

IUT DE DIJON-AUXERRE
DUT GEA 1^{ÈRE} ANNÉE
ANNÉE UNIVERSITAIRE 2020-2021

Notes du cours de Mathématiques

M 12 05 : Mathématiques pour la gestion et statistiques, Semestre 1

M 22 06 : Probabilités, Semestre 2

M 22 07 : Mathématiques financières, Semestre 2

ARNAUD ROUSSELLE
arnaud.rouselle@iut-dijon.u-bourgogne.fr

Table des matières

1	Algèbre matricielle et systèmes linéaires	1
1.1	Premières définitions	1
1.2	Calcul matriciel	2
1.2.1	Somme et différence	2
1.2.2	Multiplication par un scalaire	3
1.2.3	Multiplication	4
1.2.4	Transposition	7
1.2.5	Déterminant	8
1.3	Systèmes d'équations linéaires	10
1.3.1	Écriture matricielle	11
1.3.2	Résolution	12
1.3.3	Application à l'inversion de matrice	18
1.4	Prolongement : introduction à la programmation linéaire	21
1.4.1	Méthode graphique	24
1.4.2	Forme standard d'un problème d'optimisation linéaire	25
1.4.3	Introduction à la méthode du simplexe	28
2	Statistiques descriptives univariées	41
2.1	Vocabulaire statistique	41
2.2	Cas qualitatif	42
2.3	Cas quantitatif discret sans regroupement en classes	44
2.3.1	Tableau statistique	45
2.3.2	Représentations graphiques	46
2.3.3	Paramètres statistiques	46
2.4	Cas quantitatif continu ou discret avec regroupement en classes	50
2.4.1	Tableau statistique	50
2.4.2	Représentations graphiques	52
2.4.3	Paramètres statistiques	55
2.5	Complément : d'autres moyennes	57
3	Statistiques bivariées	59
3.1	Présentation et traitement des données	59
3.2	Étude et mesure des liens entre les variables	62
3.2.1	Indépendance	62
3.2.2	Coefficient de Cramer	63
3.2.3	Coefficient de corrélation linéaire	65

3.3	Régression	67
3.4	Complément : test du χ^2 d'indépendance	70
4	Analyse à une variable réelle	75
4.1	Premières définitions	75
4.2	Fonctions usuelles	76
4.2.1	Fonctions linéaires et affines	76
4.2.2	Fonctions quadratiques	77
4.2.3	Fonctions polynomiales	79
4.2.4	Fonction inverse	79
4.2.5	Fonctions rationnelles	79
4.2.6	Fonctions exponentielles	80
4.2.7	Fonctions logarithmes	82
4.2.8	Fonctions puissances et racines	84
4.2.9	Fonctions indicatrices	85
4.2.10	Fonction valeur absolue	86
4.3	Limites	86
4.3.1	Définitions formelles	86
4.3.2	Limites usuelles	91
4.3.3	Opérations sur les limites	92
4.3.4	Comparaison de limites	93
4.3.5	Formes indéterminées	94
4.4	Continuité	96
4.5	Dérivabilité	97
4.5.1	Définitions	97
4.5.2	Dérivées usuelles	98
4.5.3	Opérations sur les dérivées	98
4.5.4	Tangente à une courbe en un point	99
4.5.5	Applications à l'étude des variations et à la recherche d'extrema	100
4.6	Applications en économie	102
4.6.1	Coût marginal	102
4.6.2	Coût moyen, coût marginal et optimum technique	103
4.6.3	Revenu marginal	103
4.6.4	Élasticité	104
4.7	Complément : autour du lien entre variations d'une fonction et signe de sa dérivée	106
5	Suites numériques	109
5.1	Premières définitions	109
5.2	Convergence et limites de suites	110
5.3	Suites particulières	112
5.3.1	Suites définies par récurrence	112
5.3.2	Suites arithmétiques	112
5.3.3	Suites géométriques	117
5.3.4	Suites arithmético-géométriques	121
5.4	Complément : la méthode de Newton	122

6	Mathématiques financières	125
6.1	Intérêts, capitalisation et actualisation	125
6.1.1	Intérêts simples	125
6.1.2	Intérêts composés	127
6.1.3	Taux nominaux, périodiques et effectifs	129
6.1.4	Capitalisation	130
6.1.5	Actualisation	130
6.2	Versements périodiques : les annuités	132
6.2.1	Annuités de début de période	132
6.2.2	Annuités de fin de période	134
6.2.3	Cas des annuités constantes	135
6.3	Emprunts <i>indivis</i>	137
6.3.1	Emprunts <i>in fine</i>	141
6.3.2	Emprunts à annuités constantes	141
6.3.3	Emprunts à amortissements constants	142
6.3.4	Taux annuel effectif global (TAEG)	143
6.4	Rentabilité d'un investissement	144
6.4.1	Valeur actuelle nette	144
6.4.2	Taux de rentabilité interne	145
7	Séries chronologiques	147
7.1	Premières définitions, exemples et motivations	147
7.1.1	Motivations et objectifs	148
7.2	Analyse d'une série chronologique	149
7.2.1	Décomposition d'une série chronologique	149
7.2.2	Modèle additif	150
7.2.3	Modèle multiplicatif	150
7.2.4	Quelques manipulations avec un tableur	151
7.3	Estimation des composantes d'une série chronologique	152
7.3.1	Estimation de la tendance générale	152
7.3.2	Estimation de la composante saisonnière	154
7.4	Prédiction	157
7.5	Quelques manipulations avec un tableur	159
8	Ensembles et dénombrement	161
8.1	Éléments de la théorie des ensembles	161
8.1.1	Ensembles, sous-ensembles	161
8.1.2	Opérations sur les ensembles	162
8.2	Combinatoire	167
8.2.1	p -listes	167
8.2.2	Permutations	168
8.2.3	Arrangements	169
8.2.4	Combinaisons	169

9	Calcul élémentaire des probabilités	173
9.1	Vocabulaire probabiliste	173
9.2	Équiprobabilité	176
9.3	Indépendance	177
9.4	Probabilités conditionnelles	178
10	Variables aléatoires	183
10.1	Premières définitions et exemples	183
10.2	Variables aléatoires discrètes	184
10.2.1	Loi d'une variable aléatoire discrète	184
10.2.2	Moments d'une variable aléatoire discrète	184
10.2.3	Lois discrètes usuelles	185
10.3	Variables aléatoires continues	191
10.3.1	Loi d'une variable aléatoire continue et densité	191
10.3.2	Moments d'une variable aléatoire continue	192
10.3.3	Lois à densités usuelles	193
10.4	Complément : quelques mots sur les théorèmes limite en probabilité	199
10.4.1	Loi forte des grands nombres	199
10.4.2	Théorème Central Limite	200
10.4.3	Applications en statistique inférentielle	200
A	Table de la fonction de répartition de la loi normale centrée réduite	205
A.1	Lecture dans la table	205
A.1.1	Lecture directe	205
A.1.2	Cas des valeurs non indiquées dans la table	205

Chapitre 1

Algèbre matricielle et systèmes linéaires

Lorsque l'on a plusieurs/un grand nombre de données, il est pratique de pouvoir les synthétiser sous la forme d'un tableau, d'une matrice. Les objectifs de ce chapitre sont d'introduire la notion de matrice et les opérations sur les matrices ainsi que d'établir un lien entre matrices et systèmes linéaires. On discutera également des méthodes de résolution de systèmes linéaires.

1.1 Premières définitions

Définition 1.1. On appelle matrice de taille $m \times n$ et à coefficients réels un tableau de nombres réels constitué de m lignes et n colonnes.

Notations 1.1. 1. On note $\mathcal{M}_{m,n}(\mathbf{R})$ l'ensemble des matrices de taille $m \times n$ et à coefficients réels.

2. Soit $A \in \mathcal{M}_{m,n}(\mathbf{R})$. Pour $i \in \{1, \dots, m\}$ et $j \in \{1, \dots, n\}$, on note $a_{i,j}$ le coefficient situé sur la i^e ligne et dans la j^e colonne de la matrice A . Celle-ci s'écrit alors sous la forme $A = (a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n}$.

Définition 1.2. Soit $A \in \mathcal{M}_{m,n}(\mathbf{R})$.

1. Si $m = 1$, A est appelée vecteur ligne.
2. Si $n = 1$, A est appelée vecteur colonne.
3. Si $m = n$, A est appelée matrice carrée d'ordre n .

Exemple 1.1. Soient

$$A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \\ 9 & 7 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 5 & 9 & 13 \\ 2 & 6 & 10 & 14 \\ 3 & 7 & 11 & 15 \\ 4 & 8 & 12 & 16 \end{pmatrix}, \quad X = (1 \ 2 \ 3 \ 4 \ 5) \quad \text{et} \quad Y = \begin{pmatrix} 6 \\ 7 \\ 8 \\ 9 \end{pmatrix}$$

La matrice A est de taille 3×2 , B est une matrice carrée d'ordre 4, X un vecteur ligne de taille 5 et Y un vecteur colonne de taille 4.

Définition 1.3. Soit $A \in \mathcal{M}_{n,n}(\mathbf{R})$ une matrice carrée.

1. La matrice A est dite diagonale si $a_{i,j} = 0$, pour tous $i, j \in \{1, \dots, n\}$ tels que $i \neq j$.
Si de plus, pour tout $i \in \{1, \dots, n\}$, $a_{i,i} = 1$, la matrice A est appelée matrice identité d'ordre n et est notée I_n .
2. La matrice A est dite triangulaire supérieure si $a_{i,j} = 0$, pour tous $i, j \in \{1, \dots, n\}$ tels que $i > j$.
3. La matrice A est dite triangulaire inférieure si $a_{i,j} = 0$, pour tous $i, j \in \{1, \dots, n\}$ tels que $i < j$.

Ainsi, toute matrice diagonale D est de la forme

$$D = \begin{pmatrix} d_{1,1} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & d_{n,n} \end{pmatrix},$$

la matrice identité d'ordre n est de la forme

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

et toute matrice triangulaire supérieure U ou triangulaire inférieure L est de la forme :

$$U = \begin{pmatrix} u_{1,1} & \cdots & \cdots & u_{1,n} \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & u_{n,n} \end{pmatrix} \quad \text{et} \quad L = \begin{pmatrix} l_{1,1} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ l_{n,1} & \cdots & \cdots & l_{n,n} \end{pmatrix}.$$

1.2 Calcul matriciel

1.2.1 Somme et différence

Définition 1.4. Soient $A = (a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n}$, $B = (b_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} \in \mathcal{M}_{m,n}(\mathbf{R})$.

La somme $A + B$ (resp. différence $A - B$) de A et de B est la matrice de taille $m \times n$ dont les coefficients sont donnés par $a_{i,j} + b_{i,j}$ (resp. $a_{i,j} - b_{i,j}$).

Remarque 1.1. Pour que leur somme et leur différence soient définies, les matrices A et B doivent être de même tailles.

Exemple 1.2. On a :

$$\begin{pmatrix} 1 & 3 & 4 \\ 2 & 5 & 6 \end{pmatrix} + \begin{pmatrix} 1 & 4 & 1 \\ 2 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 1+1 & 3+4 & 4+1 \\ 2+2 & 5+4 & 6+2 \end{pmatrix} = \begin{pmatrix} 2 & 7 & 5 \\ 4 & 9 & 8 \end{pmatrix}$$

et

$$\begin{pmatrix} 1 & 3 & 4 \\ 2 & 5 & 6 \end{pmatrix} - \begin{pmatrix} 1 & 4 & 1 \\ 2 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 1-1 & 3-4 & 4-1 \\ 2-2 & 5-4 & 6-2 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 3 \\ 0 & 1 & 4 \end{pmatrix}.$$

Proposition 1.1. 1. L'addition de matrices est associative, c'est-à-dire que, pour tous $A, B, C \in \mathcal{M}_{m,n}(\mathbf{R})$, on a :

$$(A + B) + C = A + (B + C).$$

On notera simplement cette somme $A + B + C$.

2. L'addition de matrices est commutative, c'est-à-dire que, pour tous $A, B \in \mathcal{M}_{m,n}(\mathbf{R})$, on a :

$$A + B = B + A.$$

3. La matrice $\mathbf{0}_{m,n} \in \mathcal{M}_{m,n}(\mathbf{R})$ dont tous les coefficients sont nuls est l'élément neutre de l'addition dans $\mathcal{M}_{m,n}(\mathbf{R})$, c'est-à-dire que, pour tout $A \in \mathcal{M}_{m,n}(\mathbf{R})$, on a :

$$A + \mathbf{0}_{m,n} = \mathbf{0}_{m,n} + A = A.$$

Preuve :

1.

$$\begin{aligned} (A + B) + C &= (a_{i,j} + b_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} + (c_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} \\ &= (a_{i,j} + b_{i,j} + c_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} \\ &= (a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} + (b_{i,j} + c_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} = A + (B + C). \end{aligned}$$

2.

$$A + B = (a_{i,j} + b_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} = (b_{i,j} + a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} = B + A.$$

3.

$$A + \mathbf{0}_{m,n} = (a_{i,j} + 0)_{1 \leq i \leq m, 1 \leq j \leq n} = A$$

et

$$\mathbf{0}_{m,n} + A = (0 + a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} = A.$$

□

Définition 1.5. Soit $A = (a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} \in \mathcal{M}_{m,n}(\mathbf{R})$.

On appelle opposée de A la matrice $-A = (-a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} \in \mathcal{M}_{m,n}(\mathbf{R})$.

Remarque 1.2. On a ainsi :

$$A + (-A) = \mathbf{0}_{m,n}.$$

1.2.2 Multiplication par un scalaire

Définition 1.6. Soient $A = (a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} \in \mathcal{M}_{m,n}(\mathbf{R})$ et $\alpha \in \mathbf{R}$.

Le produit αA est la matrice de taille $m \times n$ dont les coefficients sont donnés par $\alpha a_{i,j}$.

Remarque 1.3. En particulier $(-1) \times A = -A$.

Exemple 1.3. On a :

$$7 \times \begin{pmatrix} 1 & 4 \\ 2 & 5 \end{pmatrix} = \begin{pmatrix} 7 & 28 \\ 14 & 35 \end{pmatrix}.$$

Proposition 1.2. Soient $A, B \in \mathcal{M}_{m,n}(\mathbf{R})$ et $\alpha, \beta \in \mathbf{R}$.

On a :

1. $(\alpha + \beta)A = \alpha A + \beta A$;
2. $\alpha(A + B) = \alpha A + \alpha B$.

Preuve :

1.

$$\begin{aligned} (\alpha + \beta)A &= ((\alpha + \beta)a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} = (\alpha a_{i,j} + \beta a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} \\ &= (\alpha a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} + (\beta a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} = \alpha A + \beta A. \end{aligned}$$

2.

$$\begin{aligned} \alpha(A + B) &= \alpha(a_{i,j} + b_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} = (\alpha(a_{i,j} + b_{i,j}))_{1 \leq i \leq m, 1 \leq j \leq n} \\ &= (\alpha a_{i,j} + \alpha b_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} \\ &= (\alpha a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} + (\alpha b_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} = \alpha A + \alpha B. \end{aligned}$$

□

1.2.3 Multiplication

Définition 1.7. Soit $A = (a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} \in \mathcal{M}_{m,n}(\mathbf{R})$ et $B = (b_{i,j})_{1 \leq i \leq n, 1 \leq j \leq p} \in \mathcal{M}_{n,p}(\mathbf{R})$.

On appelle produit de A par B la matrice $C = (c_{i,j})_{1 \leq i \leq m, 1 \leq j \leq p} \in \mathcal{M}_{m,p}(\mathbf{R})$ dont les coefficients sont donnés par :

$$c_{i,j} = \sum_{k=1}^n a_{i,k} b_{k,j}, \quad 1 \leq i \leq m, 1 \leq j \leq p.$$

Cette matrice est notée AB ou $A \times B$.

Remarque 1.4.

1. Pour que le produit AB soit défini, il est nécessaire que le nombre de colonnes de la matrice de gauche soit égal au nombre de lignes de la matrice de droite.
2. le coefficient $c_{i,j}$ est le produit (scalaire) de la i^{e} ligne de A par la j^{e} colonne de B .

Exemple 1.4. Soient

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 0 \\ 1 & -1 \end{pmatrix} \quad \text{et} \quad C = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 7 & 4 \end{pmatrix}.$$

On a :

$$AB = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 \times 2 + 2 \times 1 & 1 \times 0 + 2 \times (-1) \\ 3 \times 2 + 4 \times 1 & 3 \times 0 + 4 \times (-1) \end{pmatrix} = \begin{pmatrix} 4 & -2 \\ 10 & -4 \end{pmatrix}$$

et

$$BA = \begin{pmatrix} 2 & 0 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 2 \times 1 + 0 \times 3 & 2 \times 2 + 0 \times 4 \\ 1 \times 1 + (-1) \times 3 & 1 \times 2 + (-1) \times 4 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ -2 & -2 \end{pmatrix}.$$

On observe que $AB \neq BA$; cet exemple montre que, même si les produits AB et BA sont définis et de même taille, ils ne sont, en général, pas égaux. Le produit matriciel **n'est pas commutatif**.

Le produit CB n'est pas défini mais le produit BC l'est et vaut :

$$\begin{aligned} BC &= \begin{pmatrix} 2 & 0 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 3 & 7 & 4 \end{pmatrix} \\ &= \begin{pmatrix} 2 \times 1 + 0 \times 3 & 2 \times 2 + 0 \times 7 & 2 \times 0 + 0 \times 4 \\ 1 \times 1 + (-1) \times 3 & 1 \times 2 + (-1) \times 7 & 1 \times 0 + (-1) \times 4 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 0 \\ -2 & -5 & -4 \end{pmatrix}. \end{aligned}$$

Proposition 1.3. 1. *Le produit matriciel n'est pas commutatif, c'est-à-dire qu'en général $AB \neq BA$.*

2. *La matrice identité d'ordre n , I_n , est l'élément neutre du produit matriciel dans $\mathcal{M}_{n,n}(\mathbf{R})$, c'est-à-dire que, pour tout $A \in \mathcal{M}_{n,n}(\mathbf{R})$, on a :*

$$AI_n = I_nA = A.$$

3. *Lorsque tous les produits sont définis, on a :*

$$(AB)C = A(BC).$$

On notera simplement ce produit ABC .

Preuve :

1. Un contre exemple a déjà été donné dans l'Exemple 1.4.
2. C'est un calcul immédiat.
3. Soient $A = (a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n}$, $B = (b_{i,j})_{1 \leq i \leq n, 1 \leq j \leq p}$ et $C = (c_{i,j})_{1 \leq i \leq p, 1 \leq j \leq q}$.

$$\begin{aligned} (AB)C &= \left(\sum_{k=1}^n a_{i,k} b_{k,j} \right)_{1 \leq i \leq m, 1 \leq j \leq p} (c_{i,j})_{1 \leq i \leq p, 1 \leq j \leq q} \\ &= \left(\sum_{l=1}^p \left(\sum_{k=1}^n a_{i,k} b_{k,l} \right) c_{l,j} \right)_{1 \leq i \leq m, 1 \leq j \leq q} \\ &= \left(\sum_{k=1}^n a_{i,k} \left(\sum_{l=1}^p b_{k,l} c_{l,j} \right) \right)_{1 \leq i \leq m, 1 \leq j \leq q} \\ &= (a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} \left(\sum_{l=1}^p b_{i,l} c_{l,j} \right)_{1 \leq i \leq n, 1 \leq j \leq q} \\ &= A(BC). \end{aligned}$$

□

Définition 1.8. Soit $A \in \mathcal{M}_{n,n}(\mathbf{R})$ une matrice carrée.

Sous réserve d'existence, on appelle inverse de A la matrice B telle que $AB = BA = I_n$. Si elle existe, cette matrice est notée A^{-1} .

Remarque 1.5. En fait, il suffit que l'une des identités $AB = I_n$ ou $BA = I_n$ soit vérifiée pour pouvoir affirmer que $B = A^{-1}$.

Exemple 1.5. La matrice

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

admet pour inverse

$$A^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}.$$

En effet,

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 2 \times 1 + 1 \times (-1) & 2 \times (-1) + 1 \times 2 \\ 1 \times 1 + 1 \times (-1) & 1 \times (-1) + 1 \times 2 \end{pmatrix} = I_2.$$

La matrice

$$B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

n'est, quant à elle, pas inversible. Vérifions ceci en raisonnant par l'absurde. Supposons qu'il existe a, b, c et d tels que

$$B \begin{pmatrix} a & b \\ c & d \end{pmatrix} = I_2.$$

En effectuant le produit

$$B \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+c & b+d \\ a+c & b+d \end{pmatrix}$$

puis en identifiant coefficient par coefficient avec I_2 , on obtient que $1 = a + b = 0$. Absurde. Ainsi, B n'est pas inversible.

Proposition 1.4. Soient $A, B \in \mathcal{M}_{n,n}(\mathbf{R})$ deux matrices inversibles.

La matrice AB est inversible et on a :

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Preuve : On a :

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n.$$

□

Notation 1.2. Pour $n \in \mathbf{N}$, on note :

$$A^n = \underbrace{A \times \dots \times A}_{n \text{ fois}}$$

avec la convention $A^0 = I$.

1.2.4 Transposition

Définition 1.9. Soit $A = (a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} \in \mathcal{M}_{m,n}(\mathbf{R})$.

On appelle transposée de A la matrice $B = (b_{i,j})_{1 \leq i \leq n, 1 \leq j \leq m} \in \mathcal{M}_{n,m}(\mathbf{R})$ telle que pour tous i, j , $b_{i,j} = a_{j,i}$. Cette matrice est notée ${}^t A$ ou A^T .

