse que β est une fonction croissante de α . En conclusion R est une fonction de α donnée par :

$$R = \frac{\beta(\alpha)S}{\mu + \alpha}.$$

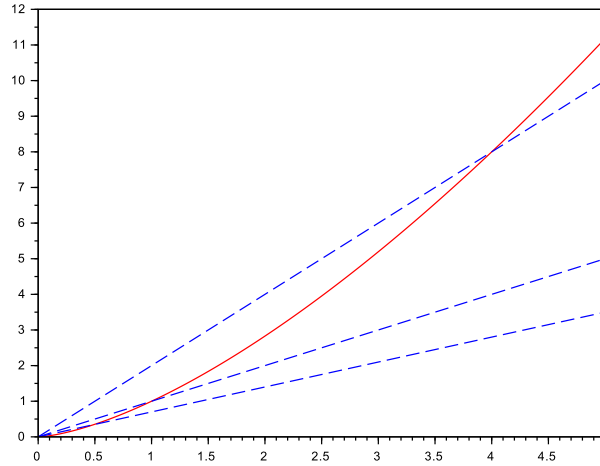
Pourquoi la virulence α de la plupart des pathogènes tend-elle à se stabiliser autour d'une valeur peu élevée, et pourquoi certains conservent-ils une virulence extrêmement forte ? Les lois de l'évolution indiquent que la fonction R a tendance à prendre la valeur la plus grande possible. En termes mathématiques, cela signifie qu'il faut étudier les maxima de R et trouver les valeurs de α pour lesquelles ces maxima sont atteints. Selon l'expression précise de la fonction $\beta(\alpha)$ nous allons voir que des phénomènes très différents peuvent se produire.

1.7.1 Premier cas : β sur-linéaire

Ici nous supposons que $\beta(\alpha) = r\alpha^c$ avec r et c des constantes, $r > 0$ et $c > 1$. Le nombre α est strictement positif, et le domaine de R est donc $D_R =]0, +\infty[$. On rappelle que

$$\beta(\alpha) = r\alpha^c = r \exp(c \ln(\alpha))$$

et β croît vite lorsque α augmente : pour n'importe quelle constante K , le graphe de β est au-dessus de celui de la fonction linéaire $\alpha \mapsto K\alpha$ lorsque α est grand. C'est ce qu'on appelle un comportement *sur-linéaire*. Sur le graphe suivant on a tracé en rouge $\alpha \mapsto \alpha^{1.5}$ et en bleu pointillé diverses fonctions $\alpha \mapsto K\alpha$:



On fixe maintenant

$$R(\alpha) = S \frac{\beta(\alpha)}{\mu + \alpha} = rS \frac{\alpha^c}{\mu + \alpha} = rS \frac{\exp(c \ln(\alpha))}{\mu + \alpha}.$$

Elle ressemble à une fraction rationnelle à ceci près que c n'est pas forcément une puissance entière. Elle est définie sur $]0, +\infty[$ et continue et dérivable sur cet ensemble (par composition et quotient). Dérivons β (par la formule pour les fonctions composées) :

$$\beta'(\alpha) = r \exp'(c \ln(\alpha))(c \ln(\alpha))' = r \exp(c \ln(\alpha)) \frac{c}{\alpha} = rc \frac{\alpha^c}{\alpha} = rc\alpha^{c-1}.$$

Puis par dérivation d'un quotient :

$$\begin{aligned} R'(\alpha) &= \frac{(\mu + \alpha)S\beta'(\alpha) - S\beta(\alpha)(\mu + \alpha)'}{(\mu + \alpha)^2} = \frac{(\mu + \alpha)Src\alpha^{c-1} - r\alpha^c S}{(\mu + \alpha)^2} \\ &= rS\alpha^{c-1} \left(\frac{(\mu + \alpha)c - \alpha}{(\mu + \alpha)^2} \right) = rS\alpha^{c-1} \left(\frac{c\mu + (c-1)\alpha}{(\mu + \alpha)^2} \right). \end{aligned}$$

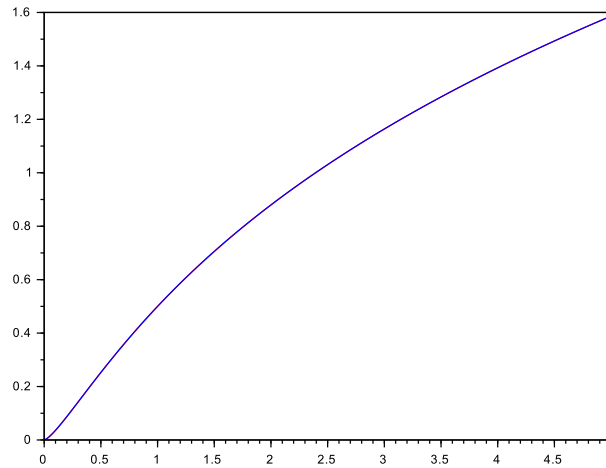
Rappelons que $c > 1$, $\mu > 0$ et $\alpha > 0$. Donc $c\mu + (c-1)\alpha > 0$. R est donc strictement croissante sur $]0, +\infty[$ et $R(0) = 0$ tandis que

$$\begin{aligned} \lim_{\alpha \rightarrow +\infty} R(\alpha) &= \lim_{\alpha \rightarrow +\infty} rS \frac{\alpha^c}{\mu + \alpha} \\ &= \lim_{\alpha \rightarrow +\infty} rS\alpha^{c-1} \frac{1}{1 + \mu/\alpha} = +\infty. \end{aligned}$$

On obtient donc le tableau de variations suivant :

α	0	$+\infty$
$R'(\alpha)$	+	
$R(\alpha)$	0	$+\infty$

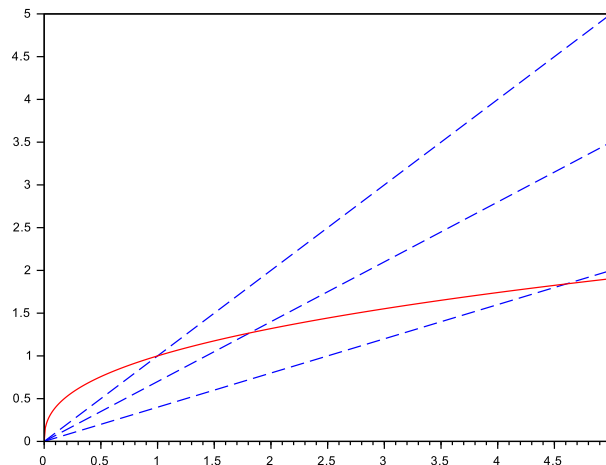
Et le graphe de R est ainsi :



On constate ici que R n'a pas de maximum : ses valeurs sont d'autant plus grandes que le taux de virulence α est élevé. Ce modèle correspond aux quelques maladies rares (Ebola, etc.) qui ont évolué vers une virulence extrême. Ces pathologies ont une contagiosité naturelle $\beta(\alpha)$ tellement élevée qu'il ne leur importe pas de préserver la vie de leur hôte, ce dernier pouvant spontanément infecter beaucoup d'individus, même en un temps très court.

1.7.2 Second cas : β sous-linéaire

Nous supposons toujours que $\beta(\alpha) = r\alpha^c$ mais cette fois avec $0 < c < 1$. La représentation graphique suivante montre que pour toute constante K , la fonction β est en dessous de $\alpha \mapsto K\alpha$ lorsque α est grand : ce caractère est dit *sous-linéaire*.