Exemple 1.6. La transposée de la matrice

$$A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \\ 9 & 7 \end{pmatrix}$$

est la matrice

$$A^T = \begin{pmatrix} 1 & 2 & 9 \\ 4 & 3 & 7 \end{pmatrix}.$$

Proposition 1.5. Soient A, B deux matrices et $\alpha \in \mathbf{R}$

1. On a :

$$(\alpha A)^T = \alpha A^T \text{ et } (A^T)^T = A.$$

2. Si A et B sont de même taille, on a :

$$(A + B)^T = A^T + B^T.$$

3. Si le produit AB est bien défini, on a :

$$(AB)^T = B^T A^T.$$

4. Si A est inversible, alors sa transposée est inversible et on a :

$$(A^T)^{-1} = (A^{-1})^T.$$

Preuve : Les deux premiers points sont évidents. Pour le troisième, il suffit d'écrire :

$$\left[(AB)^T \right]_{i,j} = (AB)_{j,i} = \sum_k a_{j,k} b_{k,i} = \sum_k b_{k,i} a_{j,k} = \sum_k (b^T)_{i,k} (a^T)_{k,j} = [B^T A^T]_{i,j};$$

et pour le quatrième :

$$A^T (A^{-1})^T = (A^{-1} A)^T = I^T = I.$$

□

1.2.5 Déterminant

Définition 1.10. Soit $A \in \mathcal{M}_{2,2}(\mathbf{R})$ une matrice carrée de taille 2.

On appelle déterminant de la matrice A la quantité :

$$\det(A) = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}. \quad (1.2.1)$$

Définition[-Théorème] 1.11. Soit $A \in \mathcal{M}_{n,n}(\mathbf{R})$ une matrice carrée de taille n .

On appelle déterminant de la matrice A la quantité :

$$\det(A) = \sum_{i=1}^n (-1)^{i+j} a_{i,j} \det(A_{i,j}), \quad (1.2.2)$$

où la matrice $A_{i,j}$ est obtenue à partir de A en supprimant la i^e et la j^e colonne. Cette quantité ne dépend pas de j .

Remarque 1.6.

1. La définition précédente est récursive. Pour calculer le déterminant d'une matrice carrée de taille $n \geq 3$, on le développe le long de colonnes à l'aide de la formule (1.2.2) pour se ramener à une combinaison linéaire de déterminant 2×2 que l'on calcule avec la formule (1.2.1).
2. On peut également développer un déterminant le long des lignes avec la formule :

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{i,j} \det(A_{i,j}), \quad (1.2.3)$$

où la matrice $A_{i,j}$ est obtenue à partir de A en supprimant la i^e et la j^e colonne.

Exemple 1.7. On a :

$$\det \begin{pmatrix} 2 & 5 \\ -3 & 7 \end{pmatrix} = 2 \times 7 - 5 \times (-3) = 29.$$

En développant le long de la première colonne, on obtient :

$$\begin{aligned} \det \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix} &= 1 \times \det \begin{pmatrix} 5 & 8 \\ 6 & 9 \end{pmatrix} - 2 \times \det \begin{pmatrix} 4 & 7 \\ 6 & 9 \end{pmatrix} + 3 \times \det \begin{pmatrix} 4 & 7 \\ 5 & 8 \end{pmatrix} \\ &= 1 \times (5 \times 9 - 8 \times 6) - 2 \times (4 \times 9 - 7 \times 6) + 3 \times (4 \times 8 - 7 \times 5) \\ &= -3 + 12 - 9 = 0. \end{aligned}$$

Lorsque $n \geq 3$, on peut profiter de l'éventuelle présence de 0 sur une ligne ou une colonne pour économiser des calculs lors du développement du déterminant. Considérons par exemple la matrice

$$A = \begin{pmatrix} 4 & 3 & 2 & 0 \\ 0 & 0 & 2 & 0 \\ 2 & 0 & 1 & 0 \\ 13 & 12 & 11 & 10 \end{pmatrix}.$$

On a tout intérêt à commencer par développer le long de la 4^e colonne puisque les 3 premiers termes du développement disparaîtront du fait des 0. On a :

$$\begin{aligned} \det(A) &= \det \begin{pmatrix} 4 & 3 & 2 & 0 \\ 0 & 0 & 2 & 0 \\ 2 & 0 & 1 & 0 \\ 13 & 12 & 11 & 10 \end{pmatrix} \\ &= -0 \times \det \begin{pmatrix} 0 & 0 & 2 \\ 2 & 0 & 1 \\ 13 & 12 & 11 \end{pmatrix} + 0 \times \det \begin{pmatrix} 4 & 3 & 2 \\ 2 & 0 & 1 \\ 13 & 12 & 11 \end{pmatrix} \\ &\quad - 0 \times \det \begin{pmatrix} 4 & 3 & 2 \\ 0 & 0 & 2 \\ 13 & 12 & 11 \end{pmatrix} + 10 \times \det \begin{pmatrix} 4 & 3 & 2 \\ 0 & 0 & 2 \\ 2 & 0 & 1 \end{pmatrix} \\ &= 10 \times \det \begin{pmatrix} 4 & 3 & 2 \\ 0 & 0 & 2 \\ 2 & 0 & 1 \end{pmatrix}. \end{aligned}$$

On profite ensuite des 0 de la 2^e colonne pour obtenir que :

$$\begin{aligned} \det(A) &= 10 \times \det \begin{pmatrix} 4 & 3 & 2 \\ 0 & 0 & 2 \\ 2 & 0 & 1 \end{pmatrix} \\ &= 10 \times \left(-3 \times \det \begin{pmatrix} 0 & 2 \\ 2 & 1 \end{pmatrix} + 0 \times \det \begin{pmatrix} \star & \star \\ \star & \star \end{pmatrix} - 0 \times \det \begin{pmatrix} \star & \star \\ \star & \star \end{pmatrix} \right) \\ &= -30 \det \begin{pmatrix} 0 & 2 \\ 2 & 1 \end{pmatrix} = -30 \times (0 \times 1 - 2 \times 2) = 120. \end{aligned}$$

Comme le suggèrent les exemples précédents, le calcul d'un déterminant peut être long et couteux en calculs. La proposition suivante contient des règles de simplification permettant de calculer un déterminant plus efficacement.

Proposition 1.6. 1. L'échange de deux lignes ou deux colonnes, change le signe du déterminant.

2. Si une ligne ou une colonne est nulle, le déterminant est nul.

3. Si on multiplie tous les termes d'une ligne ou d'une colonne par un réel α , le déterminant est multiplié par α .

4. Si l'on ajoute à une colonne (ou une ligne) un multiple d'une **autre** colonne (ou d'une **autre** ligne), la valeur du déterminant ne change pas. En particulier, si deux lignes ou deux colonnes sont identiques, le déterminant est nul.

5. Le déterminant d'une matrice triangulaire est le produit des coefficients diagonaux. En particulier, $\det(I) = 1$.

6. Si $A, B \in \mathcal{M}_{n,n}(\mathbf{R})$, on a :

$$\det(AB) = \det(A) \det(B).$$

En particulier, si A est inversible $\det(A^{-1}) = \det(A)^{-1}$.

La règle de calcul 4. se révèle extrêmement efficace dans la pratique comme l'illustre l'exemple suivant.

Exemple 1.8. Considérons la matrice :

$$B = \begin{pmatrix} 4 & 3 & 2 & 2 \\ 0 & 0 & 2 & 2 \\ 2 & 0 & 1 & 1 \\ 13 & 12 & 11 & 21 \end{pmatrix}.$$

En soustrayant la 3^e colonne à la 4^e on obtient que :

$$\det(B) = \det \begin{pmatrix} 4 & 3 & 2 & 2 \\ 0 & 0 & 2 & 2 \\ 2 & 0 & 1 & 1 \\ 13 & 12 & 11 & 21 \end{pmatrix} = \det \begin{pmatrix} 4 & 3 & 2 & 0 \\ 0 & 0 & 2 & 0 \\ 2 & 0 & 1 & 0 \\ 13 & 12 & 11 & 10 \end{pmatrix}.$$

On a ainsi fait apparaître trois 0 dans la dernière colonne qui vont largement simplifier les calculs. En reconnaissant, dans le membre de droite de la dernière identité, le déterminant de la matrice A de l'Exemple 1.7, on obtient que $\det(B) = 120$.

La notion de déterminant permet d'énoncer un critère simple pour l'inversibilité des matrices.

Théorème 1.1 (Admis). *Soit $A \in \mathcal{M}_{n,n}(\mathbf{R})$.*

Les assertions suivantes sont équivalentes :

1. A est inversible,
2. $\det(A) \neq 0$.

Il est possible d'exprimer l'inverse d'une matrice en fonction de son déterminant. Toutefois cette expression conduit à de trop lourds calculs lorsque la matrice est de taille plus grande que 3×3 (le coût algorithmique étant de l'ordre de $n!$). On aura, dans ces cas de figure, recours aux liens entre l'inversion de matrices et les systèmes linéaires ainsi qu'aux méthodes de résolution de systèmes présentées dans la section suivante (celles-ci ont un coût algorithmique de l'ordre de n^2).

Proposition 1.7. *Soit $A \in \mathcal{M}_{2,2}(\mathbf{R})$ telle que $\det(A) \neq 0$.*

Alors, l'inverse de A est :

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} a_{2,2} & -a_{1,2} \\ -a_{2,1} & a_{1,1} \end{pmatrix}.$$

Preuve : Exercice. □

1.3 Systèmes d'équations linéaires

Définition 1.12. *On appelle système de m équations (linéaires) à n inconnues toute collection de m équations faisant intervenir n variables x_1, \dots, x_n . Un tel système se présente sous la*

Remarque 1.8. Si $\det(A) = 0$ le système n'admet pas de solution ou admet une infinité de solution.

Certains systèmes linéaires sont extrêmement simples à résoudre. C'est le cas des systèmes diagonaux et triangulaires. En effet, la solution d'un système diagonal :

$$\begin{cases} a_{1,1}x_1 = y_1 \\ \vdots \\ a_{n,n}x_n = y_n \end{cases}$$

est clairement

$$\begin{cases} x_1 = \frac{y_1}{a_{1,1}} \\ \vdots \\ x_n = \frac{y_n}{a_{n,n}} \end{cases}$$

sous réserve que tous les $a_{i,i}$ soient non nuls. Si le système est triangulaire supérieur, avec tous les $a_{i,i}$ non nuls,

$$\begin{cases} a_{1,1}x_1 + \dots + a_{1,n}x_n = y_1 \\ \quad \quad \quad \ddots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ \quad \quad \quad \quad \quad \quad \quad a_{n,n}x_n = y_n \end{cases}$$

la dernière équation permet de déterminer x_n . On peut ensuite déterminer x_{n-1} à l'aide de la valeur de x_n maintenant connue et de l'avant dernière équation. En remontant de proche en proche, on peut déterminer les valeurs de x_1, \dots, x_n en ne résolvant qu'une équation du premier degré à chaque étape.

Le fait que les systèmes diagonaux et triangulaires soient particulièrement simples à résoudre a inspiré des méthodes de résolution de systèmes carrés généraux telles que la *méthode de Gauss* et la *méthode de Gauss-Jordan*. L'idée est, dans les deux cas, de se ramener par des opérations élémentaires sur les lignes du système à un système équivalent et triangulaire ou diagonal. Ces méthodes seront présentées dans les deux sous-paragraphes suivants.

Définition 1.15. On appelle opérations élémentaires sur les lignes les opérations suivantes :

1. l'interversion de deux lignes,
2. la multiplication une ligne par un scalaire,
3. l'ajout du produit d'un scalaire et d'une ligne à une **autre** ligne.

Proposition 1.8. Les opérations élémentaires sur les lignes transforment un système en un système équivalent.

Les méthodes de Gauss et Gauss-Jordan consistent à effectuer des opérations élémentaires sur les lignes du système pour le réduire à un système triangulaire ou diagonal. Par souci de concision et de légèreté des notations, on présentera le système lors de ces étapes sous la forme d'un tableau faisant apparaître dans sa partie gauche la matrice A associée au système et dans sa partie droite le membre de droite du système, les deux étant séparés par une barre verticale. Le système dont l'écriture matricielle est :

$$\begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

sera donc représenté par :

$$\begin{array}{ccc|c} a_{1,1} & \cdots & a_{1,n} & y_1 \\ \vdots & & \vdots & \vdots \\ a_{n,1} & \cdots & a_{n,n} & y_n \end{array}$$

Attention : les opérations sur les lignes s'entendent sur l'intégralité de la ligne (membre de droite y compris).

Méthode de Gauss

Cette méthode se déroule en deux temps. Le principe est de transformer, dans un premier temps, le système présenté sous la forme d'un tableau :

$$\begin{array}{ccc|c} a_{1,1} & \cdots & a_{1,n} & y_1 \\ \vdots & & \vdots & \vdots \\ a_{n,1} & \cdots & a_{n,n} & y_n \end{array}$$

en un système triangulaire équivalent avec des 1 sur la diagonale :

$$\begin{array}{cccc|c} 1 & \star & \cdots & \star & \star \\ & \ddots & \ddots & \vdots & \vdots \\ & & & \ddots & \star \\ & & & & 1 & \star \end{array}$$

La deuxième phase est une phase de remontée permettant d'obtenir la solution de ce dernier système qui est la même que celle du système d'origine.

L'algorithme itératif suivant présente la phase de réduction lors de la résolution d'un système $n \times n$ d'équations linéaires par la méthode de Gauss. Avec un léger abus de notation, on notera $a_{i,j}$ et y_i les coefficients apparaissant dans le tableau à l'étape courante. Les valeurs des $a_{i,j}$ et y_i changent donc d'une étape à l'autre. On notera L_i la i^{e} ligne du tableau et C_j sa j^{e} colonne.

Algorithme 1.1 (Méthode de Gauss - Phase de réduction).

Pour i variant de 1 à n :

1. Si $a_{i,i} = 0$, échanger L_i avec une des lignes L_j pour un $j > i$ telle que $a_{j,i} \neq 0$; si un tel échange n'est pas possible, s'arrêter et conclure que le système n'admet pas une unique solution.
2. Remplacer la L_i par $\frac{1}{a_{i,i}}L_i$.
3. Pour j variant de $i + 1$ à n :
Remplacer L_j par $L_j - a_{j,i}L_i$.

Remarque 1.9.

1. À la i^{e} itération, le coefficient $a_{i,i}$ est appelé i^{e} pivot.
2. Si le système admet une unique solution, l'algorithme ne s'arrêtera pas durant l'étape 1.. On peut, pour s'en assurer, calculer le déterminant de la matrice associée au système.
3. À la i^{e} itération, l'étape 2 permet de mettre un 1 au niveau du i^{e} coefficient diagonal.

4. À la i^e itération, l'étape 3 permet de mettre à 0 tous les coefficients de la colonne C_i situés en dessous du coefficient diagonal.
5. À la n^e itération, l'étape 3 ne contient aucune opération (il n'y a rien à faire).

La sortie de cet algorithme est un tableau de la forme :

$$\begin{array}{cccc|c} 1 & a'_{1,2} & \dots & a'_{1,n} & b_1 \\ & \ddots & \ddots & \vdots & \vdots \\ & & \ddots & a'_{n-1,n} & \vdots \\ & & & 1 & b_n \end{array}$$

correspondant au système :

$$\begin{cases} x_1 + a'_{1,2}x_2 + \dots + a'_{1,n}x_n = b_1 \\ \dots \dots \dots \dots \dots \dots \dots \\ x_{n-1} + a'_{n-1,n}x_n = b_{n-1} \\ x_n = b_n \end{cases}$$

Un tel système se résout aisément par remontée (on connaît déjà la valeur de x_n !) en utilisant l'algorithme suivant.

Algorithme 1.2 (Méthode de Gauss - Phase de remontée).

Affecter à x_n la valeur b_n .

Pour i descendant de $n - 1$ à 1 :

1. Remplacer x_{i+1}, \dots, x_n par leurs valeurs déjà trouvées dans L_i .
2. Résoudre en x_i l'équation du premier degré apparaissant dans la ligne L_i .

Sortie : x_1, \dots, x_n solution du système.

Remarque 1.10. Alternativement, on peut remplacer les étapes 1. et 2. par : Affecter à x_i la valeur :

$$b_i - a'_{i,i+1}x_{i+1} - \dots - a'_{i,n}x_n.$$

Exemple 1.10 (suite). Reprenons l'exemple du système (S₂) de l'Exemple 1.10. Rappelons que celui-ci s'écrit matriciellement sous la forme $AX = Y$ avec :

$$A = \begin{pmatrix} 2 & 8 & 12 \\ 1 & 4 & 8 \\ 3 & 4 & 2 \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad \text{et} \quad Y = \begin{pmatrix} 34 \\ 19 \\ 19 \end{pmatrix}.$$

Nous allons le résoudre en utilisant l'algorithme de Gauss et une présentation sous forme de tableaux. La phase de réduction/triangularisation débute avec le tableau :

$$\begin{array}{ccc|c} 2 & 8 & 12 & 34 \\ 1 & 4 & 8 & 19 \\ 3 & 4 & 2 & 19 \end{array}$$

Le premier pivot est **2** (non nul). On effectue l'opération $L_1 \leftarrow \frac{1}{2}L_1$ pour obtenir :

$$\begin{array}{ccc|c} \mathbf{1} & 4 & 6 & 17 \\ 1 & 4 & 8 & 19 \\ 3 & 4 & 2 & 19 \end{array} ,$$

puis les opérations $L_2 \leftarrow L_2 - L_1$ et $L_3 \leftarrow L_3 - 3L_1$, pour obtenir :

$$\begin{array}{ccc|c} \mathbf{1} & 4 & 6 & 17 \\ 0 & \mathbf{0} & 2 & 2 \\ 0 & -8 & -16 & -32 \end{array} .$$

Le candidat à être le deuxième pivot est nul ; on intervertit donc les lignes L_2 et L_3 :

$$\begin{array}{ccc|c} 1 & 4 & 6 & 17 \\ 0 & \mathbf{-8} & -16 & -32 \\ 0 & 0 & 2 & 2 \end{array} .$$

Le deuxième pivot est alors **-8** (non nul). On effectue l'opération $L_2 \leftarrow -\frac{1}{8}L_2$ pour obtenir :

$$\begin{array}{ccc|c} 1 & 4 & 6 & 17 \\ 0 & \mathbf{1} & 2 & 4 \\ 0 & 0 & \mathbf{2} & 2 \end{array} .$$

Le coefficient sous le coefficient diagonal dans la deuxième colonne est déjà 0 ; on passe directement à l'étape suivante. Le troisième pivot est **2**. On divise donc les coefficients de la ligne L_3 par 2 pour obtenir :

$$\begin{array}{ccc|c} 1 & 4 & 6 & 17 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & \mathbf{1} & 1 \end{array} .$$

Le tableau est sous forme triangulaire et n'a que des 1 sur la diagonale ; la première phase est terminée. La phase de remontée débute en écrivant le système correspondant à ce tableau (et équivalent à (S_2)) :

$$\begin{cases} x_1 + 4x_2 + 6x_3 = 17 \\ x_2 + 2x_3 = 4 \\ x_3 = 1 \end{cases} .$$

On sait déjà que $x_3 = 1$. En remplaçant x_3 par 1 dans la deuxième ligne, on déduit que $x_2 = 2$. Ensuite, en remplaçant x_2 par 2 et x_3 par 1 dans la première ligne, on déduit que $x_1 = 3$. On obtient ainsi que la solution de ce système est :

$$X = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} .$$

Méthode de Gauss-Jordan

Cette méthode se déroule en une seule phase, mais chaque étape nécessite un peu plus de calculs que pour la méthode de Gauss. Le principe est de transformer le système présenté sous la forme d'un tableau :

$$\begin{array}{ccc|c} a_{1,1} & \dots & a_{1,n} & y_1 \\ \vdots & & \vdots & \vdots \\ a_{n,1} & \dots & a_{n,n} & y_n \end{array}$$

en un système diagonal équivalent. En fait, on fera en sorte qu'à la fin de l'exécution de l'algorithme, la partie gauche du tableau soit la matrice identité I_n . La solution du système sera alors simplement le membre de droite du tableau

L'algorithme itératif suivant la résolution d'un système $n \times n$ d'équations linéaires par la méthode de Gauss-Jordan. Toujours avec le même abus de notation, $a_{i,j}$ et y_i désignerons les coefficients apparaissant dans le tableau à l'étape courante. Les valeurs des $a_{i,j}$ et y_i changent donc d'une étape à l'autre.

Algorithme 1.3 (Méthode de Gauss-Jordan).

Pour i variant de 1 à n :

1. Si $a_{i,i} = 0$, échanger L_i avec une des lignes L_j pour un $j > i$ telle que $a_{j,i} \neq 0$; si un tel échange n'est pas possible, s'arrêter et conclure que le système n'admet pas une unique solution.
2. Remplacer la L_i par $\frac{1}{a_{i,i}}L_i$.
3. Pour j variant de 1 à n à l'exception de i :
Remplacer L_j par $L_j - a_{j,i}L_i$.

Remarque 1.11.

1. À la i^{e} itération, le coefficient $a_{i,i}$ est appelé i^{e} pivot.
2. Si le système admet une unique solution, l'algorithme ne s'arrêtera pas durant l'étape 1.. On peut pour s'en assurer calculer le déterminant de la matrice associée au système.
3. À la i^{e} itération, l'étape 2 permet de mettre un 1 au niveau du i^{e} coefficient diagonal.
4. À la i^{e} itération, l'étape 3 permet de mettre à 0 tous les coefficients de la colonne C_i à l'exception du coefficient diagonal.

Exemple 1.10 (suite). Reprenons l'exemple du système (S_2) de l'Exemple 1.10 et résolvons le, cette fois, en utilisant l'algorithme de Gauss-Jordan et une présentation sous forme de tableaux. On débute avec le tableau :

$$\begin{array}{ccc|c} 2 & 8 & 12 & 34 \\ 1 & 4 & 8 & 19 \\ 3 & 4 & 2 & 19 \end{array} .$$

Le premier pivot est 2 (non nul). On effectue l'opération $L_1 \leftarrow \frac{1}{2}L_1$ pour obtenir :

$$\begin{array}{ccc|c} 1 & 4 & 6 & 17 \\ 1 & 4 & 8 & 19 \\ 3 & 4 & 2 & 19 \end{array} ,$$

puis les opérations $L_2 \leftarrow L_2 - L_1$ et $L_3 \leftarrow L_3 - 3L_1$, pour obtenir :

$$\begin{array}{ccc|c} 1 & 4 & 6 & 17 \\ 0 & 0 & 2 & 2 \\ 0 & -8 & -16 & -32 \end{array} .$$

Le candidat à être le deuxième pivot est nul ; on intervertit donc les lignes L_2 et L_3 :

$$\begin{array}{ccc|c} 1 & 4 & 6 & 17 \\ 0 & -8 & -16 & -32 \\ 0 & 0 & 2 & 2 \end{array} .$$

Le deuxième pivot est alors -8 (non nul). On effectue l'opération $L_2 \leftarrow -\frac{1}{8}L_2$ pour obtenir :

$$\begin{array}{ccc|c} 1 & 4 & 6 & 17 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 2 & 2 \end{array} .$$

Remarquons que c'est à partir d'ici que la méthode de Gauss-Jordan diffère de la méthode de Gauss dans cet exemple. On réalise l'opération $L_1 \leftarrow L_1 - 4L_2$ pour obtenir :

$$\begin{array}{ccc|c} 1 & 0 & -2 & 1 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 2 & 2 \end{array} .$$

Le troisième pivot est 2 . On divise donc les coefficients de la ligne L_3 par 2 pour obtenir :

$$\begin{array}{ccc|c} 1 & 0 & -2 & 1 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 1 \end{array} .$$

En effectuant les opérations $L_1 \leftarrow L_1 + 2L_3$ et $L_2 \leftarrow L_2 - 2L_3$, on obtient :

$$\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{array} .$$

On retrouve que, comme nous l'avons vu avec la méthode de Gauss, la solution de ce système est :

$$X = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} .$$

1.3.3 Application à l'inversion de matrice

Si une matrice A est inversible, la i^e colonne de son inverse A^{-1} n'est autre que la solution de l'équation matricielle $AX = e_i$, où e_i est le i^e vecteur de la base canonique de \mathbf{R}^n , c'est-

à-dire, le vecteur :

$$e_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i^{\text{e}} \text{ ligne.}$$

Ainsi, déterminer l'inverse d'une matrice revient à résoudre les n systèmes linéaires dont les écritures matricielles sont $AX = e_1, \dots, AX = e_n$. Ceci peut être fait simultanément en utilisant la méthode de Gauss-Jordan avec pour membre de droite la matrice identité I_n . En d'autre terme, on débutera la méthode de Gauss-Jordan avec le tableau :

$$\begin{array}{cccc|cccc} a_{1,1} & \dots & \dots & a_{1,n} & 1 & 0 & \dots & 0 \\ \vdots & \ddots & & \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \vdots & \vdots & \ddots & \ddots & 0 \\ a_{n,1} & \dots & \dots & a_{n,n} & 0 & \dots & 0 & 1 \end{array}.$$

À la fin de l'exécution de l'algorithme de Gauss-Jordan à partir de ce tableau, on obtiendra un tableau contenant dans son membre de droite l'inverse de la matrice A

$$\begin{array}{cccc|cccc} 1 & 0 & \dots & 0 & \star & \dots & \dots & \star \\ 0 & \ddots & \ddots & \vdots & \vdots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & 1 & \star & \dots & \dots & \star \end{array} \underbrace{\hspace{10em}}_{=A^{-1}}$$

Exemple 1.10 (suite). Reprenons l'exemple du système (S_2) de l'Exemple 1.10 et déterminons l'inverse de la matrice A en utilisant l'algorithme de Gauss-Jordan et une présentation sous forme de tableaux. On débute avec le tableau :

$$\begin{array}{ccc|ccc} 2 & 8 & 12 & 1 & 0 & 0 \\ 1 & 4 & 8 & 0 & 1 & 0 \\ 3 & 4 & 2 & 0 & 0 & 1 \end{array}.$$

Le premier pivot est 2 (non nul). On effectue l'opération $L_1 \leftarrow \frac{1}{2}L_1$ pour obtenir :

$$\begin{array}{ccc|ccc} 1 & 4 & 6 & \frac{1}{2} & 0 & 0 \\ 1 & 4 & 8 & 0 & 1 & 0 \\ 3 & 4 & 2 & 0 & 0 & 1 \end{array},$$

puis les opérations $L_2 \leftarrow L_2 - L_1$ et $L_3 \leftarrow L_3 - 3L_1$, pour obtenir :

$$\begin{array}{ccc|ccc} 1 & 4 & 6 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 2 & -\frac{1}{2} & 1 & 0 \\ 0 & -8 & -16 & -\frac{3}{2} & 0 & 1 \end{array}.$$

Le candidat à être le deuxième pivot est nul ; on intervertit donc les lignes L_2 et L_3 :

$$\begin{array}{ccc|ccc} 1 & 4 & 6 & \frac{1}{2} & 0 & 0 \\ 0 & -8 & -16 & -\frac{3}{2} & 0 & 1 \\ 0 & 0 & 2 & -\frac{1}{2} & 1 & 0 \end{array} .$$

Le deuxième pivot est alors -8 (non nul). On effectue l'opération $L_2 \leftarrow -\frac{1}{8}L_2$ pour obtenir :

$$\begin{array}{ccc|ccc} 1 & 4 & 6 & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 2 & \frac{3}{16} & 0 & -\frac{1}{8} \\ 0 & 0 & 2 & -\frac{1}{2} & 1 & 0 \end{array} .$$

On réalise, en suite, l'opération $L_1 \leftarrow L_1 - 4L_2$ pour obtenir :

$$\begin{array}{ccc|ccc} 1 & 0 & -2 & -\frac{1}{4} & 0 & \frac{1}{2} \\ 0 & 1 & 2 & \frac{3}{16} & 0 & -\frac{1}{8} \\ 0 & 0 & 2 & -\frac{1}{2} & 1 & 0 \end{array} .$$

Le troisième pivot est 2 . On divise donc les coefficients de la ligne L_3 par 2 pour obtenir :

$$\begin{array}{ccc|ccc} 1 & 0 & -2 & -\frac{1}{4} & 0 & \frac{1}{2} \\ 0 & 1 & 2 & \frac{3}{16} & 0 & -\frac{1}{8} \\ 0 & 0 & 1 & -\frac{1}{4} & \frac{1}{2} & 0 \end{array} .$$

En effectuant les opérations $L_1 \leftarrow L_1 + 2L_3$ et $L_2 \leftarrow L_2 - 2L_3$, on obtient :

$$\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{3}{4} & 1 & \frac{1}{2} \\ 0 & 1 & 0 & \frac{11}{16} & -1 & -\frac{1}{8} \\ 0 & 0 & 1 & -\frac{1}{4} & \frac{1}{2} & 0 \end{array} .$$

Ainsi,

$$A^{-1} = \begin{pmatrix} -\frac{3}{4} & 1 & \frac{1}{2} \\ \frac{11}{16} & -1 & -\frac{1}{8} \\ -\frac{1}{4} & \frac{1}{2} & 0 \end{pmatrix} .$$

On retrouve que la solution de (S_2) lorsque

$$Y = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$$

est

$$X = A^{-1}Y = \begin{pmatrix} -\frac{3}{4} & 1 & \frac{1}{2} \\ \frac{11}{16} & -1 & -\frac{1}{8} \\ -\frac{1}{4} & \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} 34 \\ 19 \\ 19 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} .$$

Le fait d'avoir déterminé l'inverse de A permet de trouver rapidement, sans trop d'efforts, la solution de (S_2) si l'on change arbitrairement son membre de droite. Plus précisément, ceci permet de résoudre, par un simple produit matriciel, tout système de la forme :

$$(S) \begin{cases} 2x_1 + 8x_2 + 12x_3 = y_1 \\ x_1 + 4x_2 + 8x_3 = y_2 \\ 3x_1 + 4x_2 + 2x_3 = y_3 \end{cases} \quad \text{pour } y_1, y_2, y_3 \in \mathbf{R} .$$

- x_1 le nombre de lingots de A_1 fabriqués par l'industriel ce mois ci,
- x_2 le nombre de lingots de A_2 fabriqués par l'industriel ce mois ci.

Définition de l'objectif : L'objectif est de maximiser la marge brute mensuelle de l'industriel. Celle-ci dépend du nombre x_1 de lingots de A_1 et du nombre x_2 de lingots de A_2 fabriqués ce mois. Puisque la marge brute réalisée sur la vente d'un lingot de A_1 est de 15€ et la marge brute réalisée sur la vente d'un lingot de A_2 est de 17€, la marge brute totale de l'industriel est donnée par :

$$15x_1 + 17x_2.$$

L'objectif est donc de maximiser $z = 15x_1 + 17x_2$.

Définition des contraintes :

- *Contrainte liée à la quantité de M_1 disponible :* Dans un lingot de A_1 , 20% des 500g qu'il pèse, c'est-à-dire 100g, est du M_1 . Dans un lingot de A_2 , 60% des 500g qu'il pèse, c'est-à-dire 300g, est du M_1 . La masse totale de M_1 utilisée $100x_1 + 300x_2g$ ne peut pas excéder la quantité totale de M_1 disponible dans les stocks, c'est-à-dire 30000g (soit 30Kg). On obtient donc une première contrainte :

$$100x_1 + 300x_2 \leq 30000$$

qui est équivalente à :

$$x_1 + 3x_2 \leq 300.$$

- *Contrainte liée à la quantité de M_2 disponible :* Dans un lingot de A_1 , 80% des 500g qu'il pèse, c'est-à-dire 400g, est du M_2 . Dans un lingot de A_2 , 40% des 500g qu'il pèse, c'est-à-dire 200g, est du M_2 . La masse totale de M_2 utilisée $400x_1 + 200x_2g$ ne peut pas excéder la quantité totale de M_2 disponible dans les stocks, c'est-à-dire 50000g (soit 50Kg). On obtient donc une première contrainte :

$$400x_1 + 200x_2 \leq 50000$$

qui est équivalente à :

$$4x_1 + 2x_2 \leq 500.$$

- *Contrainte technique :* On sait que l'industriel ne peut produire plus de 150 lingots par mois. On doit donc avoir :

$$x_1 + x_2 \leq 150.$$

- *Contraintes de « bon sens » :* Bien entendu, l'industriel ne peut fabriquer qu'un nombre entier positif ou nul de lingots de chaque sorte et on doit avoir $x_1, x_2 \in \mathbf{N}$. Pour des raisons techniques, nous relâchons cette contrainte en $x_1, x_2 \geq 0$.

Formulation du problème : Nous pouvons maintenant résumer ce que nous venons de voir en formulant le problème :

$$(P) \left\{ \begin{array}{l} \text{maximiser} \quad z = 15x_1 + 17x_2 \\ \text{s.l.c.} \\ \quad x_1 + 3x_2 \leq 300 \\ \quad 4x_1 + 2x_2 \leq 500 \\ \quad x_1 + x_2 \leq 150 \\ \quad x_1, x_2 \geq 0 \end{array} \right. .$$

S'il est non vide, le domaine défini par les contraintes d'un problème d'optimisation linéaire (à n variables) est un polyèdre (polygone si $n = 2$) convexe \mathcal{D} . L'ensemble \mathcal{D} est appelé *ensemble admissible*. Les sommets de ce polygone *saturent* au moins n contraintes du problème (*i.e.* leurs coordonnées vérifient toutes les contraintes et au moins n contraintes sont vérifiées avec égalité). En fait, exactement n contraintes sont saturées sauf dans les cas pathologiques.

On peut montrer le résultat suivant.

Proposition 1.9. *Si le domaine défini par les contraintes d'un problème d'optimisation linéaire (P) est borné, alors le problème (P) admet une solution atteinte en au moins un sommet de ce domaine.*

1.4.1 Méthode graphique

Lorsque le nombre n de variables d'un problème d'optimisation linéaire est faible (essentiellement $n = 2$), il est possible de le résoudre de façon graphique. Nous travaillerons ici en dimension 2, c'est-à-dire avec deux variables x_1 et x_2 , pour optimiser, sous contraintes linéaires, une fonction objectif de la forme $z = \alpha_1 x_1 + \alpha_2 x_2$.

Pour cela, on commence par tracer le polygone \mathcal{D} défini par les contraintes : on trace, pour chaque contrainte la droite représentant le cas de l'égalité, puis on détermine le demi-espace vérifiant la contrainte (celui situé « au-dessus » ou « au-dessous » de la droite).

Exemple 1.11 (suite). Reprenons l'Exemple 1.11. La première contrainte est $x_1 + 3x_2 \leq 300$; après avoir tracé la droite limite d'équation $x_1 + 3x_2 = 300$ (ou de manière équivalente $x_2 = 100 - \frac{1}{3}x_1$), on repère la partie de l'espace située en-dessous de cette droite puisque les variables doivent vérifier une inégalité du type « \leq ».

La deuxième contrainte est $4x_1 + 2x_2 \leq 500$; après avoir tracé la droite limite d'équation $4x_1 + 2x_2 = 500$ (ou de manière équivalente $x_2 = 250 - 2x_1$), on repère la partie de l'espace située en-dessous de cette droite puisque les variables doivent vérifier une inégalité du type « \leq ».

La troisième contrainte est $x_1 + x_2 \leq 150$; après avoir tracé la droite limite d'équation $x_1 + x_2 = 150$ (ou de manière équivalente $x_2 = 150 - x_1$), on repère la partie de l'espace située en-dessous de cette droite puisque les variables doivent vérifier une inégalité du type « \leq ».

N'oublions pas les contraintes $x_1, x_2 \geq 0$ qui indiquent que les variables se situent dans le premier quadrant (quart nord-est) du plan.

Le domaine \mathcal{D} est celui constitué des points vérifiant toutes les conditions listées ci-dessus ; il est représenté dans la Figure 1.1.

Une fois ce domaine tracé, on trace la ligne de niveau $z = 0$ de la fonction objectif, c'est-à-dire la droite d_0 d'équation $0 = z = \alpha_1 x_1 + \alpha_2 x_2$ ainsi qu'un vecteur normal à cette droite (*i.e.* une « flèche perpendiculaire à cette droite ») orienté dans le sens de la direction d'amélioration. Il faut alors faire attention à la nature du problème :

- s'il s'agit d'un problème de *maximisation*, on veut aller dans le sens des niveaux z de plus en plus élevés ;
- s'il s'agit d'un problème de *minimisation*, on veut aller dans le sens des niveaux z de moins en moins élevés.

Il suffit alors de « décaler » la droite d_0 dans le sens de la flèche et de s'arrêter juste avant de quitter le domaine \mathcal{D} . On trace en fait différentes lignes de niveau d_{z^*} d'équation $z = \alpha_1 x_1 + \alpha_2 x_2$. La valeur optimale est alors le niveau z^* de la dernière droite tracée et est atteinte en tout point de $\mathcal{D} \cap d_{z^*}$. Dans les bons cas, $\mathcal{D} \cap d_{z^*}$ est réduit à un point et la solution est unique.

Exemple 1.11 (suite). Reprenons l'Exemple 1.11. On a déjà tracé le polygone \mathcal{D} défini par les contraintes. Après avoir tracé la ligne de niveau 0 de la fonction objectif $(x_1, x_2) \mapsto 15x_1 + 17x_2$, c'est-à-dire $d_0 : 15x_1 + 17x_2 = 0$ (en bleu sur la Figure 1.2), on a déterminé la direction d'amélioration, représentée par une flèche rouge sur cette même figure. Pour cela on s'est souvenu qu'il s'agit d'un problème de maximisation et que l'on cherche donc à « faire monter » le niveau z des lignes de niveau considérées. En « poussant » progressivement la ligne de niveau dans le sens d'amélioration, on voit que le meilleur niveau possible est $z^* = 2400$. La ligne de niveau correspondante est représentée en vert sur la Figure 1.2. On voit aussi que cette ligne de niveau ne rencontre \mathcal{D} qu'en le point $(x_1^*; x_2^*) = (75; 75)$; l'unique solution du

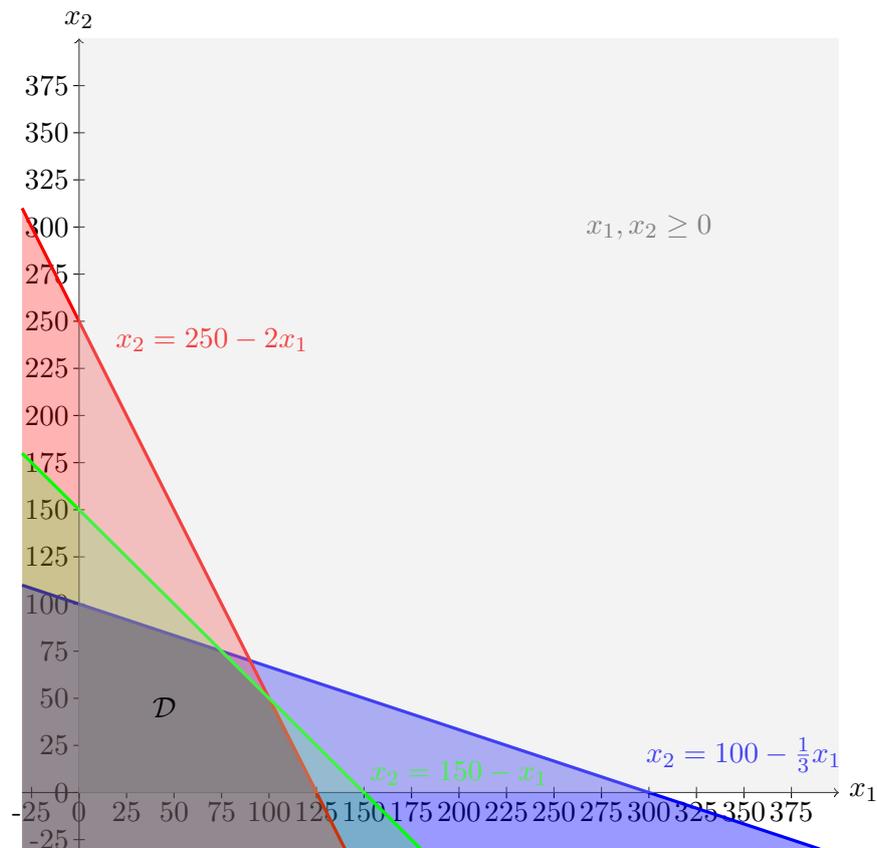


FIGURE 1.1 – Domaine \mathcal{D} (gris) du problème d'optimisation linéaire de l'Exemple 1.11 défini par les contraintes $x_1 + 3x_2 \leq 300$ (bleu), $4x_1 + 2x_2 \leq 500$ (rouge), $x_1 + x_2 \leq 150$ (vert) et $x_1, x_2 \geq 0$ (gris clair).

problème d'optimisation est donc atteinte en ce point.

Ainsi, pour maximiser sa marge brute mensuelle, l'industriel devra produire 75 lingots de chacun des deux alliages.

Remarque 1.12. La méthode graphique est facilement réalisable en dimension deux, mais difficilement applicable en dimension plus grande. Il existe d'autres méthodes, basées sur des algorithmes, permettant une résolution plus efficace de problèmes d'optimisation. Une introduction à l'une d'elle, la *méthode du simplexe*, se trouve dans la Sous-section 1.4.3.

1.4.2 Forme standard d'un problème d'optimisation linéaire

Définition 1.16. On dit qu'un problème d'optimisation linéaire est sous forme standard si toutes les contraintes intervenant dans ce problème sont des contraintes d'égalité à l'exception des contraintes de positivité des variables.