L'expression de R' est la même que précédemment et cette dérivée a toujours le signe de $c\mu + (c - 1)\alpha$:

$$R'(\alpha) = rS\alpha^{c-1} \left(\frac{c\mu + (c - 1)\alpha}{(\mu + \alpha)^2} \right).$$

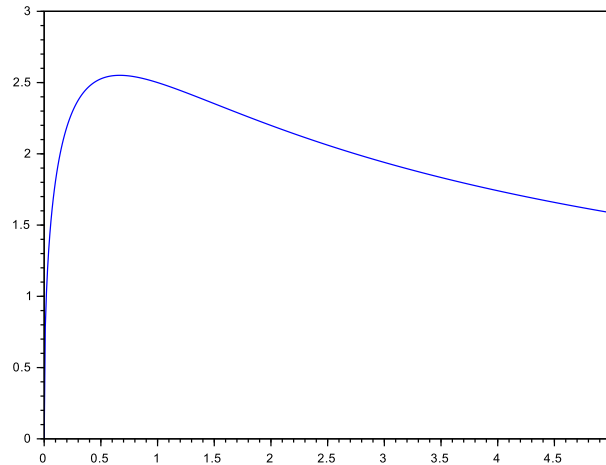
Or

$$c\mu + (c - 1)\alpha \geq 0 \Leftrightarrow c\mu \geq -(c - 1)\alpha = (1 - c)\alpha \Leftrightarrow \alpha \leq \frac{c\mu}{1 - c},$$

car $1 - c > 0$. Donc contrairement au cas sur-linéaire, R' change de signe sur D_R au point $\frac{c\mu}{1-c}$. De plus comme $c < 1$, on a $\lim_{\alpha \rightarrow +\infty} R(\alpha) = 0$, d'où le tableau de variation suivant :

α	0	$\frac{c\mu}{1-c}$	$+\infty$
$R'(\alpha)$	+	0	-
$R(\alpha)$	0	$M = \max R$	
		\nearrow	\searrow
			0

La fonction R admet donc ici un maximum M en $\frac{c\mu}{1-c}$.



Les pathogènes les plus favorisés par l'évolution ne sont pas ceux d'une virulence trop grande, mais plutôt ceux dont la virulence est très proche du maximum, qui laissent du temps à leurs hôtes pour infecter le plus grand nombre d'individus sains. C'est le cas de la majeure partie des maladies virales connues (grippes, gastro-entérites, méningites, varicelle, rougeole, etc.).

Chapitre 2

Probabilités

« La théorie des probabilités en tant que discipline mathématique peut et doit être développée à partir d'axiomes de la même manière qu'en géométrie et en algèbre. »

A. N. Kolmogorov (1903-1987).

Le calcul des probabilités est la science qui modélise les expériences dont l'issue n'est pas prévisible *a priori*. Le jet d'un dé, le tirage du Loto sont des exemples classiques de ces expériences, dites aléatoires. Une modélisation implique une simplification des phénomènes observés dans ces expériences, mais cette simplification conduit à une quantification, donc à la possibilité de faire des calculs et de prédire. La modélisation du calcul des probabilités a été inventée par A. N. Kolmogorov dans un livre (intitulé « Les fondements de la théorie des probabilités ») paru en 1933. Dans ce chapitre nous allons donner les définitions fondamentales des probabilités, ainsi que les bases du calcul des probabilités. L'accent sera mis sur le cas discret fini. Pour l'instant aucune connaissance particulière n'est exigée sur la notion de série.

En biologie on peut citer : sexe d'un nouveau né, couleur des cheveux, ou risque qu'un fœtus développe le syndrome de Down (trisomie 21). Dans ce dernier exemple, il existe des tests prénataux permettant de détecter si le fœtus est atteint par la maladie. Ces tests sont cependant effectués sur la base de prélèvements invasifs (analyse de cellules du trophoblaste, ponction de liquide amniotique) et peuvent, en cas d'utilisation massive, provoquer davantage d'avortements qu'ils ne détectent de victimes du syndrome. Il est donc essentiel d'être en mesure d'évaluer les risques de développement de la trisomie 21, la stratégie médicale consistant alors à n'effectuer un test plus poussé que pour les grossesses à risque et non pas systématiquement.

Il est facile de connaître le nombre moyen d'individus atteints par la syndrome de Down : en moyenne une naissance sur 700 si l'on choisit de ne pas tenir compte de l'âge de la future mère (et de celui du père). La méthode la plus courante pour évaluer si une grossesse est à risque consiste alors à effectuer des dosages de certaines substances comme par exemple l'hormone HCG ou l'alpha-fœtoprotéine, via de simples prises de sang. En effet la présence de

la trisomie peut parfois se traduire par des taux anormalement élevés de ces substances dans le sang. Néanmoins ce mécanisme n'est pas systématique et bien d'autres facteurs incontrôlables (liés au métabolisme de la mère ou du fœtus) peuvent aussi engendrer des taux s'écartant fortement de la moyenne attendue. Des études médicales ont montré qu'une trisomie sur quatre engendre des taux anormaux tandis qu'un taux hors norme se rencontre dans une grossesse exempte du syndrome sur cent. Dans de telles circonstances, quelles conclusions faut-il alors tirer d'un test sanguin indiquant un dosage anormalement élevé ?

2.1 Espace de probabilité

2.1.1 L'espace des observables Ω

Nous conviendrons qu'effectuer une expérience aléatoire, c'est sélectionner par un procédé quelconque un élément ω dans un ensemble Ω . Par exemple, jeter un dé revient à sélectionner un élément de $\Omega = \{1, 2, 3, 4, 5, 6\}$. Cet ensemble Ω est appelé l' **univers** (ou **espace des observables** ou **espace des issues** ou espace des événements élémentaires). Ses points ω sont appelés **possibles**, **aléas**, **éventualités** ou **observables** ou **événements élémentaires**.

Il est très important qu'il soit clairement défini afin de regrouper au moins tous les résultats possibles de l'expérience aléatoire.

Exemple 2.1 On considère un dé à 6 faces, numérotées de 1 à 6. L'ensemble Ω est l'ensemble des résultats possibles d'un lancer, soit $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Exemple 2.2 On considère le même dé, sauf que sur la sixième face le nombre 6 est remplacé par le 5. Il y a donc deux faces où 5 est inscrit. L'ensemble des résultats est différent : $\tilde{\Omega} = \{1, 2, 3, 4, 5\}$.

Attention : il n'est pas toujours possible de définir Ω de façon rigoureuse. Par exemple quel univers attacher au problème aléatoire du « temps qu'il fait » ?

2.1.2 Les événements

Le résultat d'une expérience aléatoire est par définition imprévisible. Toutefois, une fois fixé A un sous-ensemble de l'espace des observables Ω , on s'intéresse à la possibilité qu'a le résultat ω de l'expérience de tomber dans A . Les parties de Ω pour lesquelles on se pose ce genre de question sont appelées des **événements**. Un événement est donc un **sous-ensemble** ou une **partie** de Ω . On note $\mathcal{P}(\Omega)$ l'**ensemble des parties** de Ω et on supposera que toute partie de Ω est un événement ¹.

1. Un des premiers points délicats de la théorie est qu'on ne doit pas considérer tous les sous ensembles de Ω comme des événements, si l'univers n'est pas dénombrable. Il faut

Considérons l'expérience aléatoire consistant à lancer deux dés (un rouge et un bleu). L'ensemble des résultats possibles est $\Omega = \{(1, 1), (2, 1), (3, 6), \dots\}$ à 36 éléments. L'événement « la somme des valeurs des deux dés est supérieure ou égale à 10 » est l'ensemble des résultats suivants :

$$\begin{aligned} A &= \{(4, 6), (5, 5), (5, 6), (6, 4), (6, 5), (6, 6)\} \\ &= \{(i, j), i + j \geq 10, 1 \leq i \leq 6, 1 \leq j \leq 6\}. \end{aligned}$$

Dire que A s'est réalisé signifie simplement que le résultat ω que l'on observe appartient effectivement à A .