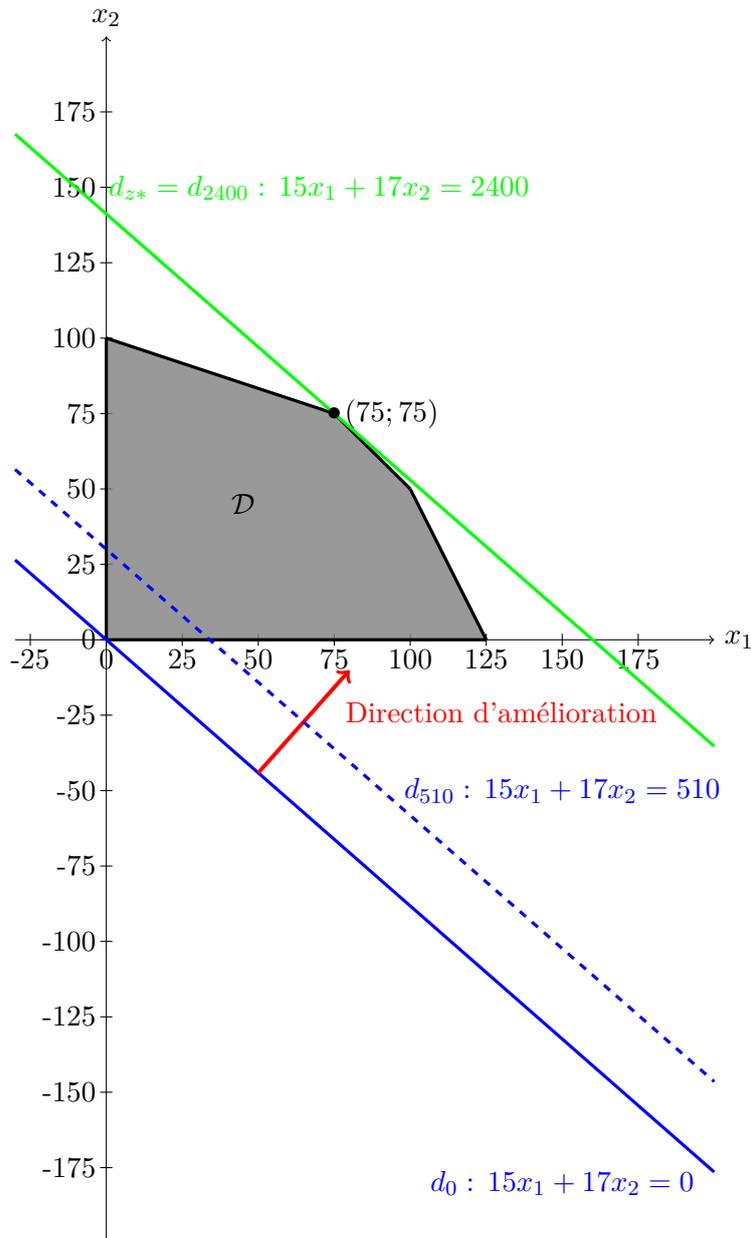


FIGURE 1.2 – Illustration de la résolution du problème d'optimisation linéaire de l'Exemple 1.11 par la méthode graphique.

Remarque 1.13.

1. Quitte à y introduire des variables supplémentaires qui n'interviendront pas dans l'expression de l'objectif, dites *variables d'écart*, tout problème d'optimisation linéaire peut être ramené à un problème sous forme standard. En effet, l'inégalité :

$$a_{i,1}x_1 + \cdots + a_{i,n}x_n \leq b_i,$$

est équivalente à $x_1 + x_2 + e_3 = 150$ avec $e_3 \geq 0$. Ainsi, le problème standard équivalent à (P) est :

$$(P_S) \left\{ \begin{array}{l} \mathbf{maximiser} \quad z = 15x_1 + 17x_2 \\ \mathbf{s.l.c.} \\ x_1 + 3x_2 + e_1 = 300 \\ 4x_1 + 2x_2 + e_2 = 500 \\ x_1 + x_2 + e_3 = 150 \\ x_1, x_2, e_1, e_2, e_3 \geq 0 \end{array} \right. .$$

Celui-ci s'écrit matriciellement sous la forme :

$$(P) \left\{ \begin{array}{l} \mathbf{minimiser} \text{ ou } \mathbf{maximiser} \quad z = CX \\ \mathbf{sous les contraintes} \\ AX = B \\ x_1, \dots, x_n \geq 0 \end{array} \right. ,$$

où

$$A = \begin{pmatrix} 1 & 3 & 1 & 0 & 0 \\ 4 & 2 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 300 \\ 500 \\ 150 \end{pmatrix}, \quad C = (15 \quad 17 \quad 0 \quad 0 \quad 0) \quad \text{et} \quad X = \begin{pmatrix} x_1 \\ x_2 \\ e_1 \\ e_2 \\ e_3 \end{pmatrix} .$$

1.4.3 Introduction à la méthode du simplexe

La *méthode du simplexe*, introduite par Dantzig en 1947, est basée sur l'idée suivante. Si l'ensemble admissible \mathcal{D} d'un problème d'optimisation linéaire est non vide et borné, on sait que la solution est atteinte en au moins un sommet de \mathcal{D} . La première étape est donc de déterminer un sommet de départ S . Ensuite, on regarde s'il est possible d'améliorer (strictement) l'objectif le long d'une des arêtes de \mathcal{D} émanant de S . Il faut être attentif à la nature du problème lors de cette étape : s'il s'agit d'un problème de maximisation (resp. minimisation), on cherche à augmenter (resp. diminuer) la valeur de la fonction objectif. Si une telle arête existe, on détermine l'arête permettant la plus forte amélioration et on sélectionne le sommet situé à l'autre extrémité de l'arête. Si une telle arête n'existe pas, on s'arrête et on conclue que l'on a trouvé la solution optimale. On itère cette procédure en partant du dernier sommet sélectionné jusqu'à s'arrêter et avoir trouvé la solution optimale.

La sélection d'un sommet initial n'est, en général, pas triviale. Elle est pourtant aisée lorsque toutes les contraintes sont du type « \leq » et tous les coefficients des membres de droite b_i sont strictement positifs. En effet, dans ce cas l'origine $(0, \dots, 0)$ est un sommet de \mathcal{D} ! Nous nous plaçons donc, pour toute la suite, dans le cadre d'un problème d'optimisation dont toutes les contraintes sont du type « \leq ».

Remarque 1.14. Lorsque le problème présente des contraintes du type \geq ou mixte, on peut utiliser une variante de l'algorithme du simplexe connue sous le nom de *méthode des deux phases*. La première phase correspond à la détermination d'un sommet de départ ; la seconde est essentiellement une application de la méthode du simplexe. Nous ne présenterons pas la méthode des deux phases dans cette sous-section qui ne se veut qu'introductive. Les éventuels problèmes liés à des sommets dégénérés ne seront pas non plus discutés ici.

- (b) noter j_0 le numéro de la colonne correspondante ;
- (c) noter \mathbf{e} la variable correspondante (variable entrante) ;
- (d) dans chacune des lignes 1 à k calculer le quotient du coefficient de la colonne b_i sur celui de la colonne j_0 ;
- (e) déterminer le plus petit des quotients parmi les positifs et noter i_0 sa ligne ;
- (f) noter \mathbf{s} la variable correspondante (variable sortante) ;
- (g) mettre à jour \mathbf{HB} et \mathbf{B} : $\mathbf{HB} \leftarrow \mathbf{HB} \cup \{\mathbf{s}\} \setminus \{\mathbf{e}\}$ et $\mathbf{B} \leftarrow \mathbf{B} \cup \{\mathbf{e}\} \setminus \{\mathbf{s}\}$ (penser à le faire dans le tableau) ;
- (h) remplacer L_{i_0} par $\frac{1}{a_{i_0, j_0}} L_{i_0}$;
- (i) pour tout $i \in \{1, \dots, k+1\} \setminus \{i_0\}$, remplacer L_i par $L_i - a_{i, j_0} L_{i_0}$.

Sortie : Valeurs du maximum z et des coordonnées du point où il est réalisé x_1, \dots, x_n lues dans le dernier tableau.

Remarque 1.15.

1. Afin de mieux comprendre le fonctionnement de cet algorithme, la lecture des deux exemples ci-dessous est vivement recommandée !
2. Dans les étapes 2. et 3., les choix des variables hors base (et de base) traduisent le fait que l'algorithme démarre du sommet de coordonnées $(x_1; \dots; x_n) = (0; \dots; 0)$.
3. Dans les tableaux comme celui écrit dans le point 4., chaque ligne représente une équation. Par exemple la première ligne représente l'équation $a_{1,1}x_1 + \dots + a_{1,n}x_n + e_1 = b_1$. Cette ligne débute par e_1 , ce qui indique que dans cette équation la seule variable non nulle est e_1 ; ceci permet de lire directement la valeur de e_1 dans l'intersection de cette ligne et de la colonne des b_i . Afin de toujours pouvoir faire une telle lecture directe, on fera en sorte de toujours avoir des 1 dans les intersections des lignes et colonnes correspondant aux variables de base ainsi qu'à $-z$. La dernière ligne du tableau T représente l'équation $z = \alpha_1x_1 + \dots + \alpha_nx_n$ (soit $\alpha_1x_1 + \dots + \alpha_nx_n + (-z) = 0$).
4. Les étapes 5a.-5c. permettent de déterminer la variable dont l'augmentation de la valeur produit la plus grande amélioration marginale ; cette variable est appelée *variable entrante*.
5. Les étapes 5d.-5f. permettent de déterminer dans quelle mesure on peut augmenter la valeur de la variable entrante pour continuer à respecter les contraintes et quelle variable il faut faire sortir de la base pour permettre cette augmentation.
6. S'il y a des ex-æquo lors de la détermination de la variable entrante ou sortante, on choisira par convention la première dans l'ordre lexicographique.
7. Les étapes 5h.-5i. reviennent à l'exécution d'une itération de l'algorithme de Gauss-Jordan. Elle permettent de ré-exprimer les variables de base et z en fonction des variables hors base.
8. Minimiser z revient à maximiser $-z$, il n'est donc pas indispensable de disposer d'un algorithme de minimisation. Nous l'écrivons pourtant ci-dessous.

(h) remplacer L_{i_0} par $\frac{1}{a_{i_0, j_0}} L_{i_0}$;

(i) pour tout $i \in \{1, \dots, k+1\} \setminus \{i_0\}$, remplacer L_i par $L_i - a_{i, j_0} L_{i_0}$.

Sortie : Valeurs du minimum z et des coordonnées du point où il est réalisé x_1, \dots, x_n lues dans le dernier tableau.

Exemple 1.11 (suite). Terminons l'Exemple 1.11 en résolvant le problème (P) par la méthode du simplexe. Avant de réaliser l'exécution de la méthode du simplexe en utilisant la présentation par tableaux et afin d'en comprendre précisément le fonctionnement, on détaillera le cheminement suivi par l'algorithme sur cet exemple sans utiliser de tableau. Ceci permettra de mettre en avant l'effet concret de chaque étape de l'algorithme sur les variables et contraintes du problème. La Figure 1.3 représente le chemin suivi par l'algorithme pour explorer différents sommets de l'ensemble admissible du problème jusqu'à avoir trouvé le sommet où est réalisée la solution optimale.

Méthode du simplexe - présentation sans tableaux

Le problème à résoudre est :

$$(P) \left\{ \begin{array}{l} \text{maximiser } z = 15x_1 + 17x_2 \\ \text{s.l.c.} \\ x_1 + 3x_2 \leq 300 \\ 4x_1 + 2x_2 \leq 500 \\ x_1 + x_2 \leq 150 \\ x_1, x_2 \geq 0 \end{array} \right. .$$

On n'est pas habitué à résoudre des systèmes d'inéquations ; on se ramène à un système d'équations en écrivant le problème standard associé :

$$(P_S) \left\{ \begin{array}{l} \text{maximiser } z = 15x_1 + 17x_2 \\ \text{s.l.c.} \\ x_1 + 3x_2 + e_1 = 300 \\ 4x_1 + 2x_2 + e_2 = 500 \\ x_1 + x_2 + e_3 = 150 \\ x_1, x_2, e_1, e_2, e_3 \geq 0 \end{array} \right. .$$

On peut réécrire ceci sous la forme :

$$\begin{array}{l} \text{maximiser } z \text{ avec } x_1, x_2, e_1, e_2, e_3 \geq 0 \text{ vérifiant} \\ \left\{ \begin{array}{l} x_1 + 3x_2 + e_1 = 300 \\ 4x_1 + 2x_2 + e_2 = 500 \\ x_1 + x_2 + e_3 = 150 \\ z = 15x_1 + 17x_2 \end{array} \right. . \end{array} \quad (1.4.1)$$

Dans le système précédent nous avons 6 variables x_1, x_2, e_1, e_2, e_3 et z mais seulement 4 équations. Le système est sous-déterminé et admet une infinité de solutions. Pour déterminer une solution particulière du système, on n'a pas d'autre choix que de fixer les valeurs de $6 - 4 = 2$ des variables et d'exprimer les autres en fonction de ces variables. En fait, on fixe à 0 les valeurs des variables hors base. On commence par choisir pour variables hors base initiales x_1 et x_2 . Ce choix correspond au fait que l'algorithme démarre à l'origine. Ce point

est un sommet de l'ensemble admissible dès lors que toutes les contraintes sont du type « \leq » avec des membres de droite strictement positifs.

Le système (1.4.1) permet d'exprimer e_1, e_2, e_3 et z en fonction de x_1 et x_2 comme suit :

$$\begin{aligned} & \text{maximiser } z \text{ avec } x_1, x_2, e_1, e_2, e_3 \geq 0 \text{ vérifiant} \\ & \begin{cases} e_1 = 300 - x_1 - 3x_2 \\ e_2 = 500 - 4x_1 - 2x_2 \\ e_3 = 150 - x_1 - x_2 \\ z = 15x_1 + 17x_2 \end{cases} . \end{aligned} \quad (1.4.2)$$

Puisque l'on a choisi de fixer $x_1 = x_2 = 0$, on obtient $e_1 = 300$, $e_2 = 500$, $e_3 = 150$ et $z = 0$. Notre but est de maximiser z . La dernière ligne du système précédant nous indique que si l'on augmente x_1 de 1, la valeur de z augmentera de 15 et si l'on augmente x_2 de 1, la valeur de z augmentera de 17. On a clairement intérêt à augmenter la valeur de x_2 plutôt que celle de x_1 pour augmenter le plus possible la valeur de z .

Il faut maintenant voir dans quelle mesure on peut augmenter x_2 en faisant sortir une des variables e_1, e_2 ou e_3 de la base pour continuer à respecter les contraintes. La première ligne nous indique que si l'on fait sortir e_1 de la base, on peut augmenter x_2 jusqu'à $x_2 = \frac{1}{3}(300 - x_1 - e_1) = 100$ avec $e_1 = x_1 = 0$ puisque ces variables seraient hors base. Si l'on augmentait x_2 de façon plus importante, il faudrait prendre e_1 négatif ce qui n'est pas possible. La deuxième ligne nous indique que si l'on fait sortir e_2 de la base, on peut augmenter x_2 jusqu'à $x_2 = \frac{1}{2}(500 - 4x_1 - e_2) = 250$ avec $e_2 = x_1 = 0$ puisque ces variables seraient hors base. Si l'on augmentait x_2 de façon plus importante, il faudrait prendre e_2 négatif ce qui n'est pas possible. La troisième ligne nous indique que si l'on fait sortir e_3 de la base, on peut augmenter x_2 jusqu'à $x_2 = 150 - x_1 - e_3 = 150$ avec $e_3 = x_1 = 0$ puisque ces variables seraient hors base. Si l'on augmentait x_2 de façon plus importante, il faudrait prendre e_3 négatif ce qui n'est pas possible. Ainsi, le meilleur choix possible est de faire sortir e_1 de la base, ce qui conduira à $x_2 = 100$. Le choix d'une valeur supérieure à 100 pour x_2 et le respect des trois premières égalités du système contredirait au moins la positivité de e_1 . Les nouvelles variables hors base étant x_1 et e_1 , on exprime les variables de base x_2, e_2, e_3 ainsi que z en fonction de x_1 et e_1 en partant de (1.4.2)

$$\begin{aligned} & \text{maximiser } z \text{ avec } x_1, x_2, e_1, e_2, e_3 \geq 0 \text{ vérifiant} \\ & \begin{cases} e_1 = 300 - x_1 - 3x_2 \\ e_2 = 500 - 4x_1 - 2x_2 \\ e_3 = 150 - x_1 - x_2 \\ z = 15x_1 + 17x_2 \end{cases} \\ & \text{maximiser } z \text{ avec } x_1, x_2, e_1, e_2, e_3 \geq 0 \text{ vérifiant} \\ \iff & \begin{cases} x_2 = 100 - \frac{1}{3}x_1 - \frac{1}{3}e_1 \\ e_2 = 300 - \frac{10}{3}x_1 + \frac{2}{3}e_1 \\ e_3 = 50 - \frac{2}{3}x_1 + \frac{1}{3}e_1 \\ z = 1700 + \frac{28}{3}x_1 - \frac{17}{3}e_1 \end{cases} . \end{aligned} \quad (1.4.3)$$

Puisque x_1 et e_1 sont hors base, $x_1 = e_1 = 0$ et on obtient que $x_2 = 100$, $e_2 = 300$, $e_3 = 50$ et $z = 1700$. On souhaite encore augmenter la valeur de z si c'est possible. La dernière ligne du système précédant nous indique que si on augmente x_1 de 1, la valeur de z augmentera de $\frac{28}{3}$ et si on augmente e_1 de 1, la valeur de z diminuera de $\frac{17}{3}$. Ainsi, seule l'augmentation de la valeur de x_1 permet d'améliorer la valeur de z . On choisit de faire entrer x_1 dans la base.

Il faut maintenant voir dans quelle mesure on peut augmenter x_1 en faisant sortir une des variables x_2 , e_2 ou e_3 de la base pour continuer à respecter les contraintes. La première ligne nous indique que si l'on fait sortir x_2 de la base, on peut augmenter x_1 jusqu'à $x_1 = 3(100 - x_2 - \frac{1}{3}e_1) = 300$ avec $e_1 = x_2 = 0$ puisque ces variables seraient hors base. Si l'on augmentait x_1 de façon plus importante, il faudrait prendre x_2 négatif ce qui n'est pas possible. La deuxième ligne nous indique que si l'on fait sortir e_2 de la base, on peut augmenter x_1 jusqu'à $x_1 = \frac{3}{10}(300 + \frac{2}{3}x_2 - e_2) = 90$ avec $e_1 = e_2 = 0$ puisque ces variables seraient hors base. Si l'on augmentait x_1 de façon plus importante, il faudrait prendre e_2 négatif ce qui n'est pas possible. La troisième ligne nous indique que si l'on fait sortir e_3 de la base, on peut augmenter x_1 jusqu'à $x_1 = \frac{3}{2}(50 + \frac{1}{3}e_1 - e_3) = 75$ avec $e_1 = e_3 = 0$ puisque ces variables seraient hors base. Si l'on augmentait x_1 de façon plus importante, il faudrait prendre e_3 négatif ce qui n'est pas possible. Ainsi, le meilleur choix possible est de faire sortir e_3 de la base, ce qui conduira à $x_1 = 75$. Le choix d'une valeur supérieure à 75 pour x_1 et le respect des trois premières égalités du système contredirait à la positivité de e_3 . Les nouvelles variables hors base étant e_1 et e_3 , on exprime les variables de base x_1 , e_2 , x_2 ainsi que z en fonction de e_1 et e_3 en partant de (1.4.3) :

$$\begin{aligned} & \text{maximiser } z \text{ avec } x_1, x_2, e_1, e_2, e_3 \geq 0 \text{ vérifiant} \\ & \begin{cases} x_2 = 100 - \frac{1}{3}x_1 - \frac{1}{3}e_1 \\ e_2 = 300 - \frac{10}{3}x_1 + \frac{2}{3}e_1 \\ e_3 = 50 - \frac{2}{3}x_1 + \frac{1}{3}e_1 \\ z = 1700 + \frac{28}{3}x_1 - \frac{17}{3}e_1 \end{cases} \\ & \text{maximiser } z \text{ avec } x_1, x_2, e_1, e_2, e_3 \geq 0 \text{ vérifiant} \\ \iff & \begin{cases} x_2 = 75 - \frac{1}{2}e_1 + \frac{1}{2}e_3 \\ e_2 = 50 - e_1 + 5e_3 \\ x_1 = 75 + \frac{1}{2}e_1 - \frac{3}{2}e_3 \\ z = 2400 - e_1 - 14e_3 \end{cases} \end{aligned} \tag{1.4.4}$$

Puisque e_1 et e_3 sont hors base, $e_1 = e_3 = 0$ et donc $x_1 = x_2 = 75$, $e_2 = 50$ et $z = 2400$. Dans la dernière ligne, on voit que le fait d'augmenter e_1 de 1 conduit à une perte de 1 pour z et le fait d'augmenter e_3 de 1 conduit à une perte de 3 sur z . Il n'y a donc plus de moyen pour augmenter la valeur de z . On s'arrête et on conclue que la solution optimale est $z^* = 2400$ atteinte en $(x_1^*; x_2^*) = (75; 75)$.

Méthode du simplexe - présentation par tableaux

Nous avons déjà écrit le problème sous forme standard ; le tableau initial est le suivant et les variables de base sont e_1 , e_2 et e_3 .

	v.	x_1	x_2	e_1	e_2	e_3	$-z$	b_i
v.b.		x_1	x_2	e_1	e_2	e_3	$-z$	b_i
e_1		1	3	1	0	0	0	300
e_2		4	2	0	1	0	0	500
e_3		1	1	0	0	1	0	150
$-z$		15	17	0	0	0	1	0

La première ligne représente l'équation :

$$1 \times x_1 + 3 \times x_2 + 1 \times e_1 + 0 \times e_2 + 0 \times e_3 + 0 \times (-z) = 300,$$

soit

$$x_1 + 3x_2 + e_1 = 300.$$

De même, les deuxième et troisième lignes représentent respectivement les équations :

$$4x_1 + 2x_2 + e_2 = 500 \quad \text{et} \quad x_1 + x_2 + e_3 = 150.$$

La quatrième ligne représente l'équation :

$$15x_1 + 17x_2 - z = 0$$

c'est-à-dire l'objectif à maximiser : $z = 15x_1 + 17x_2$.

Ce tableaux correspond à (1.4.2). Puisque $x_1 = x_2 = 0$, on a $e_1 = 300$, $e_2 = 500$, $e_3 = 150$ et $z = 0$, ce qui se lit directement dans le tableau. Les variables de base sont e_1 , e_2 et e_3 , les contraintes saturées sont $x_1 \geq 0$ et $x_2 \geq 0$ et l'algorithme démarre du sommet $S_0 = (0; 0)$.

Le coefficient le plus grand parmi les positifs dans la ligne de $-z$ (colonne de $-z$ exceptée) est 17 et correspond à la variable x_2 , on décide donc de faire rentrer x_2 dans la base.

Il faut maintenant faire sortir une variable de la base pour y laisser une place à x_2 . Pour cela, on calcule, dans chaque ligne correspondant à une variable de base, le quotient du coefficient de la colonne des b_i sur le coefficient de la colonne de la variable entrante x_2 .

v. \ v.b.	x_1	x_2	e_1	e_2	e_3	$-z$	b_i	quotients
e_1	1	3	1	0	0	0	300	$\frac{300}{3} = 100$
e_2	4	2	0	1	0	0	500	$\frac{500}{2} = 250$
e_3	1	1	0	0	1	0	150	$\frac{150}{1} = 150$
$-z$	15	17	0	0	0	1	0	

Le quotient le plus petit (parmi les positifs) est 100 et correspond à la variable e_1 que l'on fait sortir de la base et que l'on remplace par x_2 .

On repère, ensuite, le pivot situé dans la ligne de la variable sortante et la colonne de la variable entrante et on effectue une itération de la méthode de Gauss-Jordan pour mettre à 1 la valeur du pivot et à 0 les autres coefficients de la colonne de la variable entrante. En partant du tableau :

v. \ v.b.	x_1	x_2	e_1	e_2	e_3	$-z$	b_i
\leftarrow e_1 x_2	1	3	1	0	0	0	300
e_2	4	2	0	1	0	0	500
e_3	1	1	0	0	1	0	150
$-z$	15	17	0	0	0	1	0

↑

on effectue l'opération $L_1 \leftarrow L_1/3$ pour obtenir :

v.b. \ v.	x_1	x_2	e_1	e_2	e_3	$-z$	b_i
x_2	$\frac{1}{3}$	1	$\frac{1}{3}$	0	0	0	100
e_2	4	2	0	1	0	0	500
e_3	1	1	0	0	1	0	150
$-z$	15	17	0	0	0	1	0

puis les opérations $L_2 \leftarrow L_2 - 2L_1$, $L_3 \leftarrow L_2 - L_1$ et $L_4 \leftarrow L_4 - 17L_1$, pour obtenir :

v.b. \ v.	x_1	x_2	e_1	e_2	e_3	$-z$	b_i
x_2	$\frac{1}{3}$	1	$\frac{1}{3}$	0	0	0	100
e_2	$\frac{10}{3}$	0	$-\frac{2}{3}$	1	0	0	300
e_3	$\frac{2}{3}$	0	$-\frac{1}{3}$	0	1	0	50
$-z$	$\frac{28}{3}$	0	$-\frac{17}{3}$	0	0	1	-1700

On peut lire dans ce tableau (correspondant à (1.4.3)) que lorsque $x_1 = 0$, $x_2 = 100$, $e_1 = 0$, $e_2 = 300$ et $e_3 = 50$, la valeur de la fonction objectif z est 1700 ($-z = -1700$).

Dans le tableau :

v.b. \ v.	x_1	x_2	e_1	e_2	e_3	$-z$	b_i
x_2	$\frac{1}{3}$	1	$\frac{1}{3}$	0	0	0	100
e_2	$\frac{10}{3}$	0	$-\frac{2}{3}$	1	0	0	300
e_3	$\frac{2}{3}$	0	$-\frac{1}{3}$	0	1	0	50
$-z$	$\frac{28}{3}$	0	$-\frac{17}{3}$	0	0	1	-1700

nous pouvons voir qu'il reste un coefficient strictement positif dans la ligne de $-z$. Celui-ci est $\frac{28}{3}$ et correspond à la variable x_1 que l'on va faire entrer dans la base.

Pour déterminer la variable sortante, on calcule, dans chaque ligne correspondant à une variable de base, le quotient du coefficient de la colonne des b_i sur le coefficient de la colonne de la variable entrante x_1 .

v.b. \ v.	x_1	x_2	e_1	e_2	e_3	$-z$	b_i	quotients
x_2	$\frac{1}{3}$	1	$\frac{1}{3}$	0	0	0	100	$\frac{100}{\frac{1}{3}} = 300$
e_2	$\frac{10}{3}$	0	$-\frac{2}{3}$	1	0	0	300	$\frac{300}{\frac{10}{3}} = 90$
e_3	$\frac{2}{3}$	0	$-\frac{1}{3}$	0	1	0	50	$\frac{50}{\frac{2}{3}} = 75$
$-z$	$\frac{28}{3}$	0	$-\frac{17}{3}$	0	0	1	-1700	

Le quotient le plus faible (parmi les positifs) est 75 et correspond à la variable e_3 que l'on fait sortir de la base et que l'on remplace par x_1 .

On repère, ensuite, le pivot situé dans la ligne de la variable sortante et la colonne de la variable entrante et on effectue une itération de la méthode de Gauss-Jordan pour mettre à 1 la valeur du pivot et à 0 les autres coefficients de la colonne de la variable entrante. En partant du tableau :

v.b. \ v.	x_1	x_2	e_1	e_2	e_3	$-z$	b_i
x_2	$\frac{1}{3}$	1	$\frac{1}{3}$	0	0	0	100
e_2	$\frac{10}{3}$	0	$-\frac{2}{3}$	1	0	0	300
x_1	$\frac{2}{3}$	0	$-\frac{1}{3}$	0	1	0	50
$-z$	$\frac{28}{3}$	0	$-\frac{17}{3}$	0	0	1	-1700

↑

on effectue l'opération $L_3 \leftarrow \frac{3}{2}L_3$, pour obtenir :

v.b. \ v.	x_1	x_2	e_1	e_2	e_3	$-z$	b_i
x_2	$\frac{1}{3}$	1	$\frac{1}{3}$	0	0	0	100
e_2	$\frac{10}{3}$	0	$-\frac{2}{3}$	1	0	0	300
x_1	1	0	$-\frac{1}{2}$	0	$\frac{3}{2}$	0	75
$-z$	$\frac{28}{3}$	0	$-\frac{17}{3}$	0	0	1	-1700

puis les opérations $L_1 \leftarrow L_1 - \frac{1}{3}L_3$, $L_2 \leftarrow L_2 - \frac{10}{3}L_3$ et $L_4 \leftarrow L_4 - \frac{28}{3}L_3$, pour obtenir :

v.b. \ v.	x_1	x_2	e_1	e_2	e_3	$-z$	b_i
x_2	0	1	$\frac{1}{2}$	0	$-\frac{1}{2}$	0	75
e_2	0	0	1	1	-5	0	50
x_1	1	0	$-\frac{1}{2}$	0	$\frac{3}{2}$	0	75
$-z$	0	0	-1	0	-14	1	-2400

L'effet de cette étape est exprimer x_2 , e_2 , x_1 et $-z$ en fonction des variables hors base e_1 et e_3 (ceci correspond à 1.4.4).

Dans le dernier tableau, la ligne de $-z$ (colonne de $-z$ exceptée) ne comporte pas de coefficient strictement positif. L'algorithme s'arrête et on peut lire dans le tableau que la valeur optimale est $z = 2400$ ($-z = -2400$) atteinte lorsque $x_1 = x_2 = 75$ ($e_1 = e_3 = 0$ et $e_2 = 50$). On retrouve le résultat obtenu plus haut par la méthode graphique.

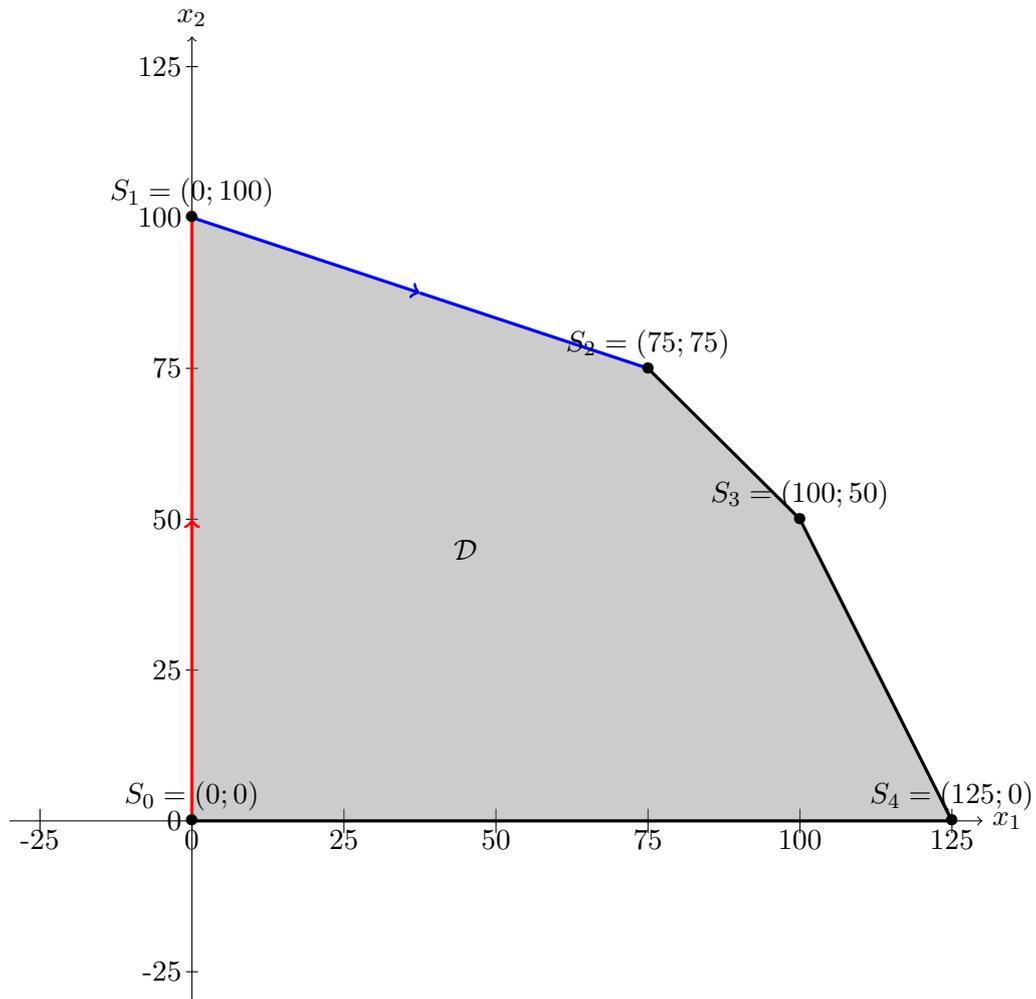


FIGURE 1.3 – Pour résoudre le problème d’optimisation linéaire de l’Exemple 1.11, l’algorithme du simplexe démarre du sommet S_0 . Lors de la première itération (rouge), il décide d’explorer le sommet S_1 puis, lors de la deuxième itération (bleu), le sommet S_2 . L’algorithme s’arrête ensuite puisqu’il est arrivé sur le sommet réalisant le maximum de la fonction objectif sur le domaine \mathcal{D} . Les sommets S_2 , S_3 et S_4 ne sont pas visités par l’algorithme.

Exemple 1.12. Résolvons par la méthode du simplexe le problème :

$$(P) \left\{ \begin{array}{l} \text{minimiser } z = 2x_1 - 3x_2 - x_3 \\ \text{s.l.c.} \\ x_1 + x_2 + 2x_3 \leq 100 \\ 3x_1 - 2x_2 \leq 50 \\ x_1 - 4x_3 \leq 20 \\ x_1, x_2, x_3 \geq 0 \end{array} \right. .$$

Le problème standard associé à (P) :

$$(P_S) \begin{cases} \text{minimiser} & z = 2x_1 - 3x_2 - x_3 \\ \text{s.l.c.} & \\ & x_1 + x_2 + 2x_3 + e_1 = 100 \\ & 3x_1 - 2x_2 + e_2 = 50 \\ & x_1 - 4x_3 + e_3 = 20 \\ & x_1, x_2, x_3, e_1, e_2, e_3 \geq 0 \end{cases}$$

L'algorithme du simplexe démarre ici avec pour variables de base e_1, e_2 et e_3 et le tableau :

v.b. \ v.	x_1	x_2	x_3	e_1	e_2	e_3	$-z$	b_i
e_1	1	1	2	1	0	0	0	100
e_2	3	-2	0	0	1	0	0	50
e_3	1	0	-4	0	0	1	0	20
$-z$	2	<u>-3</u>	-1	0	0	0	1	0

Il s'agit d'un problème de **minimisation** et le coefficient le **plus grand en valeur absolue parmi les négatifs** dans la ligne de $-z$ (colonne de $-z$ exceptée) est -3 et correspond à la variable x_2 que l'on va faire rentrer dans la base. Pour déterminer la variable sortante, on calcule les quotients des coefficients de la colonne des b_i sur ceux de la colonne de la variable entrante x_2 .

v.b. \ v.	x_1	x_2	x_3	e_1	e_2	e_3	$-z$	b_i	quotients
e_1 x_2	1	1	2	1	0	0	0	100	$\frac{100}{1} = 100$
e_2	3	-2	0	0	1	0	0	50	$\frac{50}{-2} = -25$
e_3	1	0	-4	0	0	1	0	20	∞
$-z$	2	-3	-1	0	0	0	1	0	

Le quotient le plus petit parmi les positifs est 100 et correspond à la variable e_1 que l'on fait sortir de la base et que l'on remplace par x_2 . Le pivot est **2** les opérations $L_2 \leftarrow L_2 + 2L_1$ et $L_4 \leftarrow L_4 + 3L_1$ conduisent à :

v.b. \ v.	x_1	x_2	x_3	e_1	e_2	e_3	$-z$	b_i
x_2	1	1	2	1	0	0	0	100
e_2	5	0	4	2	1	0	0	250
e_3	1	0	-4	0	0	1	0	20
$-z$	5	0	5	3	0	0	1	300

Il ne reste plus de coefficient strictement négatif dans la ligne de $-z$ (colonne de $-z$ exceptée). Puisqu'il s'agit d'un problème de minimisation, l'algorithme s'arrête. La solution optimale est donc $z^* = -300$ et est atteinte en $(x_1^*; x_2^*; x_3^*) = (0; 100; 0)$.

Chapitre 2

Statistiques descriptives univariées

La *statistique descriptive* regroupe un panel de méthodes permettant de représenter et synthétiser des données collectées. Ces données sont obtenues soit sur une population entière soit sur un échantillon choisi au hasard dans cette population. Une fois les données collectées, le but est de fournir une visualisation ou une description simple d'un phénomène à l'aide d'un nombre limité de valeurs. On a pour cela recours à des représentations graphiques et à des caractéristiques de la série de données appelées *paramètres statistiques*.

S'il est nécessaire de préciser statistiques *descriptives*, c'est qu'il existe d'autres méthodes statistiques dont les objectifs sont différents. Celles-ci sont regroupées sous la terminologie *statistiques inférentielles* et ont pour objet l'aide à la décision concernant des caractéristiques de (très) grandes populations en ne se basant que sur des observations obtenues sur un échantillon, généralement petit, de la population.

Les statistiques inférentielles s'appuient de manière intensive sur la théorie des probabilités, ce qui les rend délicates à appréhender. Au contraire, les statistiques descriptives (*a fortiori* dans le cas d'une seule dimension) ne demandent pas de prérequis particulier et sont l'objet de ce chapitre.

2.1 Vocabulaire statistique

Définition 2.1. *On appelle :*

1. population tout ensemble étudié par la statistique ;
2. individu tout élément de la population ;
3. effectif total le nombre d'individus dans la population.

Notations 2.1. *On note \mathcal{P} la population et N l'effectif total de cette population.*

Exemple 2.1.

1. Si la population \mathcal{P} est l'ensemble des étudiants de la promotion, un individu est un étudiant de la promotion.
2. Si la population \mathcal{P} est l'ensemble des jours de septembre 2016, chaque jour de ce mois est un individu et l'effectif total est $N = 30$.

Définition 2.2. Soit \mathcal{P} une population.

On appelle variable statistique (ou caractère) une quantité ou qualité définie sur \mathcal{P} susceptible de varier d'un individu à l'autre. On appelle modalités les différentes valeurs ou aspects pris par cette variable.

On distingue :

1. les variables qualitatives pour lesquelles les modalités ne sont pas objectivement comparables ;
2. les variables quantitatives (ou ordinales) dont les modalités sont mesurables et comparables deux à deux.

Parmi les variables quantitatives, on distingue :

1. les variables quantitatives discrètes dont les valeurs possibles sont isolées ;
2. les variables quantitatives continues pouvant prendre toutes les valeurs contenues dans un intervalle.

Notation 2.2. Les variables statistiques sont désignées par des lettres majuscules, généralement X ou Y .

Exemple 2.1 (suite).

1. Si la population \mathcal{P} est l'ensemble des étudiants de la promotion, une variable statistique peut être la couleur des yeux (variable qualitative) ou son âge (variable quantitative discrète).
2. Si la population \mathcal{P} est l'ensemble des jours de septembre 2016, une variable statistique peut être la hauteur totale des précipitations journalières relevées à Dijon (variable quantitative continue).

Dans les trois sections suivantes, on présente les outils et méthodes nécessaires à la réalisation d'une étude statistique (descriptive) en distinguant les cas :

- qualitatif,
- quantitatif discret sans regroupement en classes,
- quantitatif continu ou discret avec regroupement en classes.

2.2 Cas qualitatif

Après le dépouillement d'une étude, on synthétise généralement les résultats sous la forme d'un tableau faisant apparaître les modalités x_1, x_2, \dots, x_r et les effectifs de ces modalités n_1, n_2, \dots, n_r .

Modalités	x_1	x_2	\dots	x_i	\dots	x_r
Effectifs	n_1	n_2	\dots	n_i	\dots	n_r

Ici, r est le nombre de modalités (valeurs) prises par la variable statistique étudiée X , n_i représentent le nombre d'individus (effectif) pour lesquels la variable statistique prend la modalité x_i . L'effectif total est :

$$N = n_1 + n_2 + \dots + n_r,$$

ce qui se réécrit de façon compacte sous la forme :

$$N = \sum_{i=1}^r n_i.$$

Remarque 2.1. La notation $\sum_{i=1}^r n_i$ ci-dessus, que l'on lit « somme des n_i pour i allant de 1 à n », présente deux avantages principaux dans notre contexte. Elle est compacte (courte à écrire) et induit une façon simple et efficace de saisir un calcul dans Excel en utilisant la fonction `SOMME(...)` et en sélectionnant la plage de données à sommer.

Définition 2.3. Soit X une variable statistique de modalités x_1, x_2, \dots, x_r ayant pour effectifs n_1, n_2, \dots, n_r dans une population \mathcal{P} d'effectif total N . La fréquence f_i de la modalité x_i est définie par :

$$f_i = \frac{n_i}{N}.$$

Remarque 2.2.

1. On a :

$$\sum_{i=1}^r f_i = f_1 + \dots + f_r = 1 (= 100\%).$$

2. L'utilisation des fréquences présente l'avantage de pouvoir comparer facilement les résultats d'études portant sur le même caractère dans deux populations de tailles différentes.
3. Dans Excel, le symbole \$ permet de « bloquer » une ligne (devant la lettre) ou une colonne (devant le nombre). Ainsi, il est particulièrement utile pour étendre une formule comme celle du calcul des fréquences : on laisse libre les lettres et nombre repérant n_i mais on bloque ceux repérant N avant d'étendre la formule que l'on n'aura pas besoin de réécrire d'une ligne à l'autre (voir aussi fichier « Aid mémoire Excel »).

Les représentations graphiques portent généralement sur les fréquences et rarement sur les effectifs. Dans le cas de variables qualitatives, on représente ces fréquences sous forme de :

- *diagramme en bâtons* : la hauteur du bâton associé à une modalité est simplement sa fréquence ;
- *diagramme circulaire* : l'angle du secteur associé à la modalité x_i est $\theta_i = f_i \times 360^\circ$.

Exemple 2.2. Une étude a porté sur l'ensemble des 150 étudiants d'une promotion. Chacun devait répondre à la question « *Reprenez-vous le contenu de vos cours le soir ?* » par « *Jamais* », « *Rarement* », « *Souvent* » ou « *Toujours* ». Les résultats ont été les suivants :

Modalités (réponses) x_i	Jamais	Rarement	Souvent	Toujours
Effectifs n_i	9	60	66	15

L'effectif total est, comme annoncé, $N = 9 + 60 + 66 + 15 = 150$. On peut dresser le tableau suivant contenant les fréquences de chaque modalité.

Modalités (réponses) x_i	Jamais	Rarement	Souvent	Toujours	Total
Effectifs n_i	9	60	66	15	150
Fréquence $f_i = \frac{n_i}{N}$	0,06	0,4	0,44	0,1	1
Fréquence $f_i = \frac{n_i}{N}$ (en %)	6	40	44	10	100

Les Figures 2.1 et 2.2 fournissent des représentations graphiques de ces données.

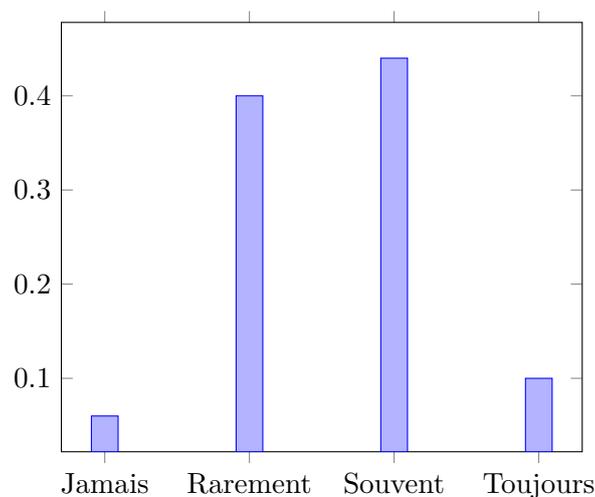


FIGURE 2.1 – Représentation des fréquences par un diagramme en bâtons (Exemple 2.2).