Le langage de la théorie des ensembles permet des calculs systématiques sur les événements. Toutefois, il faut savoir que le langage courant, que nous utilisons dans une première étape pour décrire des événements a sa traduction ensembliste.

Définition 2.1 (Opérations sur les ensembles)

- Ensemble Ω : *événement certain*,
- Ensemble vide \emptyset : *événement impossible*,
- **Réunion** de A et B , $A \cup B$: A ou B sont réalisés (« ou » non exclusif),
- **Intersections** de A et B , $A \cap B$: A et B sont réalisés,
- A et B sont **disjoints**, i.e. $A \cap B = \emptyset$: les événements A et B sont **incompatibles**,
- $\bar{A} = A^c = \Omega \setminus A$: événement **contraire** de A ou **complémentaire** de A .

Ainsi un possible ω est dans $A \cup B$ si ω est dans A ou dans B (ou les deux!). L'aléa ω est dans $A \cap B$ si ω est à la fois dans A et dans B .

2.1.3 La probabilité \mathbb{P}

Pour l'instant l'univers Ω , espace probabilisable, est une image incomplète de l'expérience, car il manque les probabilités. Reprenons l'exemple du lancer d'un dé équilibré (non truqué). Toutes les faces ont donc la même probabilité d'apparition : on dit qu'elles sont **équiprobables**. Il y a **équiprobabilité** parce qu'on a admis qu'il n'y avait aucune raison de tomber sur une face plutôt qu'une autre. Cette hypothèse fréquente en pratique est due à notre ignorance : « il n'y a pas de raison de privilégier l'un plutôt que l'autre ». On peut l'imposer a priori pour aborder un calcul direct, puis après informations supplémentaires, faire une correction a posteriori (voir section 2.4.2). L'équiprobabilité dans un univers *fini* permet d'obtenir l'expression bien connue de la probabilité d'un événement (cf. section 2.2.1).

L'équiprobabilité n'est pas la règle générale. Considérons le cas du lancer de deux dés (un rouge et un bleu ou lancés l'un après l'autre). On a

parler de **tribu**.

vu que Ω était composé de 36 éléments et si les deux dés sont équilibrés et qu'il n'y a pas de tricherie, tous les résultats sont équiprobables. Supposons maintenant que les deux dés sont identiques et lancés simultanément et qu'on obtienne le résultat (3,5). Il est impossible de savoir quel dé a donné 3. Par rapport à la situation précédente, il faut regrouper tous les résultats du type (i, j) avec $i \neq j$. Ainsi en choisissant de mettre d'abord le nombre le plus petit, Ω devient

$$\Omega = \{(1, 1), (1, 2), (1, 3), \dots, (2, 2), (2, 3), \dots\} = \{(i, j), 1 \leq i \leq j \leq 6\},$$

et a 21 éléments. Ces éléments ne sont pas équiprobables : (1,2) a une chance sur 18 de sortir, alors que (1,1) n'en a qu'une sur 36.

On quantifie désormais la possibilité qu'a le résultat ω de l'expérience de tomber dans un événement $A \in \mathcal{P}(\Omega)$. La puissance du calcul des probabilités réside dans le fait qu'il ne manipule pas directement le résultat ω de l'expérience aléatoire mais mesure les chances d'arrivée d'un événement A du au hasard. La définition qui suit est la définition mathématiquement rigoureuse d'une probabilité.

Définition 2.2 *Étant donné un espace d'observables Ω , une probabilité \mathbb{P} est une application de $\mathcal{P}(\Omega)$ dans $[0, 1]$, donc une fonction qui associe à tout événement A un nombre $\mathbb{P}(A)$ compris entre 0 et 1 appelé **probabilité** de A , et qui satisfait aux axiomes suivants² :*

1. L'événement certain est de probabilité 1 : $\mathbb{P}(\Omega) = 1$.
2. Pour deux événements A et B disjoints, $A \cap B = \emptyset$,

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

Le couple (Ω, \mathbb{P}) est alors appelé un **espace de probabilité** ou **espace probabilisé**.

Avant de voir les propriétés conséquences de ces axiomes, faisons quelques remarques sur cette définition.

1. Une probabilité est une mesure ou un poids d'un événement. Elle se doit d'être **positive**.
2. La condition $\mathbb{P}(\Omega) = 1$ est une **normalisation**. La probabilité que n'importe quel aléa ω se réalise est 1.
3. La seconde condition est une **règle d'addition** pour les probabilités. Nous l'avons utilisé plusieurs fois en ajoutant des probabilités d'événements élémentaires pour obtenir celle d'un événement à plusieurs aléas.

2. En toute rigueur, la seconde assertion devrait être : pour toute suite $A_1, A_2, \dots, A_n, \dots$ d'événements qui sont deux à deux disjoints, c'est-à-dire tels que $A_k \cap A_j = \emptyset$ si $k \neq j$, alors la série $\sum_{k=1}^{\infty} \mathbb{P}(A_k)$ converge et a pour somme $\mathbb{P}\left(\bigcup_{k \geq 1} A_k\right)$.

Voici quelques conséquences immédiates des axiomes.

Propriétés 2.1 Soit (Ω, \mathbb{P}) un espace de probabilité. Alors

1. $\mathbb{P}(\emptyset) = 0.$
2. Si A_1, A_2, \dots, A_n dans $\mathcal{P}(\Omega)$ sont deux à deux disjoints, alors

$$\mathbb{P}(A_1 \cup \dots \cup A_n) = \mathbb{P}(A_1) + \dots + \mathbb{P}(A_n);$$

en particulier $\mathbb{P}(A^c) = 1 - \mathbb{P}(A).$

3. Si A et B sont dans $\mathcal{P}(\Omega)$ et si $A \subset B$ alors $\mathbb{P}(A) \leq \mathbb{P}(B).$
4. Si A et B sont dans $\mathcal{P}(\Omega)$, mais ne sont pas nécessairement disjoints, alors

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

Si les A_1, A_2, \dots, A_n dans $\mathcal{P}(\Omega)$ ne sont pas nécessairement deux à deux disjoints, alors $\mathbb{P}(A_1 \cup \dots \cup A_n) \leq \mathbb{P}(A_1) + \dots + \mathbb{P}(A_n).$

5. Soit (A_1, \dots, A_n) une partition finie de Ω (i.e. les $A_k \in \mathcal{P}(\Omega)$ sont deux à deux disjoints, et leur union est Ω). Alors pour tout $B \in \mathcal{P}(\Omega)$,

$$\mathbb{P}(B) = \mathbb{P}(B \cap A_1) + \dots + \mathbb{P}(B \cap A_n).$$

La dernière propriété est souvent utilisée comme suit. Soit A un événement et A^c son complémentaire. Alors $A \cap A^c = \emptyset$ et $A \cup A^c = \Omega$: c'est une partition de Ω . Donc pour tout événement B :

$$\mathbb{P}(B) = \mathbb{P}(B \cap A) + \mathbb{P}(B \cap A^c).$$

2.2 Probabilités discrètes finies, équiprobabilité

On considère un espace fini $\Omega = \{\omega_n, 1 \leq n \leq N\}$ à N éléments.