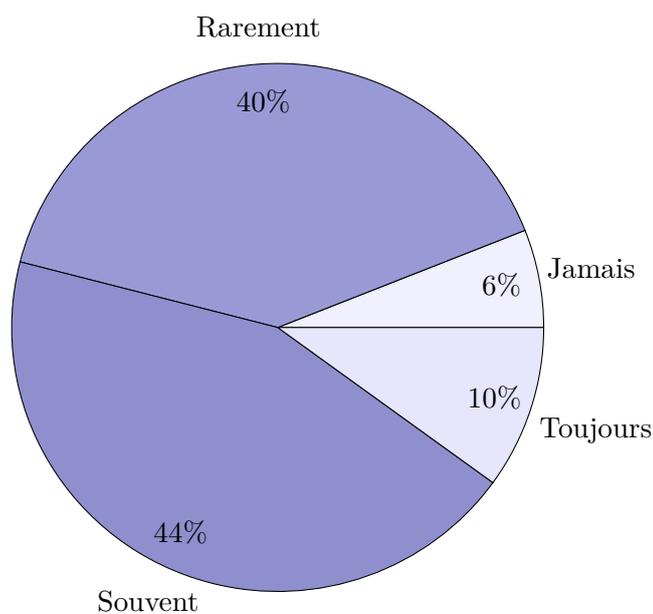


FIGURE 2.2 – Représentation des fréquences par un diagramme circulaire (Exemple 2.2).

2.3 Cas quantitatif discret sans regroupement en classes

Dans le cas où les modalités d'une variable statistique sont des valeurs numériques isolées (discrètes) et ne sont pas trop nombreuses, on peut effectuer un traitement statistique des données sans regroupement en classes. Dans toute cette section, on suppose que l'on étudie une variable statistique quantitative discrète X sur une population \mathcal{P} d'effectif total N . On suppose également que les modalités x_1, x_2, \dots, x_r sont ordonnées par ordre croissant ($x_1 < x_2 < \dots < x_r$) et que n_1, n_2, \dots, n_r sont les effectifs associés à ces modalités.

2.3.1 Tableau statistique

Outre les modalités, effectifs et fréquences (définies dans la Définition 2.3), le tableau statistique d'une telle série contient les *effectifs cumulés croissants* et les *fréquences cumulées croissantes*.

Définition 2.4. *Le i^e effectif cumulé croissant est $N_i = n_1 + \dots + n_i = \sum_{k=1}^i n_k$ et la i^e fréquence cumulée croissante est $F_i = f_1 + \dots + f_i = \sum_{k=1}^i f_k$.*

Remarque 2.3.

1. Les quantités N_i et F_i représentent l'effectif (resp. la fréquence) des individus de la population pour lesquels la variable statistique prend une valeur inférieure ou égale à la modalité x_i .
2. On a :

$$N_r = N, \quad F_r = 1 = 100\% \quad \text{et} \quad F_i = \frac{n_1 + \dots + n_i}{N} = \frac{N_i}{N}.$$

Notation 2.3. *On note $\mathbf{P}[X < t]$ la fréquence totale des modalités x_i telles que $x_i < t$. On peut définir de manière analogue $\mathbf{P}[X \leq t]$, $\mathbf{P}[X > t]$, $\mathbf{P}[X \geq t]$, $\mathbf{P}[t_1 < X < t_2]$, ...*

Définition 2.5. *On appelle fonction de répartition de la variable X la fonction définie sur \mathbf{R} et à valeurs dans $[0, 1]$:*

$$F : t \mapsto F(t) = \mathbf{P}[X \leq t].$$

Proposition 2.1. *Soit X une variable statistique discrète dont les modalités sont x_1, \dots, x_r .*

1. *La fonction de répartition F de X est croissante.*
2. *On a :*

$$F(t) = \begin{cases} 0 & \text{si } t < x_1 \\ F_i & \text{si } x_i \leq t < x_{i+1}, i = 1, \dots, r-1 \\ 1 & \text{si } t \geq x_r \end{cases}.$$

En particulier la fonction de répartition d'une variable discrète est une fonction en escalier.

3. *On a :*

$$\begin{aligned} \mathbf{P}[X > t] &= 1 - \mathbf{P}[X \leq t] = 1 - F(t), \\ \mathbf{P}[t_1 < X \leq t_2] &= \mathbf{P}[X \leq t_2] - \mathbf{P}[X \leq t_1] = F(t_2) - F(t_1), \\ \mathbf{P}[X = x_i] &= f_i \quad \text{et} \quad \mathbf{P}[X = t] = 0 \quad \text{si } t \notin \{x_1, \dots, x_r\}, \\ \mathbf{P}[X < t] &= \mathbf{P}[X \leq t] - \mathbf{P}[X = t] = \begin{cases} F(x_i) - f(x_i) = F(x_{i-1}) & \text{si } t = x_i \\ F(t) & \text{si } t \notin \{x_1, \dots, x_r\} \end{cases}. \end{aligned}$$

Exemple 2.3. Une enquête réalisée auprès d'une population de 100 femmes de 40 ans a recensé le nombre d'enfant de chacune. Les résultats sont consignés dans le tableau suivant.

Nombre d'enfant(s) x_i	0	1	2	3	4	5	6
Nombre de femmes n_i	9	28	32	24	4	2	1

La population \mathcal{P} est le contingent de femmes interrogées. L'effectif total est $N = 100$. La variable statistique étudiée X est le nombre d'enfant par femme ; il s'agit d'une variable quantitative discrète. Les modalités sont $x_1 = 0, x_2 = 1, \dots, x_7 = 6$ et ont pour effectifs respectifs $n_1 = 9, n_2 = 28, \dots, n_7 = 1$. Le tableau statistique complet de cette série prend la forme suivante.

Modalités x_i	0	1	2	3	4	5	6
Effectifs n_i	9	28	32	24	4	2	1
Fréquences $f_i = \frac{n_i}{N}$	0,09	0,28	0,32	0,24	0,04	0,02	0,01
Effectifs cumulés croissants $N_i = n_1 + \dots + n_i$	9	37	69	93	97	99	100
Fréquences cumulées croissantes $F_i = \frac{N_i}{N}$	0,09	0,37	0,69	0,93	0,97	0,99	1

On peut, par exemple, s'intéresser à la proportion de femmes ayant au plus 2 enfants. Celle-ci est :

$$\mathbf{P}[X \leq 2] = F(2) = F_3 = 0,69 = 69\%.$$

La proportion de femmes ayant (strictement) plus de 3 enfants est quant à elle :

$$\mathbf{P}[X > 3] = 1 - \mathbf{P}[X \leq 3] = 1 - F(3) = 1 - F_4 = 1 - 0,93 = 0,07 = 7\%,$$

alors que la proportion de femmes ayant entre 1 et 3 enfants est :

$$\begin{aligned} \mathbf{P}[1 \leq X \leq 3] &= \mathbf{P}[X \leq 3] - \mathbf{P}[X < 1] = F(3) - \mathbf{P}[X \leq 0] \\ &= F(3) - F(0) = F_4 - F_1 = 0,93 - 0,09 = 0,84 = 84\%. \end{aligned}$$

2.3.2 Représentations graphiques

Histogramme des fréquences

Comme dans le cas qualitatif, l'histogramme des fréquences prend la forme d'un diagramme en bâtons (la hauteur du bâton associé à une modalité est simplement sa fréquence). On peut commenter à l'aide de cet histogramme la symétrie ou dissymétrie (traîne étalée vers la droite ou la gauche) de la distribution.

La Figure 2.3 représente l'histogramme des fréquences obtenu dans l'Exemple 2.3. On observe dans ce cas une *traîne* (ou *queue de distribution*) étalée vers la droite.

Fonction de répartition

Comme il a été annoncé dans la section précédente, la fonction de répartition F d'une variable discrète est une fonction en escalier. Cette fonction est nulle pour des valeurs de la variable t inférieures à la plus petite modalité x_1 , vaut F_i sur tout intervalle de la forme $[x_i, x_{i+1}[$, $i \in \{1, \dots, r-1\}$ et vaut 1 sur l'intervalle $[x_r, +\infty[$.

La Figure 2.4 représente la fonction de répartition obtenue dans l'Exemple 2.3.

2.3.3 Paramètres statistiques

Dans toute cette section, on considère une série statistique quantitative discrète dont les modalités sont x_1, x_2, \dots, x_r d'effectifs respectifs n_1, n_2, \dots, n_r . On note N l'effectif total de la population. On illustrera les notions introduites à l'aide de l'Exemple 2.3.

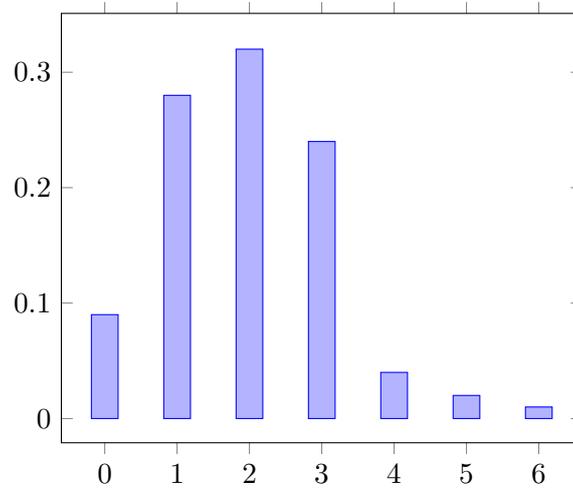


FIGURE 2.3 – Représentation des fréquences par un diagramme en bâtons (Exemple 2.3).

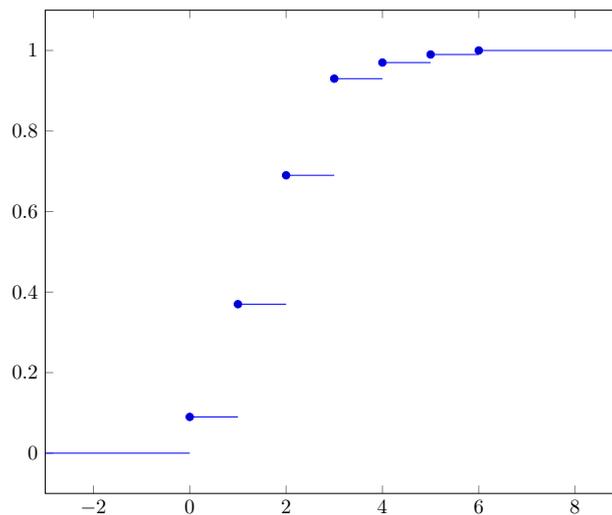


FIGURE 2.4 – Fonction de répartition d'un variable discrète (Exemple 2.3).

Paramètres de position

Le paramètre statistique le plus connu est certainement la *moyenne arithmétique* (empirique). Celle-ci représente la valeur qu'obtiendrait chaque individu de la population si ceux-ci se répartissaient les ressources équitablement.

Définition 2.6. La moyenne arithmétique \bar{X} (ou simplement moyenne) d'une série statistique est définie par :

$$\bar{X} = \frac{1}{N} \sum_{i=1}^r n_i x_i = \frac{n_1 x_1 + \dots + n_r x_r}{N},$$

où x_1, x_2, \dots, x_r sont les modalités, n_1, n_2, \dots, n_r leurs effectifs respectifs et N l'effectif total de la population.

En combinant cette définition et la Définition 2.3, on obtient immédiatement une réécriture de la moyenne.

Proposition 2.2.

$$\bar{X} = \sum_{i=1}^r f_i x_i.$$

Remarque 2.4. La moyenne est très sensible aux valeurs extrêmes (grandes dans les positifs ou les négatifs).

Exemple 2.3 (suite). On a, avec les données de l'Exemple 2.3 :

$$\bar{X} = \frac{9 \times 0 + 28 \times 1 + 32 \times 2 + 24 \times 3 + 4 \times 4 + 2 \times 5 + 1 \times 6}{100} = 1,96.$$

Un deuxième paramètre de position est le *mode*.

Définition 2.7. On appelle mode d'une série statistique discrète toute modalité x_i dont l'effectif est maximal parmi tous les effectifs.

Remarque 2.5. Le mode n'est en général pas unique; une série statistique peut admettre plusieurs modes (cas d'effectif ex-æquo).

Exemple 2.3 (suite). La modalité la plus fréquente dans cet exemple est 2 (l'effectif correspondant est 32). Il s'agit du mode de cette série.

Pour obtenir un indicateur de position moins sensible aux valeurs extrêmes que la moyenne, on peut chercher à séparer la série en deux parties de même effectif. Cette idée conduit à la définition de la médiane.

Définition 2.8. On appelle médiane d'une série statistique toute valeur me telle que :

$$\mathbf{P}[X \leq me] \geq 0,5 \quad \text{et} \quad \mathbf{P}[X \geq me] \geq 0,5.$$

Cette définition peut être généralisée pour définir la notion de *quantile*.

Définition 2.9. Soit $p \in]0; 1[$. On appelle quantile d'ordre p toute valeur q_p telle que :

$$\mathbf{P}[X \leq q_p] \geq p \quad \text{et} \quad \mathbf{P}[X \geq q_p] \geq 1 - p.$$

Remarque 2.6. On peut voir que la médiane est le quantile d'ordre 0,5. Les quantiles d'ordre 0,25 et 0,75 sont appelés premier et troisième *quartiles*. Les quantiles d'ordre 0,1 et 0,9 sont appelés premier et neuvième *déciles*.

Exemple 2.3 (suite). On a, dans cet exemple, que $\mathbf{P}[X \leq 2] = F(2) = 0,69 \geq 0,5$ et $\mathbf{P}[X \geq 2] = 1 - F(1) = 1 - 0,37 = 0,63 \geq 0,5$. La médiane est donc $me = 2$. Dans la pratique, on observe que $F(1) = \mathbf{P}[X \leq 1] = 0,37 \leq 0,5$ et que $F(2) = \mathbf{P}[X \leq 2] = 0,69 \geq 0,5$. Ceci suffit pour déterminer que la médiane est $me = 2$. On obtient de la même façon que le premier quartile est $q_{0,25} = 1$ et le troisième quartile est $q_{0,75} = 3$.

Paramètres de dispersion

Certains indicateurs de dispersion ont des définitions très simples, il s'agit de l'*étendue* et de l'*étendue inter-quartiles*.

Définition 2.10. On appelle *étendue de la série statistique* la quantité $x_r - x_1$ où x_1 est la plus petite modalité et x_r la plus grande modalité.

On appelle *étendue inter-quartiles de la série statistique* la quantité $q_{0,75} - q_{0,25}$ où $q_{0,25}$ et $q_{0,75}$ sont les premier et troisième quartiles.

Exemple 2.3 (suite). On a simplement que l'étendue est $x_7 - x_1 = 6 - 0 = 6$ et l'étendue inter-quartiles $q_{0,75} - q_{0,25} = 3 - 1 = 2$.

Un paramètre de dispersion permettant de mesurer de quelle manière se disperse une série statistique autour de sa moyenne est la *variance*.

Définition 2.11. La variance $V[X]$ d'une série statistique est définie par :

$$V[X] = \frac{1}{N} \sum_{i=1}^r n_i (x_i - \bar{X})^2 = \frac{n_1(x_1 - \bar{X})^2 + \dots + n_r(x_r - \bar{X})^2}{N},$$

où x_1, x_2, \dots, x_r sont les modalités, n_1, n_2, \dots, n_r leurs effectifs respectifs, N l'effectif total de la population et \bar{X} la moyenne de la série.

Remarque 2.7. Il s'agit de la moyenne des carrés des écarts à la moyenne.

En développant les carrés et regroupant les termes, on obtient immédiatement la réécriture suivante de la variance, connue sous le nom de *formule de décentrage de la variance*.

Proposition 2.3.

$$V[X] = \frac{1}{N} \sum_{i=1}^r n_i x_i^2 - \bar{X}^2 = \frac{n_1 x_1^2 + \dots + n_r x_r^2}{N} - \bar{X}^2.$$

Du fait de l'élevation au carré, les unités se retrouvent modifiées dans la variance. Pour palier à ce problème, on introduit l'*écart-type* qui s'exprime dans la même unité que la moyenne et mesure encore la dispersion des données autour de la moyenne.

Définition 2.12. L'écart-type σ d'une série statistique est défini par :

$$\sigma = \sqrt{V[X]}.$$

Un autre problème lié aux unités est un problème d'échelle. Imaginons que l'on veuille comparer la dispersion de deux séries de longueurs la première étant exprimée en mètres et la seconde en kilomètres. La comparaison ne peut pas se faire directement sur les écarts-types du fait des unités. On introduit pour cela un *écart-type relatif* qui est quant à lui sans unité.

Définition 2.13. On appelle *écart-type relatif* (ou *coefficient de variation*) d'une série statistique la quantité :

$$c_v = \frac{\sigma}{\bar{X}} \quad \text{si } \bar{X} \neq 0.$$

Remarque 2.8. L'écart-type relatif n'est pas défini si la moyenne est nulle. Il faut être vigilant au fait que ce paramètre est sensible à la position des données.

Exemple 2.3 (suite). On a, avec les données de l'Exemple 2.3 :

$$\begin{aligned} V[X] &= \frac{9 \times (0 - 1,96)^2 + \dots + 1 \times (6 - 1,96)^2}{100} \\ &= \frac{9 \times 0^2 + \dots + 1 \times 6^2}{100} - 1,96^2 = 1,3784. \end{aligned}$$

On en déduit que :

$$\sigma = \sqrt{1,3784} \simeq 1,1741$$

puis que

$$c_v = \frac{\sqrt{1,3784}}{1,96} \simeq 0,5990.$$

La variance mesure la dispersion des données autour de la moyenne au sens des carrés des écarts. Un autre indicateur mesure la dispersion des données autour de la moyenne au sens des valeurs absolues des écarts. Il s'agit de l'*écart absolu moyen*.

Définition 2.14. L'écart absolu moyen EAM d'une série statistique est défini par :

$$\text{EAM} = \frac{1}{N} \sum_{i=1}^r n_i |x_i - \bar{X}| = \frac{n_1 |x_1 - \bar{X}| + \dots + n_r |x_r - \bar{X}|}{N},$$

où x_1, x_2, \dots, x_r sont les modalités, n_1, n_2, \dots, n_r leurs effectifs respectifs, N l'effectif total de la population et \bar{X} la moyenne de la série.

Exemple 2.3 (suite). On a, avec les données de l'Exemple 2.3 :

$$\text{EAM} = \frac{9 \times |0 - 1,96| + \dots + 1 \times |6 - 1,96|}{100} = 0,8904.$$

2.4 Cas quantitatif continu ou discret avec regroupement en classes

Lorsque la variable statistique à étudier est continue ou discrète mais les modalités sont nombreuses, il est utile de collecter les données en les regroupant en des classes. On suppose, dans cette section, se trouver dans une telle situation. On note $C_i = [b_{i-1}, b_i[$ les classes dans lesquelles ont été regroupées les données et n_i leurs effectifs respectifs. Il est important de noter que les classes n'ont pas nécessairement des amplitudes identiques. Il convient d'être attentifs aux amplitudes des classes considérées pour interpréter les données.

2.4.1 Tableau statistique

Outre les modalités, effectifs, effectifs cumulés croissants, fréquences et fréquences cumulées croissantes définis comme dans les sections précédentes, le tableau statistique décrivant une variable continue ou discrète avec regroupement en classes contient les *amplitudes*, *amplitudes relatives*, *effectifs relatifs* et *fréquences relatives*.

Définition 2.15. 1. L'amplitude de la classe $C_i = [b_{i-1}, b_i[$ est la quantité $a_i = b_i - b_{i-1}$.
 2. Étant fixé un paramètre A (arbitraire), l'amplitude relative de la classe $C_i = [b_{i-1}, b_i[$ est la quantité $a_i^r = \frac{a_i}{A}$, son effectif relatif est $n_i^r = \frac{n_i}{a_i}$ et sa fréquence relative est la quantité $f_i^r = \frac{f_i}{a_i}$.

Remarque 2.9.

1. En général, on choisit comme amplitude de référence A , l'amplitude la plus courante ou toute amplitude paraissant rendre les calculs les plus simples possibles et les représentations graphiques agréables.
2. Il est immédiat de vérifier que l'on peut réécrire f_i^r sous la forme $f_i^r = \frac{n_i^r}{N}$.
3. La quantité f_i^r représente la densité de la classe C_i et est primordiale pour les représentations et interprétations dans les cas avec regroupement en classes. On peut ici faire l'analogie avec la densité de population en géographie (il est plus judicieux de regarder les densités de populations que les effectifs pour décider quelle région est la plus peuplée ; par exemple, comparez la population de Dijon intra-muros et du reste de la Bourgogne en nombre d'habitants puis en densité de population.).

Exemple 2.4. Durant le mois de janvier 2016, on a relevé les précipitations journalières (exprimées en mm) sur Dijon. On a consigné les résultats dans le tableau suivant.

Hauteur des précipitations (en mm)	[0; 1[[1; 3[[3; 6[[6; 9[[9; 14[
Nombre de jours	17	5	5	2	2

La population étudiée est l'ensemble des jours sur le mois de janvier 2016 et la variable statistique étudiée est la hauteur des précipitations relevées à Dijon chaque jour. Il s'agit d'une variable quantitative continue. On choisit (arbitrairement) de normaliser par $A = 1$, de sorte que $a_i^r = a_i$. On dresse le tableau statistique complet de cette série en notant $N = 31$ l'effectif total et pour $i \in \{1, \dots, 5\}$:

- $C_i = [b_{i-1}; b_i[$ la i^e classe,
- $c_i = \frac{b_{i-1} + b_i}{2}$ le centre de la i^e classe,
- $a_i = b_i - b_{i-1}$ l'amplitude de la i^e classe,
- $a_i^r = \frac{a_i}{A} = a_i$,
- n_i l'effectif de la i^e classe,
- $f_i = \frac{n_i}{N}$ la fréquence de la i^e classe,
- n_i^r l'effectif relatif de la i^e classe,
- $f_i^r = \frac{n_i^r}{N}$ la fréquence relative de la i^e classe,
- N_i l'effectif cumulé croissant jusqu'à la i^e classe,
- $F_i = \frac{N_i}{N}$ la fréquence cumulée croissante jusqu'à la i^e classe.

i	C_i	c_i	$a_i = a_i^r$	n_i	f_i	n_i^r	f_i^r	N_i	F_i
1	[0; 1[0,5	1	17	0,55	17	0,55	17	0,55
2	[1; 3[2	2	5	0,16	2,5	0,08	22	0,71
3	[3; 6[4	3	5	0,16	1,67	0,05	27	0,87
4	[6; 9[7,5	3	2	0,06	0,67	0,02	29	0,94
5	[9; 14[11,5	5	2	0,06	0,4	0,01	31	1

2.4.2 Représentations graphiques

Histogramme des fréquences

Contrairement à ce qui a été fait dans les cas qualitatifs et quantitatifs sans regroupement en classes, l'histogramme des fréquences prend ici la forme d'un diagramme en rectangles tel que l'aire du rectangle associé à une classe représente sa fréquence. Le rectangle associé à une classe a pour largeur son amplitude et pour **hauteur** sa fréquence **relative**. On peut encore commenter à l'aide de cet histogramme la symétrie ou dissymétrie de la distribution.

La Figure 2.5 représente l'histogramme des fréquences relatives obtenu dans l'Exemple 2.4. On observe dans ce cas une *traine* (ou *queue de distribution*) étalée vers la droite.

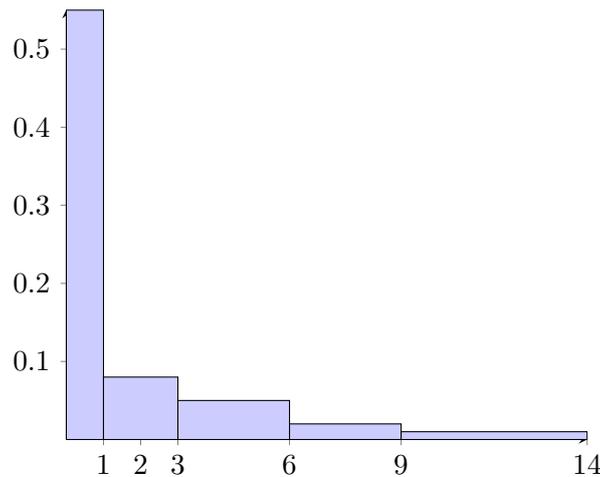


FIGURE 2.5 – Histogramme des fréquences relatives (Exemple 2.4).

Polygone des fréquences cumulées croissantes et fonction de répartition

Supposons que l'on dispose de r classes $C_1 = [b_0, b_1[$, \dots , $C_r = [b_{r-1}, b_r[$. Le polygone des fréquences cumulées est la ligne polygonale dont l'ordonnée est nulle pour $x \leq b_0$, vaut 1 pour $x \geq b_r$ et interpolant linéairement les points $(b_0, 0)$, (b_1, F_1) , \dots , (b_{r-1}, F_{r-1}) , $(b_r, F_r = 1)$ (c'est-à-dire reliant deux points successifs par un segment). Le polygone des fréquences cumulées fournit une approximation de la fonction de répartition de la variable statistique étudiée. Plus précisément, si l'on fait l'hypothèse que les données se répartissent uniformément au sein de chaque classe, le polygone des fréquences cumulées est le graphe de la fonction de répartition $t \mapsto F(t) = \mathbf{P}[X \leq t]$ de la variable aléatoire X . On fait usuellement cette hypothèse.

On peut observer qu'une variable aléatoire continue satisfait les propriétés suivantes.

Proposition 2.4. *Soit X une variable aléatoire continue. La fonction de répartition de X est croissante et continue. En particulier, on a pour tout $x \in \mathbf{R}$:*

$$\mathbf{P}[X < x] = \mathbf{P}[X \leq x] = F(x) \quad \text{et} \quad \mathbf{P}[X = x] = 0.$$

La Figure 2.6 représente la fonction de répartition obtenue dans l'Exemple 2.4.

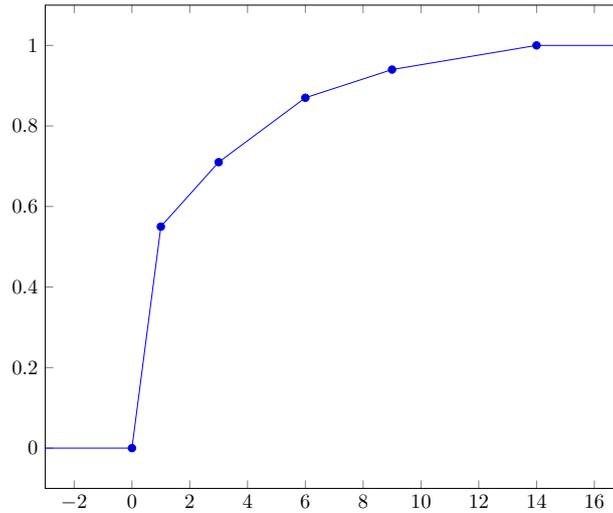


FIGURE 2.6 – Fonction de répartition d’une variable continue (Exemple 2.4).

On n’a accès, sans calcul supplémentaire, qu’aux valeurs de la fonction de répartition de X qu’en les points de $] - \infty, b_0] \cup \{b_1, \dots, b_{r-1}\} \cup [b_r, +\infty[$. Puisque les autres valeurs sont obtenues par interpolation linéaire, le théorème de Thalès permet d’obtenir facilement la valeur de $F(x)$ pour $x \in C_i = [b_{i-1}, b_i[$ (voir Figure 2.7). En effet, celui-ci permet d’affirmer que :

$$\frac{F(x) - F(b_{i-1})}{F(b_i) - F(b_{i-1})} = \frac{x - b_{i-1}}{b_i - b_{i-1}}, \text{ pour } x \in [b_{i-1}, b_i[.$$

Alternativement, puisque le graphe de la fonction de répartition est un segment de droite entre les abscisse b_{i-1} et b_i , la proposition suivante peut s’obtenir en déterminant l’équation de la droite passant par les points A de coordonnées $(x_A, y_A) = (b_{i-1}, F(b_{i-1}))$ et B de coordonnées $(x_B, y_B) = (b_i, F(b_i))$. L’équation de cette droite de la forme $y = ax + b$ est caractérisée par la donnée de son coefficient directeur

$$a = \frac{y_B - y_A}{x_B - x_A} = \frac{F(b_i) - F(b_{i-1})}{b_i - b_{i-1}}$$

et de son ordonnée à l’origine déterminée en profitant du fait qu’elle passe en particulier par le point A en écrivant que :

$$F(b_{i-1}) = y_A = ax_A + b = \frac{F(b_i) - F(b_{i-1})}{b_i - b_{i-1}} b_{i-1} + b$$

et donc

$$\begin{aligned} b &= F(b_{i-1}) - \frac{F(b_i) - F(b_{i-1})}{b_i - b_{i-1}} b_{i-1} = \frac{F(b_{i-1})(b_i - b_{i-1})}{b_i - b_{i-1}} - \frac{F(b_i)b_{i-1} - F(b_{i-1})b_{i-1}}{b_i - b_{i-1}} \\ &= \frac{F(b_{i-1})b_i - F(b_{i-1})b_{i-1} - F(b_i)b_{i-1} + F(b_{i-1})b_{i-1}}{b_i - b_{i-1}} = \frac{F(b_{i-1})b_i - F(b_i)b_{i-1}}{b_i - b_{i-1}}. \end{aligned}$$

On obtient ainsi que :

Proposition 2.5. Soit X une variable statistique continue dont la fonction de répartition est notée F . Pour tout x dans la classe $C_i = [b_{i-1}, b_i[$, on a :

$$\begin{aligned} F(x) &= \frac{x - b_{i-1}}{b_i - b_{i-1}} (F(b_i) - F(b_{i-1})) + F(b_{i-1}) \\ &= \frac{F(b_i) - F(b_{i-1})}{b_i - b_{i-1}} x + \frac{F(b_{i-1})b_i - F(b_i)b_{i-1}}{b_i - b_{i-1}} \\ &= \frac{f_i^r}{A} (x - b_{i-1}) + F(b_{i-1}), \end{aligned}$$

où A est l'amplitude de référence choisie pour la normalisation et f_i^r la fréquence relative de la classe C_i .

Preuve : La première identité est une conséquence du Théorème de Thalès ; la seconde s'obtient en écrivant l'équation de la droite passant par les points de coordonnées $(b_{i-1}, F(b_{i-1}))$ et $(b_i, F(b_i))$; la troisième s'obtient à partir de la première par le jeu d'écriture suivant :

$$\begin{aligned} F(x) &= \frac{x - b_{i-1}}{b_i - b_{i-1}} (F(b_i) - F(b_{i-1})) + F(b_{i-1}) \\ &= \frac{x - b_{i-1}}{a_i} f_i + F(b_{i-1}) \\ &= \frac{f_i}{A a_i^r} (x - b_{i-1}) + F(b_{i-1}) \\ &= \frac{f_i^r}{A} (x - b_{i-1}) + F(b_{i-1}). \end{aligned}$$

□

Remarque 2.10. Lorsque $A = 1$, f_i^r est le coefficient directeur du segment de droite correspondant à la classe C_i dans le polygone des fréquences cumulées et s'interprète alors pleinement comme la densité de cette classe. Dans ce cas, $F(t)$ est l'aire de la partie de l'histogramme de X située à gauche de la droite d'équation $x = t$.

Exemple 2.4 (suite). Illustrons la méthode d'interpolation linéaire à l'aide de l'Exemple 2.4. On a :

$$\mathbf{P}[X \geq 2] = 1 - \mathbf{P}[X < 2] = 1 - \mathbf{P}[X \leq 2] = 1 - F(2).$$

On utilise que le meilleur encadrement de 2 par des bornes de classe est $1 < 2 < 3$ et la méthode d'interpolation linéaire pour déterminer $F(2)$. Celle-ci donne que :

$$\frac{F(2) - F(1)}{F(3) - F(1)} = \frac{2 - 1}{3 - 1} = \frac{1}{2}$$

donc

$$F(2) = \frac{1}{2}(F(3) - F(1)) + F(1) = \frac{1}{2}(0,71 - 0,55) + 0,55 = 0,63.$$

Ainsi, $\mathbf{P}[X \geq 2] = 1 - F(2) = 1 - 0,63 = 0,37$ et on a relevé au moins $2mm$ de précipitations durant 37% des jours.

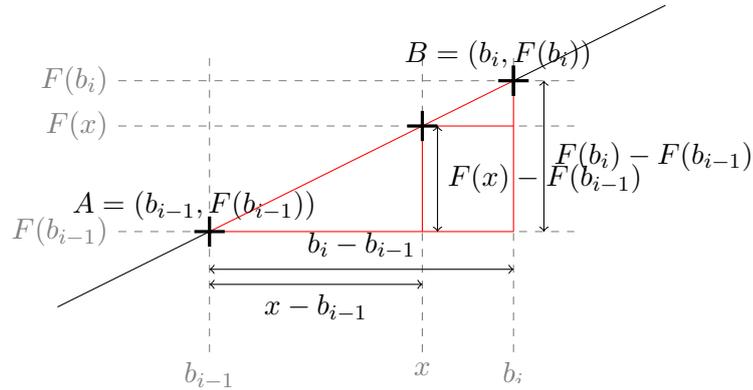


FIGURE 2.7 – Illustration de la méthode d'interpolation linéaire à l'aide du Théorème de Thalès.

2.4.3 Paramètres statistiques

Les paramètres statistiques sont définis dans le cas continu dans le même esprit que dans le cas discret. Il faut cependant faire attention à quelques subtilités. La première porte sur la définition des classes modales.

Définition 2.16. On appelle classe modale d'une série statistique continue ou discrète avec regroupements en classes toute classe $C_i = [b_{i-1}, b_i[$ dont l'effectif **relatif** est maximal parmi tous les effectifs **relatifs**.

Remarque 2.11. La classe modale n'est en général pas unique. On peut formuler la définition en remplaçant *effectif relatif* par *fréquence relative*. Il est primordial de garder en tête l'idée de « classe la plus densément peuplée » et de ne pas oublier l'adjectif **relatif** dans cette définition.

À cause du regroupement en classes, on n'aura pas accès ici aux valeurs exactes de la moyenne, de la variance, de l'écart-type, du coefficient de variation ou de l'écart-absolute moyen, mais seulement à des approximations de ceux-ci. En fait, pour faire les calculs, il faut remplacer les modalités qui apparaissent dans les formules par un représentant de chaque classe. En faisant, par convention, l'hypothèse d'une répartition uniforme au sein de chaque classe, on choisit pour représentant de la classe $C_i = [b_{i-1}, b_i[$ son centre $c_i = \frac{b_{i-1} + b_i}{2}$. Avec ces notations, ces paramètres se calculent avec les formules :

$$\bar{X} = \frac{1}{N} \sum_{i=1}^r n_i c_i = \frac{n_1 c_1 + \cdots + n_r c_r}{N},$$

$$V[X] = \frac{1}{N} \sum_{i=1}^r n_i (c_i - \bar{X})^2 = \frac{n_1 (c_1 - \bar{X})^2 + \cdots + n_r (c_r - \bar{X})^2}{N},$$

$$V[X] = \frac{1}{N} \sum_{i=1}^r n_i c_i^2 - \bar{X}^2 = \frac{n_1 c_1^2 + \cdots + n_r c_r^2}{N} - \bar{X}^2.$$

$$\sigma = \sqrt{V[X]}.$$

$$c_v = \frac{\sigma}{\bar{X}}.$$

$$\text{EAM} = \frac{1}{N} \sum_{i=1}^r n_i |c_i - \bar{X}| = \frac{n_1 |c_1 - \bar{X}| + \cdots + n_r |c_r - \bar{X}|}{N}.$$

Médiane, quantiles, étendue et étendue inter-quartile sont définis exactement comme dans la section précédente. On peut toutefois donner une définition plus maniable de la médiane et des quantiles dans le cas de variables statistiques continues.

Proposition 2.6. *Soit X une variable statistique continue et F sa fonction de répartition.*

On appelle médiane toute valeur me telle que $F(me) = 0,5$.

On appelle quantile d'ordre p toute valeur q_p telle que $F(q_p) = p$.

Remarque 2.12. Pour déterminer une médiane dans le cas continu, on commence par déterminer dans quelle classe $C_i = [b_{i-1}, b_i[$ celle-ci se trouve. On utilise, dans un deuxième temps, la méthode d'interpolation linéaire pour déterminer la valeur exacte de me grâce à la relation :

$$\frac{me - b_{i-1}}{b_i - b_{i-1}} = \frac{\frac{1}{2} - F(b_{i-1})}{F(b_i) - F(b_{i-1})}.$$

On procède de manière analogue pour déterminer un quantile.

Exemple 2.4 (suite). Les paramètres de la série statistique de l'Exemple 2.4 sont les suivants :

— la classe modale est la classe ayant la fréquence relative la plus élevée ; il s'agit ici de la classe $C_1 = [0; 1[$.

— la moyenne (arithmétique) est donnée par :

$$\bar{X} = \frac{n_1 c_1 + \cdots + n_5 c_5}{N} \simeq 2,5483871;$$

— La médiane me est la valeur telle que $F(me) = \mathbf{P}[X \leq me] = 0,5$. On a $F(0) = 0$ et $F(1) = 0,55$ et cette valeur est dans la classe $C_1 = [0; 1[$. On détermine me par interpolation linéaire :

$$\frac{me - 0}{1 - 0} = \frac{F(me) - F(0)}{F(1) - F(0)}$$

soit

$$me = \frac{0,5}{0,55} = \frac{10}{11} \simeq 0,90909;$$

— La variance est donnée par :

$$V[X] = \frac{n_1(c_1 - \bar{X})^2 + \cdots + n_5(c_5 - \bar{X})^2}{N} = \frac{n_1 c_1^2 + \cdots + n_5 c_5^2}{N} - \bar{X}^2 \simeq 9,71540;$$

— l'écart-type est donné par :

$$\sigma = \sqrt{V[X]} \simeq 3,11695;$$

— le coefficient de variation est donné par :

$$c_v = \frac{\sigma}{\bar{X}} \simeq 1,22310;$$

— l'écart-absolu moyen est donné par :

$$\text{EAM} = \frac{n_1 |c_1 - \bar{X}| + \cdots + n_5 |c_5 - \bar{X}|}{N} \simeq 2,42352.$$

2.5 Complément : d'autres moyennes

La moyenne arithmétique définie dans la Définition 2.6, est certainement la plus connue et la plus fréquemment utilisée. Il existe pourtant d'autres notions de moyenne qui sont, dans des situations particulières, plus adaptées ou représentatives.

Par exemple, supposons que pour se rendre sur son lieu de vacances, une famille réalise son trajet aller à une vitesse de 80km/h et son trajet retour à une vitesse de 60km/h. On peut chercher à déterminer la vitesse moyenne de la famille lors de cet aller-retour et nous allons voir que la moyenne arithmétique n'est pas adaptée. Notons D la distance séparant le lieu de vacances du domicile de la famille. Lors de son trajet aller, la famille a parcouru cette distance D en un temps de $T_{\text{aller}} = \frac{D}{80}$ heures. Lors de son trajet retour, la famille a parcouru cette distance D en un temps de $T_{\text{retour}} = \frac{D}{60}$ heures. Ainsi, sur l'aller-retour, la famille a parcouru une distance de $2D$ en un temps $T_{\text{aller}} + T_{\text{retour}}$. Sa vitesse moyenne sur l'aller-retour est donc :

$$\begin{aligned} \frac{2D}{T_{\text{aller}} + T_{\text{retour}}} &= \frac{2D}{\frac{D}{80} + \frac{D}{60}} = \frac{2}{\frac{1}{80} + \frac{1}{60}} \\ &= \frac{2}{\frac{3}{240} + \frac{4}{240}} = \frac{480}{7} \simeq 68,57 \text{ km/h.} \end{aligned}$$

Nous venons de construire la *moyenne harmonique* des deux vitesses.

Définition 2.17. Soient $a_1, \dots, a_N \in \mathbf{R}^*$ tels que $\sum_{i=1}^N \frac{1}{a_i} \neq 0$.

On appelle moyenne harmonique de a_1, \dots, a_N la quantité :

$$H = \frac{N}{\frac{1}{a_1} + \dots + \frac{1}{a_N}} = \frac{N}{\sum_{i=1}^N \frac{1}{a_i}}.$$

Remarque 2.13.

1. Si les différentes valeurs prises par les a_i sont x_1, \dots, x_r avec multiplicités respectives n_1, \dots, n_r , la moyenne harmonique se réécrit sous la forme :

$$H = \frac{N}{\frac{n_1}{x_1} + \dots + \frac{n_r}{x_r}} = \frac{N}{\sum_{i=1}^r \frac{n_i}{x_i}},$$

avec $N = n_1 + \dots + n_r$.

2. Lorsqu'elle est bien définie, la moyenne harmonique est l'inverse de la moyenne arithmétique des inverses.

Définition 2.18. Soient $a_1, \dots, a_N \in \mathbf{R}_+^*$.

On appelle moyenne géométrique de a_1, \dots, a_N la quantité :

$$G = (a_1 \times \dots \times a_N)^{\frac{1}{N}} = \sqrt[N]{\prod_{i=1}^N a_i}.$$

Remarque 2.14.

1. Si les différentes valeurs prises par les a_i sont x_1, \dots, x_r avec multiplicités respectives n_1, \dots, n_r , la moyenne géométrique se réécrit sous la forme :

$$G = \sqrt[N]{\prod_{i=1}^r x_i^{n_i}},$$

avec $N = n_1 + \dots + n_r$.

2. Lorsqu'elle est bien définie, la moyenne géométrique est l'exponentielle de la moyenne arithmétique des logarithmes des valeurs.
3. La moyenne géométrique doit son nom au fait que, pour tous $a, b > 0$, la longueur du carré dont la surface est la même que celle du rectangle de côtés a et b , est la moyenne géométrique de ces deux nombres.
4. Moins sensible que la moyenne arithmétique aux valeurs les plus extrêmes d'une série de données, la moyenne géométrique donne une meilleure estimation de la tendance centrale des données dans le cas d'une distribution à longue traîne. Par exemple, si l'on s'intéresse au salaire moyen dans une entreprise, la moyenne géométrique sera moins « tirée vers le haut » par les (rares) salaires les plus élevés que la moyenne arithmétique. De façon générale, l'utilisation de cette notion de moyenne sera adaptée dans les cas où les valeurs sont pour la plus part du même ordre de grandeur et que de rares valeurs extrêmement hautes par rapport aux autres viennent perturber la tendance centrale.

Chapitre 3

Statistiques bivariées

Dans ce chapitre, on introduit quelques outils d'étude de relations entre variables statistiques définies sur une **même** population. L'objectif général est de pouvoir décider si deux variables X et Y définies sur une même population sont suffisamment liées pour pouvoir « expliquer » raisonnablement l'une grâce à l'autre. En d'autres termes, on cherche à établir une relation du type $Y = f(X)$ pour prédire Y grâce à la variable X . Le cas le plus simple est celui où la fonction f est affine, c'est-à-dire s'écrit sous la forme $f(x) = ax + b$. On parle alors de relation *affine*, *linéaire*, de *profondeur* (ou *degré*) un.