Soit $(p_n)_{1 \leq n \leq N}$ N nombres réels tels que

1. $p_n \geq 0,$ pour tout $n,$
2. $p_1 + \dots + p_N = \sum_{n=1}^N p_n = 1.$

Ces deux conditions impliquent en particulier que $0 \leq p_n \leq 1$ pour tout entier $n.$

L'application qui à tout événement A fait correspondre la valeur

$$\mathbb{P}(A) = \sum_{\substack{1 \leq n \leq N \\ \omega_n \in A}} p_n$$

est une probabilité sur Ω . En effet

$$\mathbb{P}(\Omega) = \sum_{\substack{1 \leq n \leq N \\ \omega_n \in \Omega}} p_n = \sum_{n=1}^N p_n = 1.$$

De plus si A et B sont disjoints, si $\omega_n \in A$, alors $\omega_n \notin B$ et vice-versa. Ainsi si ω_n est dans $A \cup B$, soit il est dans A , soit il est dans B , et :

$$\mathbb{P}(A \cup B) = \sum_{\substack{1 \leq n \leq N \\ \omega_n \in A \cup B}} p_n = \sum_{\substack{1 \leq n \leq N \\ \omega_n \in A}} p_n + \sum_{\substack{1 \leq n \leq N \\ \omega_n \in B}} p_n = \mathbb{P}(A) + \mathbb{P}(B).$$

Réciproquement si \mathbb{P} est une probabilité sur Ω , on calcule la probabilité d'un événement élémentaire, et on pose

$$p_n = \mathbb{P}(\{\omega_n\}) = \mathbb{P}(\omega_n).$$

Comme \mathbb{P} est une probabilité, pour tout entier n , $p_n \geq 0$ et

$$\sum_{n=1}^N p_n = \sum_{n=1}^N \mathbb{P}(\{\omega_n\}) = \mathbb{P}\left(\bigcup_{1 \leq n \leq N} \{\omega_n\}\right) = \mathbb{P}(\Omega) = 1.$$

Ce type de probabilités est appelé **probabilité discrète finie** (portée par les éléments ω_n pondérés par les poids p_n). Ainsi, si Ω est fini, toute probabilité est discrète. Elles sont représentées souvent sous forme de tableau.

Exemple 2.3 Toujours pour le dé à six faces, $\Omega = \{1, 2, \dots, 6\}$. Toute probabilité sur Ω se représente par six nombres compris entre 0 et 1, p_1, p_2, \dots, p_6 , tels que $\sum_{i=1}^6 p_i = 1$. Si le dé est équilibré alors pour tout i , $p_i = 1/6$.

2.2.1 Équiprobabilité sur les espaces finis

C'est un cas particulier où tous les poids p_k sont égaux à 1 sur le nombre d'éléments de Ω .

Définition 2.3 (Équiprobabilité) Soient un entier $N \geq 1$ et $\Omega = \{\omega_1, \dots, \omega_N\}$ un ensemble fini. L'équiprobabilité sur Ω est défini comme la probabilité :

$$\mathbb{P}(A) = \frac{1}{N} \sum_{n=1}^N \mathbf{1}_A(\omega_n) = \frac{\text{card } A}{\text{card } \Omega}.$$

La fonction $\mathbf{1}_A$ vaut 1 si $\omega \in A$, 0 sinon. Cette probabilité ne privilégie aucun élément particulier de Ω ; tous les p_n sont égaux. C'est un **cas particulier** de probabilité discrète, et ce n'est pas un cas universel. Dans ce cas, le calcul des probabilités se ramène au calcul de cardinaux d'ensembles, donc au *dénombrement*. **On suppose connus** les factoriels, les arrangements, les coefficients binomiaux et les calculs simples sur le dénombrement (voir section suivante).

Exemple 2.4 On jette trois fois une pièce de monnaie parfaite. On peut représenter l'espace Ω comme l'ensemble des applications de $\{1, 2, 3\}$ (trois jets) dans $\{P, F\}$ ($P = \text{pile}$, $F = \text{face}$). Autrement dit, $\Omega = \{P, F\}^3$. Donc $N = \text{card } \Omega = 2^3 = 8$. Il est alors facile de voir par exemple que

$$\mathbb{P}(\text{on sort exactement une fois P}) = \frac{3}{8},$$

$$\mathbb{P}(\text{on sort au moins une fois P}) = 1 - \mathbb{P}(\text{on sort trois fois F}) = 1 - \frac{1}{8} = \frac{7}{8}.$$

Exemple 2.5 Le jeu du loto consiste à choisir 6 numéros distincts parmi $\{1, 2, \dots, 49\}$. On suppose que les boules qui portent les 49 numéros sont toutes parfaites. On ne s'intéresse qu'aux résultats des 6 boules. L'espace Ω est

$$\Omega = \{(a_1, a_2, \dots, a_6), 1 \leq a_i \leq 49, \text{ les } a_i \text{ sont deux à deux différents}\}.$$

Par exemple $(1, 2, \dots, 6) = (6, 5, \dots, 1)$. On a $N = \text{card } \Omega = C_{49}^6 = \binom{49}{6}$.

Par conséquent,

$$\mathbb{P}(\text{on gagne le premier prix avec un bulletin}) = \frac{1}{C_{49}^6} \approx 7 \times 10^{-8}.$$

2.3 Dénombrement : quelques rappels

Ce court paragraphe est là pour rappeler quelques principes de base du dénombrement. Il doit être compris mais n'est pas traité en cours. On rappelle que pour tout entier $n \in \mathbb{N}$, $0! = 1$ et pour $n \geq 1$

$$n! = 1 \times 2 \times 3 \times \dots \times n.$$

2.3.1 Arrangements et permutations

Arrangements avec répétition

On effectue p tirages *non exhaustifs* dans une urne contenant n boules *distinctes*. Combien y a-t-il de résultats différents en tenant compte de l'ordre des tirages ? Au premier tirage il y a n possibilités, au second n , etc. jusqu'au p -ième. Le nombre de résultats possibles est donc n^p . Ici les boules sont discernables ainsi que les tirages. À chaque tirage est associée une boule et une seule ; il s'agit donc d'une application de l'ensemble des p tirages dans l'ensemble des n boules. Ces sont des **arrangements avec répétition**, leur nombre est $a_n^p = n^p$.

À titre d'exemple, si $n = 18$ chevaux participent à $p = 3$ courses dans une même journée et qu'un parieur s'intéresse uniquement au cheval qui arrive premier, il a $18^3 = 5832$ possibilités pour répartir sa mise.

Arrangements sans répétition

On effectue p tirages *exhaustifs* dans une urne contenant n boules *distinctes*. Le nombre de résultats possibles en tenant compte de l'ordre se calcule de la même façon : au premier tirage on a n choix, au second $n - 1$, etc. jusqu'au p -ième où il reste $n - p + 1$ boules. On obtient donc : $n(n - 1) \dots (n - p + 1)$. Boules et tirages sont discernables, mais à chaque tirage est associée une boule et une seule. C'est une **injection** de l'ensemble des p tirages dans celui des n boules. Une boule peut ne pas être tirée, et p ne peut pas être supérieur à n . On parle de tirages *exhaustifs et ordonnés*. Ce sont des **arrangements sans répétition**, leur nombre A_n^p est :

$$A_n^p = n(n - 1) \dots (n - p + 1) = \frac{n!}{(n - p)!}.$$

On rappelle que $n! = 1 \times 2 \times 3 \times \dots \times n$ avec par convention $0! = 1$.

S'il y a $n = 18$ chevaux au départ d'une course alors le nombre de tiercés possibles ($p = 3$) dans l'ordre est égal à

$$A_{18}^3 = \frac{18!}{15!} = 18 \times 17 \times 16 = 4896.$$

Permutations

C'est un arrangement sans répétition avec $p = n$. Ceci correspond aussi au nombre de numérotations des n éléments d'un ensemble, ou encore au nombre de distributions de n objets distincts dans n cases distinctes. Ce sont les **permutations** de n éléments ; il y en a :

$$P_n = n!$$

Toujours avec $n = 18$ chevaux, il y a exactement $18! = 6402373705728000$ arrivées possibles à l'issue de la course.

2.3.2 Indiscernabilité des objets : combinaisons

Combinaisons sans répétition

Soit p tirages exhaustifs dans une urne contenant n objets *distincts* ; on ne s'intéresse maintenant qu'à la nature des objets tirés et non à l'ordre des tirages³. : le nombre donné par A_n^p est donc trop important. Les tirages constitués des mêmes objets sont regroupés en classes, chacune contenant $p!$ permutations d'objets. Une classe est une **combinaison** de p éléments parmi n ; elle est ici sans répétition car les objets n'apparaissent jamais plus d'une fois. Leur nombre est :

$$C_n^p = \binom{n}{p} = \frac{n!}{p!(n - p)!}$$

qu'on appelle **coefficient binomial**. On peut attacher à ce cas :

3. le tirage peut se faire par poignées

- la distinction de p objets (atomes, particules) *indiscernables* dans n cases (sites, états) *distinctes, sans répétition* ;
- le nombre de façons de partitionner un ensemble à n éléments (distincts) en deux sous-ensembles à p et $n - p$ éléments.

Si on a encore $n = 18$ chevaux au départ d'une course et qu'on parie sur le quinté sans ordre il y a

$$C_{18}^5 = \frac{18!}{5! \times 13!} = 8568$$

combinaisons possibles.

Coefficient multinomial ou polynomial

On généralise au nombre de partitions d'un ensemble à n éléments distincts en r sous-ensembles à n_1, \dots, n_r éléments avec $n_1 + n_2 + \dots + n_r = n$. On obtient alors le **coefficient multinomial** ou **polynomial** :

$$C_n^{n_1, \dots, n_r} = \frac{n!}{n_1! n_2! \dots n_r!}.$$

Ce coefficient représente aussi le nombre de façons de tirer n_1 objets de type 1, n_2 objets de type 2, ... et n_r objets de type r lors de n tirages exhaustifs dans une urne contenant des objets de ces types.

Relations entre combinaisons

On peut prouver les relations suivantes :

$$\begin{aligned} C_n^0 &= C_n^n = 1, \\ C_n^1 &= C_n^{n-1} = n \\ C_n^p &= C_n^{n-p} \\ pC_n^p &= nC_{n-1}^{p-1} \\ C_n^p &= C_{n-1}^{p-1} + C_{n-1}^p. \end{aligned}$$

La dernière relation est dite du triangle de Pascal. On peut ajouter la formule dite de Vandermonde

$$\sum_{k=0}^n C_n^k C_m^{p-k} = C_{n+m}^p,$$

et celle dite du binôme de Newton :

$$(a + b)^n = \sum_{p=0}^n C_n^p a^p b^{n-p}.$$

2.4 Indépendance et probabilités conditionnelles

2.4.1 Indépendance

Définition 2.4 (Indépendance de deux événements) Soient A et B dans $\mathcal{P}(\Omega)$. On dit qu'ils sont *indépendants* si

$$(2.1) \quad \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

Même si l'indépendance entre deux événements est souvent intuitivement évidente, il faut toujours essayer de la démontrer rigoureusement à partir de la définition. Et il n'y a pas de raison à donner : il suffit de vérifier la relation (2.1). Il faut souligner le rôle essentiel de \mathbb{P} . Si on change de probabilité, il n'y a pas de raison pour que des événements, qui étaient indépendants, le restent sous la nouvelle probabilité. Contrairement à l'incompatibilité, l'indépendance n'est pas une propriété intrinsèque des événements.

Considérons l'exemple suivant. À chaque naissance, la probabilité d'avoir une fille est $1/2$ et d'avoir un garçon $1/2$ aussi. On souhaite étudier dans l'ensemble des familles à deux enfants, l'indépendance éventuelle entre A : « la famille a un enfant de chaque sexe » et B : « il y a au plus une fille ». Ici en tenant compte de l'ordre des naissances, l'univers est

$$\Omega = \{(f, f), (f, g), (g, f), (g, g)\}.$$

Avec l'hypothèse d'équiprobabilité, on a

$$\mathbb{P}(A) = \mathbb{P}(\{(f, g), (g, f)\}) = 1/2,$$

et

$$\mathbb{P}(B) = \mathbb{P}(\{(f, g), (g, f), (g, g)\}) = 3/4,$$

tandis que

$$\mathbb{P}(A \cap B) = \mathbb{P}(\{(f, g), (g, f)\}) = 1/2.$$

Donc $\mathbb{P}(A \cap B) \neq \mathbb{P}(A)\mathbb{P}(B)$: les deux événements ne sont pas indépendants.

Si on regarde le même problème sur des familles à trois enfants, alors

$$\Omega = \{(f, f, f), (f, f, g), (f, g, f), (f, g, g), (g, f, f), (g, f, g), (g, g, f), (g, g, g)\}$$

de cardinal $2^3 = 8$, et

$$\mathbb{P}(A) = \mathbb{P}(\{(f, f, g), (f, g, f), (f, g, g), (g, f, f), (g, f, g), (g, g, f)\}) = 6/8 = 3/4,$$

$$\mathbb{P}(B) = \mathbb{P}(\{(f, g, g), (g, f, g), (g, g, f), (g, g, g)\}) = 4/8 = 1/2,$$

$$\mathbb{P}(A \cap B) = \mathbb{P}(\{(f, g, g), (g, f, g), (g, g, f)\}) = 3/8.$$

Donc $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$: les deux événements sont indépendants.

L'indépendance n'est pas transitive : si A est indépendant de B et B de C , A n'est pas forcément indépendant de C .

Propriétés 2.2 Si A et B sont indépendants, alors A^c et B sont indépendants, de même que A et B^c ou que A^c et B^c .

Plus généralement, si (A_1, \dots, A_n) sont n événements :

Définition 2.5 (A_1, \dots, A_n) sont (mutuellement) indépendants si pour tout ensemble d'indices $I \subset \{1, 2, \dots, n\}$:

$$\mathbb{P}\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} \mathbb{P}(A_i).$$

Par exemple pour trois événements $\{A, B, C\}$ il faut vérifier que

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C), \quad \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B),$$

$$\mathbb{P}(A \cap C) = \mathbb{P}(A)\mathbb{P}(C), \quad \text{et} \quad \mathbb{P}(B \cap C) = \mathbb{P}(B)\mathbb{P}(C).$$

Par exemple si $\Omega = \{a, b, c, d\}$ muni de la probabilité uniforme (c'est-à-dire $\mathbb{P}(a) = \mathbb{P}(b) = \mathbb{P}(c) = \mathbb{P}(d) = 1/4$), et si

$$A_1 = \{a, b\}, \quad A_2 = \{a, c\}, \quad A_3 = \{a, d\},$$

alors

$$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1 \cap A_3) = \mathbb{P}(A_2 \cap A_3) = 1/4,$$

et ainsi les trois événements sont deux à deux indépendants. Mais

$$\mathbb{P}(A_1 \cap A_2 \cap A_3) = 1/4 \neq \mathbb{P}(A_1)\mathbb{P}(A_2)\mathbb{P}(A_3) = 1/8,$$

ce qui montre que les trois événements ne sont pas (mutuellement) indépendants.

2.4.2 Probabilités conditionnelles

On s'intéresse au lancer successif de deux dés. On considère les deux événements suivants : A : « la somme des deux dés est égale à 11 » et B : « le premier dé donne 6 ». On vérifie que la probabilité de A est $1/18$.

Mais si on observe le premier dé et que l'événement B est réalisé, en quoi cela affecte la probabilité d'avoir une somme égale à 11 ? Autrement dit, sachant que B est réalisé, quelle est la probabilité de A ? Pour obtenir 11, le deuxième dé doit donner 5. Et ainsi la probabilité d'avoir A , sachant B , est de $1/6$.

Cet exemple motive la définition suivante.

Définition 2.6 Soient deux événements A et B sur (Ω, \mathbb{P}) , avec $\mathbb{P}(B) > 0$. La **probabilité conditionnelle de A sachant B** est

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

Si B ne se réalise jamais, i.e. $\mathbb{P}(B) = 0$, cette définition n'a aucun sens.

Théorème 2.1 *Si $\mathbb{P}(B) > 0$, alors la fonction $\mathbb{P}(\cdot|B) : \mathcal{P}(\Omega) \rightarrow [0, 1]$ est une probabilité sur Ω .*

Preuve. D'abord on va vérifier que cette fonction est correctement définie. Soit $A \in \mathcal{P}(\Omega)$. Comme $A \cap B \subset B$, on a $\mathbb{P}(A \cap B) \leq \mathbb{P}(B)$. Donc par définition,

$$0 \leq \mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \leq 1.$$

Vérifions maintenant les deux axiomes d'une probabilité.

1. Si $A = \Omega$, $\Omega \cap B = B$. Donc $\mathbb{P}(\Omega|B) = 1$.
2. Soit A_1 et A_2 deux événements de $\mathcal{P}(\Omega)$ avec $A_1 \cap A_2 = \emptyset$. On veut calculer $\mathbb{P}\left(A_1 \cup A_2 \middle| B\right)$. On a

$$\begin{aligned} \mathbb{P}\left(A_1 \cup A_2 \middle| B\right) &= \frac{1}{\mathbb{P}(B)} \mathbb{P}[(A_1 \cup A_2) \cap B] \\ &= \frac{1}{\mathbb{P}(B)} \mathbb{P}[(A_1 \cap B) \cup (A_2 \cap B)] \\ &= \frac{1}{\mathbb{P}(B)} [\mathbb{P}(A_1 \cap B) + \mathbb{P}(A_2 \cap B)] \\ &= \frac{\mathbb{P}(A_1 \cap B)}{\mathbb{P}(B)} + \frac{\mathbb{P}(A_2 \cap B)}{\mathbb{P}(B)} \\ &= \mathbb{P}(A_1|B) + \mathbb{P}(A_2|B). \end{aligned}$$

Le passage de la première ligne à la deuxième est justifié par le fait que $(A_1 \cap B) \cap (A_2 \cap B) = (A_1 \cap A_2) \cap B = \emptyset$, c'est-à-dire $(A_1 \cap B)$ et $(A_2 \cap B)$ sont deux à deux disjoints.

□

On déduit de la définition et du théorème précédent quelques règles de calcul sur les probabilités conditionnelles.

Propriétés 2.3

1. Si $A \cap B = \emptyset$, $\mathbb{P}(A \cup B|C) = \mathbb{P}(A|C) + \mathbb{P}(B|C)$.
2. $\mathbb{P}(A^c|B) = 1 - \mathbb{P}(A|B)$.
3. $\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B) = \mathbb{P}(B|A)\mathbb{P}(A)$.

Proposition 2.1 *Soit deux événements A et B sur (Ω, \mathbb{P}) , avec $\mathbb{P}(B) > 0$. A et B sont indépendants si et seulement si*

$$\mathbb{P}(A|B) = \mathbb{P}(A),$$

ou encore avec $\mathbb{P}(A) > 0$, $\mathbb{P}(B|A) = \mathbb{P}(B)$.

Dire que deux événements sont indépendants, c'est dire que la connaissance de l'un n'influence pas l'autre.

Exemple 2.6 On modélise le lancer de deux pièces semblables indépendantes par

$$\Omega = \{\omega = (\omega_1, \omega_2); \omega_i \in \{P, F\}\} \text{ et } \mathbb{P}(\omega) = \mathbb{P}_1(\omega_1)\mathbb{P}_1(\omega_2),$$

avec $\mathbb{P}_1(P) = 1 - \mathbb{P}_1(F) = p$. On s'imagine qu'on a lancé la première pièce et que $\omega_1 = P$. On associe à chaque issue du lancer de la deuxième pièce une probabilité (sachant que $\omega_1 = P$) :

$$\mathbb{P}(\omega_2 = P | \omega_1 = P) = \frac{\mathbb{P}(\{P, P\})}{\mathbb{P}(\{\omega; \omega_1 = P\})} = \frac{pp}{p(1-p) + pp} = p.$$

L'indépendance des deux pièces signifie que la connaissance du résultat du lancer de la première ne nous apprend rien sur le résultat du lancer de la seconde.

Exemple 2.7 On suppose qu'on a deux paires de pièces différentes : la 1ère paire consiste en deux pièces identiques et pile sort 99 fois sur 100 en moyenne ; la deuxième paire consiste en deux pièces identiques mais pile sort 1 fois sur 10 en moyenne. Un jeu consiste en : (i) choisir au hasard une d'entre les deux paires, (ii) lancer les deux pièces correspondantes. On modélise l'espace des issues d'un jeu par :

$$\Omega = \{\omega = (\omega_1, \omega_2, n); \omega_i \in \{P, F\}, n = 1, 2\},$$

et

$$\mathbb{P}(\omega_1, \omega_2, 1) = \frac{1}{2}\mathbb{P}_1(\omega_1)\mathbb{P}_1(\omega_2) \text{ et } \mathbb{P}_1(P) = \alpha = 0,99$$

et

$$\mathbb{P}(\omega_1, \omega_2, 2) = \frac{1}{2}\mathbb{P}_2(\omega_1)\mathbb{P}_2(\omega_2) \text{ et } \mathbb{P}_2(P) = \beta = 0,1.$$

On calcule ici

$$\begin{aligned} \mathbb{P}(\omega_2 = P | \omega_1 = P) &= \frac{\mathbb{P}(\{\omega; \omega_2 = P, \omega_1 = P\})}{\mathbb{P}(\{\omega; \omega_1 = P\})} \\ &= \frac{\mathbb{P}(P, P, 1) + \mathbb{P}(P, P, 2)}{\mathbb{P}(P, P, 1) + \mathbb{P}(P, F, 1) + \mathbb{P}(P, P, 2) + \mathbb{P}(P, F, 2)} \\ &= \frac{\alpha^2/2 + \beta^2/2}{\alpha/2 + \beta/2} = \frac{\alpha^2 + \beta^2}{\alpha + \beta} \end{aligned}$$

Par ailleurs si on ne savait rien sur la première pièce

$$\mathbb{P}(\omega_2 = P) = \mathbb{P}(\{P, P, 1\}, \{F, P, 1\}, \{P, P, 2\}, \{F, P, 2\}) = \frac{1}{2}(\alpha + \beta).$$

On note que $\mathbb{P}(\omega_2 = P | \omega_1 = P) > \mathbb{P}(\omega_2 = P)$ (resp. 0,91 et 0,54). Intuitivement, comme pile est sorti la 1ère fois, on a plus de chance que la paire choisie soit la paire 1.

Les probabilités conditionnelles sont à l'origine de deux formules.

Proposition 2.2 (Formule des probabilités totales)

Soit (A_1, A_2, \dots, A_n) une partition de Ω , telle que pour tout i , $\mathbb{P}(A_i) > 0$.
Pour tout $B \in \mathcal{P}(\Omega)$,

$$\mathbb{P}(B) = \mathbb{P}(B|A_1)\mathbb{P}(A_1) + \dots + \mathbb{P}(B|A_n)\mathbb{P}(A_n) = \sum_{i=1}^n \mathbb{P}(B|A_i)\mathbb{P}(A_i).$$

Cette première formule provient immédiatement de la formule :

$$\mathbb{P}(B) = \sum_{i=1}^n \mathbb{P}(B \cap A_i).$$

De la formule des probabilités totales, on déduit la formule de Bayes.

Proposition 2.3 (Formule de Bayes) De plus si $\mathbb{P}(B) > 0$, on a pour tout entier k :

$$\mathbb{P}(A_k|B) = \frac{\mathbb{P}(B|A_k)\mathbb{P}(A_k)}{\sum_{i=1}^n \mathbb{P}(B|A_i)\mathbb{P}(A_i)}.$$

Elle consiste à « inverser » les probabilités conditionnelles et est à la base des statistiques dites bayésiennes. Voyons sur un exemple comment elle s'utilise.

Exemple 2.8 Soit une urne contenant des boules blanches ou noires. Pour en savoir plus sur son contenu, on effectue 6 tirages successifs avec remise qui donnent tous une boule blanche. Que peut-on dire de la composition de l'urne ?

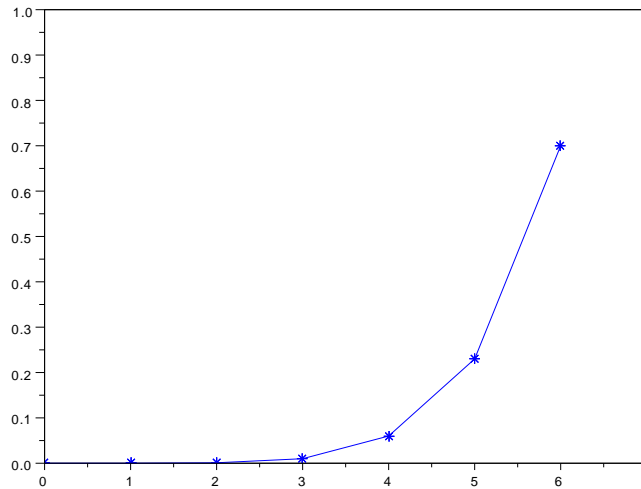
Si on sait que l'urne contient 6 boules au total, on a deux renseignements a priori (nombre et couleurs des boules) et une information a posteriori (résultat R : les 6 tirages ont tous donné une boule blanche). On peut imaginer les compositions possibles de l'urne (modélisation) : il y en a 7, depuis 0 boule blanche et 6 noires ($0B, 6N$) jusqu'à 6 boules blanches et 0 noire ($6B, 0N$), qui sont les états dits accessibles A_k . Il n'y a pas a priori de raison de privilégier l'un ou l'autre, d'où l'équiprobabilité $\mathbb{P}(A_k) = 1/7$. La probabilité de tirer une boule blanche dans une urne en contenant k est $k/6$; celle de tirer successivement 6 boules blanches est $(k/6)^6$: c'est $\mathbb{P}(R|A_k)$. Les probabilités a posteriori des compositions sont :

$$\mathbb{P}(A_k|R) = \frac{k^6}{\sum_{m=0}^6 m^6} = \frac{k^6}{67171}.$$

Ceci donne la table suivante :

Composition A_k	$0B, 6N$	$1B, 5N$	$2B, 4N$	$3B, 3N$	$4B, 2N$	$5B, 1N$	$6B, 0N$
k	0	1	2	3	4	5	6
$\mathbb{P}(A R)$	0	10^{-5}	10^{-3}	0,01	0,06	0,23	0,70

Et le graphe suivant :



Si maintenant le nombre total de boules M est inconnu, la composition A_k est

$$\{kB, (M - k)N\}.$$

Les probabilités $\mathbb{P}(A_k)$ sont inconnues mais l'équiprobabilité les rend inutile. La probabilité $\mathbb{P}(R|A_k)$ de tirer 6 boules blanches dans l'urne contenant k blanches sur M est $(k/M)^6$. Et la probabilité a posteriori cherchée est :

$$\mathbb{P}(A_k|R) = \frac{k^6}{f(M)}, \quad f(M) = \sum_{m=0}^M m^6.$$

$f(M)$ n'est pas connu, mais est indépendant de k . On a un renseignement relatif sur la composition de l'urne : les compositions contenant beaucoup de blanches sont favorisées (facteur k^6).

2.5 Évaluation d'un risque de trisomie 21 et tests médicaux

Traduisons ici les données recensées dans l'introduction. L'univers Ω est celui de toutes les grossesses⁴. On note T l'événement « le fœtus est atteint du syndrome de Down » et A « le taux repéré lors du test sanguin de la mère est anormal ». On sait que

$$\mathbb{P}(T) = \frac{1}{700}, \quad \mathbb{P}(A|T) = \frac{1}{4}, \quad \mathbb{P}(A|T^c) = \frac{1}{100}.$$

4. L'univers des possibles est ici assez flou. Comment donner un sens mathématique à l'ensemble des grossesses passées et à venir ? C'est plutôt l'ensemble des probabilités des divers événements d'intérêt (loi ou distribution) qui importe

On souhaite calculer lorsque le test révèle un taux anormal, le risque que le fœtus soit trisomique, soit $\mathbb{P}(T|A)$. Ce calcul se fait par la formule de Bayes :

$$\begin{aligned}\mathbb{P}(T|A) &= \frac{\mathbb{P}(A|T)\mathbb{P}(T)}{\mathbb{P}(A|T)\mathbb{P}(T) + \mathbb{P}(A|T^c)\mathbb{P}(T^c)} \\ &= \frac{(1/4) \times (1/700)}{(1/4) \times (1/700) + (1/100) \times (1 - 1/700)} \approx 3,5\%.\end{aligned}$$

En cas de taux anormal, le risque de présence réelle de la trisomie, n'est que de 3,5%. En effectuant des tests invasifs sur 1000 grossesses avec tests sanguins anormaux, on ne détectera (en moyenne) que 35 cas de trisomie. Cette conclusion mitigée est à mettre en rapport avec la stratégie consistant à ne pas présélectionner les grossesses et à effectuer les tests invasifs. Dans ce cas, sur 1000 patientes on ne trouvera qu'un à deux cas de trisomie 21, car $1000/700 \approx 1,4$.

En pratique plusieurs tests sont effectués en même temps pour affiner un peu plus la population à risque.

Ce problème du test médical est bien connu justement en médecine, mais s'applique à n'importe quel problème de test. De façon plus générale, on s'intéresse exclusivement à l'étude de la performance d'un test médical de dépistage T d'une maladie M . T peut être positif ($T+$), ou négatif ($T-$). Un sujet peut être atteint par la maladie ($M+$) ou sain ($M-$).

	$M+$	$M-$
$T+$	VP	FP
$T-$	FN	VN

avec VP= vrai positif, FP= faux positif, VN= vrai négatif, FN= faux négatif.

La sensibilité Se est la capacité d'un test à identifier correctement les sujets malades

$$Se = \frac{VP}{VP + FN}.$$

La spécificité Sp est la capacité d'un test à identifier correctement les sujets sains

$$Sp = \frac{VN}{VN + FP}.$$

Pour déterminer qui est malade et qui ne l'est pas, il a fallu utiliser une méthode indiscutable (point discutable en pratique) et indépendante du test T . On peut alors disposer de deux populations « indiscutables » : les sujets sains et les sujets malades.

L'ensemble des individus d'une population donnée n'aura pas un même résultat au test du fait de la variabilité biologique. Quand on compare deux populations, les malades $M+$ et les sains $M-$, il y a de grandes chances que les courbes décrivant les résultats au test T se « croisent ». Il va donc falloir déterminer un seuil qui permettra de déterminer des sujets $T-$ ou $T+$. C'est là que ça se complique, car déterminer ce seuil est une délicate question

d'équilibre entre la sensibilité et la spécificité. Il va falloir choisir de privilégier l'une ou l'autre. Plus le seuil est bas, plus on va diminuer le nombre de FN (on va donc augmenter la sensibilité), mais on va aussi augmenter le nombre de FP, et donc diminuer la spécificité. Ce choix dépend de ce que l'on veut faire du test, c'est-à-dire si on veut le rendre très spécifique ou au contraire très sensible. On peut aussi déterminer ce seuil à l'aide d'une courbe « ROC » (Receiver Operating Characteristics). Dans tous les cas ce choix d'un seuil de positivité d'un test est délicat, complexe et parfois (souvent) sujet à caution.

Dans la vie réelle, le patient, son médecin ou l'utilisateur du test se fichent un peu de savoir quelle est la sensibilité, la spécificité d'un test, ou la courbe ROC. Ce qu'ils veulent savoir c'est :

- Quelle est la probabilité d'être malade quand le test est positif, c'est à dire la valeur prédictive positive (VPP) ?
- Quelle est la probabilité d'être sain quand le test est négatif, c'est à dire la valeur prédictive négative (VPN) ?

C'est là qu'intervient le théorème de Bayes. Notons Pr la prévalence de la maladie dans la population, c'est-à-dire la proportion de la maladie dans une population définie, à un instant t :

Pr = nombre de cas observés à un instant t / population à risque au même instant t .

Alors

$$VPP = \frac{Se \times Pr}{Se \times Pr + (1 - Sp) \times (1 - Pr)}$$

$$VPN = \frac{Sp \times (1 - Pr)}{Sp \times (1 - Pr) + (1 - Se) \times Pr}$$

On remarque alors que la VPP et la VPN d'un test dépendent de la prévalence de la maladie.

- Plus la prévalence est élevée, plus la VPP est élevée, et plus la VPN est faible.
- Plus la prévalence est faible, plus la VPP est faible, et plus la VPN est élevée.

Autrement dit, si l'on effectue un test dans une population où la prévalence de la maladie est faible (dans une population à faible risque, pour exprimer les choses autrement) la probabilité qu'un patient avec un test positif soit effectivement malade sera faible, voire très faible. Au contraire, la probabilité qu'un patient soit sain en ayant un test négatif sera très élevée. Un même test aura donc une performance variable en fonction de la population étudiée.

Si on fait une série d'épreuves d'effort à la recherche d'une coronaropathie aux enfants d'une crèche, un test positif sera peu inquiétant, un test négatif très rassurant. Si on fait les mêmes épreuves d'effort à la recherche de la même chose dans une population d'hommes dépassant la cinquantaine, hypertendus, tabagiques et diabétiques non traités, un test positif sera très probablement révélateur d'une coronaropathie latente, alors qu'un test négatif sera finalement peu rassurant.

La performance d'un test dépend d'un paramètre qui lui est totalement étranger, la prévalence d'une maladie dans une population.

Prenons l'épreuve d'effort. Selon un article du *Journal of American Heart Association* de 1989, sa sensibilité est environ de 68%, sa spécificité environ de 77%. Selon un autre article du *European Heart Journal* de 2000, la prévalence d'une coronaropathie diagnostiquée à l'angiographie coronaire dans une population générale de 40 à 70 ans est de 7.4%. Cela donne une VPN à 96.79% et une VPP à 19.11%. Ce cas illustre bien le manque de pertinence d'une épreuve d'effort positive dans une population à relativement faible risque. Dans une population d'insuffisants rénaux à très haut risque, un article du *Journal of the American Society of Nephrology* de 2005 évalue la prévalence d'une coronaropathie à 53.3%. On obtient alors une VPN à 67.83% et une VPP à 77.14%.

Deuxième conséquence de la formule de Bayes, un peu plus intuitive : la valeur prédictive d'un test dépend de sa spécificité et de sa sensibilité.

— Plus la Sp est élevée, plus la VPP est élevée.

— Plus la Se est élevée, plus la VPN est élevée.

La pertinence d'un test est directement dépendante de ses caractéristiques propres.

Du point de vue médical, la non compréhension induit au quotidien des demandes statistiquement aberrantes d'examen complémentaires et aussi et surtout des interprétations tout aussi aberrantes que dangereuses pour le patient. Ce ne sont pas tellement les tests qui posent problème, mais plutôt la population dans laquelle ils sont effectués et l'interprétation qu'en fait le médecin demandeur.

Néanmoins dans d'autres domaines les choses sont moins claires. Dans une note parue dans le journal *Nature* (volume 454, issue 7205 du 7 août 2008), Donal Berry, du MD Anderson cancer center, un des plus fameux instituts de recherche et de traitement du cancer des Etats-Unis, remet en question la fiabilité des tests de dépistage du dopage. Son argument : pour des raisons de méthodologie on disculpe trop de tricheurs et on punit beaucoup d'innocents. Il avait auparavant beaucoup travaillé sur le dépistage du cancer du sein, montrant pourquoi, en termes de santé publique, le dépistage de masse ne se justifiait qu'à partir de cinquante ans. Dans cet article Berry dit que les tests de mesure des anabolisants et de la testostérone en particulier, n'ont pas été suffisamment évalués pour qu'on puisse être certain de la sensibilité et de la spécificité des mesures. Il reproche aux laboratoires de lutte contre le dopage un manque de rigueur et donc de faire « tomber » des innocents. De la même façon, il estime que l'existence de faux négatifs permet à de vrais tricheurs de passer à travers les mailles du filet. Et Berry peut enfoncer le clou car, en matière de recherche de dopage, les valeurs normales et les limites ne sont pas fixées comme pour un test de dépistage du cholestérol ou du VIH par exemple. Impossible en effet de mesurer des milliers d'échantillons. Souvent les études ont été faites après que quelques volontaires ont ingéré une substance interdite et qu'on a mesuré les taux sanguins ou urinaires de

*2.5. ÉVALUATION D'UN RISQUE DE TRISOMIE 21 ET TESTS MÉDICAUX*⁶¹

cette substance. On est donc parfois loin de la rigueur et des études exigées dans le domaine des tests diagnostiques par exemple.

Alphabet grec

Lettre minuscule	Lettre majuscule	Français	Valeur
α	A	alpha	a
β	B	beta	b
γ	Γ	gamma	g
δ	Δ	delta	d
ε	E	epsilon	é
ζ	Z	zéta	z
η	H	éta	e
θ	Θ	théta	th
ι	I	iota	i
κ	K	kappa	k
λ	Λ	lambda	l
μ	M	mu	m
ν	N	nu	n
ξ	Ξ	xi	x
o	O	omicron	o
π	Π	pi	p

Lettre minuscule	Lettre majuscule	Français	Valeur
ρ	P	rho	r,rh
σ	Σ	sigma	s
τ	T	tau	t
υ	Υ	upsilon	u
ϕ	Φ	phi	f
χ	X	khi	kh
ψ	Ψ	psi	ps
ω	Ω	omega	o