En plus des éléments mathématiques présentés dans ce chapitre, le bon sens doit nous guider lorsque l'on souhaite « expliquer » un phénomène par un autre. Par exemple, on peut raisonnablement vouloir expliquer le prix d'un appartement dans une ville donnée par sa superficie. Par contre, il n'est pas raisonnable de vouloir expliquer le montant de la dette grecque par le nombre d'utilisateur de Facebook (et pourtant, les deux quantités semblent évoluer de façon tout à fait similaire)¹.

On s'intéressera ici à l'intensité des liens entre deux variables statistiques (*corrélations*) et, lorsque ceci est possible, à la manière de prédire une variable par une autre dans le cadre de relations affines (*régression linéaire*).

3.1 Présentation et traitement des données

On considère dans tout ce chapitre un couple de variables statistiques (X, Y) définies sur la même population \mathcal{P} d'effectif total N . Chacune des variables X et Y peut être de nature différente (qualitative, quantitative discrète ou continue). On se place, pour la suite du chapitre, dans le cadre où les deux variables sont quantitatives discrètes. Le lecteur pourra effectuer par lui-même les légères modifications nécessaires pour le cas continu en s'inspirant du Chapitre 2.

Pour chaque individu $i \in \{1, \dots, N\}$, on dispose d'un couple d'observations (x_i, y_i) . Dans la pratique, N est grand et il arrive qu'un même couple d'observations soit associé à plusieurs individus. On a alors recours à un *tableau de contingence* pour synthétiser les données. Avec un léger abus de notations, on écrira x_1, \dots, x_l (resp. y_1, \dots, y_r) pour les différentes valeurs prises par la variable X (resp. Y) dans la population. Celui-ci prend la forme :

1. Exemple tiré de <http://www.bloomberg.com/news/articles/2011-12-01/correlation-or-causation>

$X \backslash Y$	y_1	y_2	\dots	y_j	\dots	y_r
x_1	$n_{1,1}$	$n_{1,2}$	\dots	$n_{1,j}$	\dots	$n_{1,r}$
x_2	$n_{2,1}$	$n_{2,2}$	\dots	$n_{2,j}$	\dots	$n_{2,r}$
\dots	\dots	\dots	\dots	\dots	\dots	\dots
x_i	$n_{i,1}$	$n_{i,2}$	\dots	$n_{i,j}$	\dots	$n_{i,r}$
\dots	\dots	\dots	\dots	\dots	\dots	\dots
x_l	$n_{l,1}$	$n_{l,2}$	\dots	$n_{l,j}$	\dots	$n_{l,r}$

où $n_{i,j}$ est le nombre d'individu pour lesquels les variables X et Y ont pris les valeurs x_i et y_j . On notera $f_{i,j} = \frac{n_{i,j}}{N}$, où $N = \sum_{i=1}^l \sum_{j=1}^r n_{i,j}$ est l'effectif total de la population, la fréquence de la caractéristique (x_i, y_j) .

Définition 3.1. Soit (X, Y) une série statistique bivariée.

1. On appelle effectifs marginaux en X les nombres $n_{i,\cdot}$, $1 \leq i \leq l$ définis par :

$$n_{i,\cdot} = n_{i,1} + \dots + n_{i,r} = \sum_{j=1}^r n_{i,j}.$$

2. On appelle effectifs marginaux en Y les nombres $n_{\cdot,j}$, $1 \leq j \leq r$ définis par :

$$n_{\cdot,j} = n_{1,j} + \dots + n_{l,j} = \sum_{i=1}^l n_{i,j}.$$

3. On appelle fréquence marginale de la caractéristique x_i , $1 \leq i \leq l$, le nombre $f_{i,\cdot} = \frac{n_{i,\cdot}}{N}$.

4. On appelle fréquence marginale de la caractéristique y_j , $1 \leq j \leq r$, le nombre $f_{\cdot,j} = \frac{n_{\cdot,j}}{N}$.

Remarque 3.1. Il découle directement des définitions précédentes que :

$$\sum_{i=1}^l n_{i,\cdot} = \sum_{j=1}^r n_{\cdot,j} = N \quad \text{et} \quad \sum_{i=1}^l f_{i,\cdot} = \sum_{j=1}^r f_{\cdot,j} = \sum_{i=1}^l \sum_{j=1}^r f_{i,j} = 1.$$

Il est pratique de compléter le tableau de contingence avec les effectifs marginaux et fréquences marginales. Celui-ci prend alors la forme :

$X \backslash Y$	y_1	y_2	\dots	y_j	\dots	y_r	$n_{i,\cdot}$	$f_{i,\cdot}$
x_1	$n_{1,1}$	$n_{1,2}$	\dots	$n_{1,j}$	\dots	$n_{1,r}$	$n_{1,\cdot}$	$f_{1,\cdot}$
x_2	$n_{2,1}$	$n_{2,2}$	\dots	$n_{2,j}$	\dots	$n_{2,r}$	$n_{2,\cdot}$	$f_{2,\cdot}$
\dots	\dots	\dots	\dots	\dots	\dots	\dots	\dots	\dots
x_i	$n_{i,1}$	$n_{i,2}$	\dots	$n_{i,j}$	\dots	$n_{i,r}$	$n_{i,\cdot}$	$f_{i,\cdot}$
\dots	\dots	\dots	\dots	\dots	\dots	\dots	\dots	\dots
x_l	$n_{l,1}$	$n_{l,2}$	\dots	$n_{l,j}$	\dots	$n_{l,r}$	$n_{l,\cdot}$	$f_{l,\cdot}$
$n_{\cdot,j}$	$n_{\cdot,1}$	$n_{\cdot,2}$	\dots	$n_{\cdot,j}$	\dots	$n_{\cdot,r}$	N	
$f_{\cdot,j}$	$f_{\cdot,1}$	$f_{\cdot,2}$	\dots	$f_{\cdot,j}$	\dots	$f_{\cdot,r}$		1

CHAPITRE 3. STATISTIQUES BIVARIÉES

Une série statistique bivariée peut être représentée graphiquement par :

- un *stéréogramme* (analogue dans le cas bivarié de l'histogramme du cas univarié en 3 dimensions),
- un nuage de points.

On utilisera ici la représentation par des nuages de points pondérés. Pour chaque couple (x_i, y_j) tel que $n_{i,j} \geq 1$, on représentera le point de coordonnées (x_i, y_j) auquel on attache l'étiquette $(n_{i,j})$. On pourra omettre l'étiquette $(n_{i,j})$ lorsque $n_{i,j} = 1$. Alternativement, on peut représenter le point de coordonnées (x_i, y_j) par un disque de rayon proportionnel à $n_{i,j}$.

Exemple 3.1. On a relevé les notes des étudiants d'un groupe de TP, durant les deux premiers TP notés d'informatique. Celle-ci sont présentées dans la liste suivante sous la forme (X, Y) où X est la note du premier TP et Y celle du deuxième :

(12, 13), (10, 9), (15, 16), (12, 13), (13, 12), (6, 7), (9, 10),
(6, 7), (16, 15), (11, 13), (12, 13), (18, 19), (1, 3), (11, 13).

Les variables X et Y sont quantitatives discrètes et le tableau de contingence complété par les effectifs marginaux et fréquences marginales s'écrit sous la forme :

$X \backslash Y$	3	7	9	10	12	13	15	16	19	$n_{i,\cdot}$	$f_{i,\cdot}$
1	1	0	0	0	0	0	0	0	0	1	$\frac{1}{14} \simeq 0,07$
6	0	2	0	0	0	0	0	0	0	2	$\frac{1}{7} \simeq 0,14$
9	0	0	0	1	0	0	0	0	0	1	$\frac{1}{14} \simeq 0,07$
10	0	0	1	0	0	0	0	0	0	1	$\frac{1}{14} \simeq 0,07$
11	0	0	0	0	0	2	0	0	0	2	$\frac{1}{7} \simeq 0,14$
12	0	0	0	0	0	3	0	0	0	3	$\frac{3}{14} \simeq 0,21$
13	0	0	0	0	1	0	0	0	0	1	$\frac{1}{14} \simeq 0,07$
15	0	0	0	0	0	0	0	1	0	1	$\frac{1}{14} \simeq 0,07$
16	0	0	0	0	0	0	1	0	0	1	$\frac{1}{14} \simeq 0,07$
18	0	0	0	0	0	0	0	0	1	1	$\frac{1}{14} \simeq 0,07$
$n_{\cdot,j}$	1	2	1	1	1	5	1	1	1	14	
$f_{\cdot,j}$	$\frac{1}{14}$	$\frac{1}{7} \simeq 0,14$	$\frac{1}{14}$	$\frac{1}{14}$	$\frac{1}{14}$	$\frac{5}{14} \simeq 0,36$	$\frac{1}{14}$	$\frac{1}{14}$	$\frac{1}{14}$		1

Le nuage de points correspondant est représenté dans la Figure 3.1. On constate que les points se répartissent essentiellement le long d'une droite et il est raisonnable de chercher une relation affine entre les variables X et Y . Les techniques permettant de faire déterminer une telle relation seront développées dans la Section 3.3.

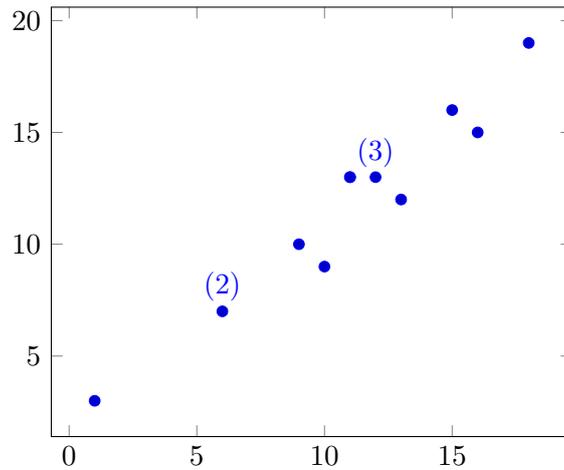


FIGURE 3.1 – Représentation de la série de l'Exemple 3.1 par un nuage de points pondérés.

3.2 Étude et mesure des liens entre les variables

3.2.1 Indépendance

Informellement, la notion d'*indépendance* exprime le fait que la variable X n'influence pas la variable Y et réciproquement.

Définition 3.2. Soit (X, Y) une série statistique bivariée.

On dit que X et Y sont indépendantes si :

$$f_{i,j} = f_{i,\cdot} \cdot f_{\cdot,j} \quad \text{pour tout } (i, j) \in \{1, \dots, l\} \times \{1, \dots, r\}.$$

Définition 3.3. Soit (X, Y) une série statistique bivariée.

1. On appelle fréquence conditionnelle de x_i sachant y_j la quantité $f_{i|j} = \frac{f_{i,j}}{f_{\cdot,j}}$.
2. On appelle fréquence conditionnelle de y_j sachant x_i la quantité $f_{j|i} = \frac{f_{i,j}}{f_{i,\cdot}}$.

Proposition 3.1. Soit (X, Y) une série statistique.

Les assertions suivantes sont équivalentes :

1. X et Y sont indépendantes,
2. pour tout $i \in \{1, \dots, l\}$, $f_{i|j}$ ne dépend pas de $j \in \{1, \dots, r\}$ (et vaut $f_{i,\cdot}$),
3. pour tout $j \in \{1, \dots, r\}$, $f_{j|i}$ ne dépend pas de $i \in \{1, \dots, l\}$ (et vaut $f_{\cdot,j}$).

Preuve : Montrons que 1. est équivalente à 2. ; la preuve du fait que 1. est équivalente à 3. est analogue. Par définition, X et Y sont indépendantes si, et seulement si, pour tout (i, j) , $f_{i,j} = f_{i,\cdot} \cdot f_{\cdot,j}$. En divisant par $f_{\cdot,j}$, on obtient que X et Y sont indépendantes si, et seulement si, pour tout (i, j)

$$f_{i|j} = \frac{f_{i,j}}{f_{\cdot,j}} = \frac{f_{i,\cdot} \cdot f_{\cdot,j}}{f_{\cdot,j}} = f_{i,\cdot}$$

□

Remarque 3.2.

1. Il découle de la Proposition 3.1 que les variables X et Y sont indépendantes si, et seulement si, les lignes ou les colonnes du tableau de contingence sont proportionnelles.
2. Lorsque chaque ligne du tableau de contingence ne contient qu'un seul effectif $n_{i,j}$ non nul, on dit que Y dépend totalement de X .
3. Dans la pratique, il n'y a jamais une indépendance « parfaite ». Pour mesurer l'intensité des liens entre X et Y , on introduit dans les sous-sections suivantes le *coefficient de Cramer* et le *coefficient de corrélation linéaire*.

3.2.2 Coefficient de Cramer

On souhaite déterminer si les composantes X et Y d'une série statistique bivariée sont indépendantes (ou proches de l'indépendance). Pour cela, on introduit une distance qui mesure l'écart entre les effectifs observés $n_{i,j}$ et les effectifs calculés théoriquement $T_{i,j}$ en faisant l'hypothèse que X et Y sont effectivement indépendantes.

Définition 3.4. Soit (X, Y) une série statistique bivariée.

On appelle chi-2 de la série statistique (X, Y) , la quantité :

$$\chi^2 = \sum_{i=1}^l \sum_{j=1}^r \frac{(n_{i,j} - T_{i,j})^2}{T_{i,j}}$$

où $T_{i,j} = \frac{n_{i.}n_{.j}}{N}$ est l'effectif théorique de la caractéristique (x_i, y_j) en faisant l'hypothèse que X et Y sont effectivement indépendantes.

Proposition 3.2. Soit (X, Y) une série statistique bivariée.

On a :

$$\chi^2 \in [0, \chi_{max}^2]$$

où $\chi_{max}^2 = N \times \min(l - 1, r - 1)$.

Remarque 3.3.

1. Si X et Y sont indépendantes, $\chi^2 = 0$; au contraire, si Y dépend totalement de X alors $\chi^2 = \chi_{max}^2$.
2. La distance du χ^2 est utilisée dans un test d'indépendance (du même nom) qui sera étudié en deuxième année.

Définition 3.5. Soit (X, Y) une série statistique bivariée.

On appelle coefficient de Cramer (ou V de Cramer) de la série statistique (X, Y) , la quantité :

$$C = \sqrt{\frac{\chi^2}{\chi_{max}^2}}$$

Proposition 3.3. Soit (X, Y) une série statistique bivariée.

On a :

$$C \in [0, 1].$$

Remarque 3.4. Lorsque C est proche de 0, X et Y sont « presque » indépendantes. Lorsque C est proche de 1, X et Y sont fortement liées ; ceci n'implique pour autant pas de relation fonctionnelle entre X et Y ni de relation de cause à effet.

Interpétation : Plus précisément, on considérera que la dépendance entre les variables est :

1. *très faible* si $C \in [0; 0,045]$,
2. *faible* si $C \in]0,045; 0,09]$,
3. *moyenne* si $C \in]0,09; 0,18]$,
4. *forte* si $C \in]0,18; 0,36]$,
5. *très forte* si $C \in]0,36; 1]$.

Exemple 3.2. On a relevé les couleurs des yeux X et pointures Y des clientes ayant acheté une paire de chaussures dans un grand magasin un jour donné. Le résultats ont été consignés dans le tableau suivant.

$X \backslash Y$	37	38	39	40
marrons	5	20	11	3
verts	2	10	10	3
bleus	5	11	15	5

On complète ce tableau en calculant les effectifs marginaux et l'effectif total N .

$X \backslash Y$	37	38	39	40	$n_{i\cdot}$
marrons	5	20	11	3	39
verts	2	10	10	3	25
bleus	5	11	15	5	36
$n_{\cdot j}$	12	41	36	11	$N = 100$

On en déduit les effectifs calculés théoriquement $T_{i,j}$ en faisant l'hypothèse que X et Y sont indépendantes. Ceux-ci sont présentés dans le tableau suivant.

$X \backslash Y$	37	38	39	40	$n_{i\cdot}$
marrons	$\frac{39 \times 12}{100} = 4,68$	15,99	14,04	4,29	39
verts	3	10,25	9	2,75	25
bleus	4,32	14,76	12,96	3,96	36
$n_{\cdot j}$	12	41	36	11	$N = 100$

Avec les deux derniers tableaux, on calcule le χ^2 de la série statistique (X, Y) :

$$\begin{aligned}\chi^2 &= \sum_{i=1}^l \sum_{j=1}^r \frac{(n_{i,j} - T_{i,j})^2}{T_{i,j}} \\ &= \frac{(5 - 4,68)^2}{4,68} + \frac{(20 - 15,99)^2}{15,99} + \dots + \frac{(5 - 3,96)^2}{3,96} \\ &\simeq 4,21.\end{aligned}$$

Puisque le nombre de lignes est $l = 3$, le nombre de colonnes $r = 4$ et l'effectif total $N = 100$, on obtient que $\chi_{\max}^2 = 100 \times \min(3 - 1, 4 - 1) = 200$. Il s'ensuit que le coefficient de Cramer de la série est :

$$C = \sqrt{\frac{\chi^2}{\chi_{\max}^2}} \simeq \sqrt{\frac{4,21}{200}} \simeq 0,15.$$

On en déduit que la couleur des yeux et la pointure sont des variables ayant une dépendance moyenne.

3.2.3 Coefficient de corrélation linéaire

Définition 3.6. Soit (X, Y) une série statistique bivariée avec X et Y quantitatives.

La covariance de (X, Y) est la quantité :

$$\text{Cov}(X, Y) = \frac{1}{N} \sum_{i=1}^l \sum_{j=1}^r n_{i,j} (x_i - \bar{X})(y_j - \bar{Y})$$

où $\bar{X} = \frac{1}{N} \sum_{i=1}^l n_{i,\cdot} x_i$ (resp. $\bar{Y} = \frac{1}{N} \sum_{j=1}^r n_{\cdot,j} y_j$) est la moyenne de la marginale en X (resp. en Y).

Remarque 3.5. Lorsque des regroupements en classes sont utilisés, il convient de remplacer dans la formule précédente les x_i et/ou y_j par les centres des classes correspondantes.

Des propriétés découlant directement de cette définition sont listées dans la proposition suivante.

Proposition 3.4. Soit (X, Y) une série statistique bivariée avec X et Y quantitatives.

1. Pour tous $a, b \in \mathbf{R}$, on a :

$$\text{Cov}(aX + b, Y) = a \text{Cov}(X, Y).$$

2. On a :

$$\text{Cov}(X, X) = V[X].$$

3. Si X et Y sont indépendantes, alors $\text{Cov}(X, Y) = 0$.

4. On a :

$$\text{Cov}(X, Y) = \frac{1}{N} \sum_{i=1}^l \sum_{j=1}^r n_{i,j} x_i y_j - \bar{X}\bar{Y}.$$

Définition 3.7. Soit (X, Y) une série statistique bivariée avec X et Y quantitatives.

Le coefficient de corrélation linéaire de (X, Y) est la quantité :

$$\text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{V}[X] \text{V}[Y]}}$$

Proposition 3.5. Soit (X, Y) une série statistique bivariée avec X et Y quantitatives.

1. On a :

$$\text{Cor}(X, Y) \in [-1, 1].$$

2. On a :

$$\text{Cor}(X, X) = 1, \quad \text{Cor}(X, -X) = -1$$

et pour tous $a, b \in \mathbf{R}$:

$$|\text{Cor}(X, aX + b)| = 1.$$

3. Si X et Y sont indépendantes, alors $\text{Cor}(X, Y) = 0$.

Interprétation : On dira qu'il y a :

- une *forte corrélation* entre X et Y si $|\text{Cor}(X, Y)| \geq 0,8$;
- une *corrélation médiocre* entre X et Y si $0,5 < |\text{Cor}(X, Y)| < 0,8$;
- une *mauvaise corrélation* entre X et Y si $|\text{Cor}(X, Y)| \leq 0,5$.

Un lien fort entre deux variables entraîne *a priori* une forte corrélation de celles-ci mais une forte corrélation ne suffit pas pour établir un lien de cause à effet entre deux variables. En effet, d'autres facteurs peuvent entrer en ligne de compte et il faut être attentif à ne pas mal interpréter ou surinterpréter les résultats d'une étude statistique. Par exemple, une étude a montré que les personnes résidant à proximité d'une centrale nucléaire sont significativement plus souvent malades que les autres. On ne peut pour autant pas affirmer que des problèmes de fuites ou autres au niveau des centrale provoquent ces maladies. On peut remarquer que les terrains situés dans ces zones sont généralement très bon marché. La santé et la pauvreté étant liées, l'étude ne permet pas, à elle seule, de conclure que les centrales nucléaires influent sur la santé.

Exemple 3.3. Une agence immobilière a relevé les surfaces et prix des 10 appartements qu'elle a vendus une semaine donnée et les a consignés dans le tableau suivant :

n° i de l'appartement	Surface x_i en m^2	Prix y_i en K€
1	20	47
2	122	230
3	45	87
4	56	98
5	18	39
6	54	90
7	77	180
8	80	176
9	68	124
10	32	45

On a ici :

$$\bar{X} = \frac{20 + \dots + 32}{10} = 57,2$$

$$\bar{Y} = \frac{47 + \dots + 45}{10} = 111,6$$

$$V[X] = \frac{(20 - 57,2)^2 + \dots + (32 - 57,2)^2}{10} = 894,36$$

$$V[Y] = \frac{(47 - 111,6)^2 + \dots + (45 - 111,6)^2}{10} = 3813,44$$

$$\text{Cov}(X, Y) = \frac{(20 - 57,2)(47 - 111,6) + \dots + (32 - 57,2)(45 - 111,6)}{10} = 1794,18$$

et donc

$$\text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{V[X] V[Y]}} \simeq 0,97.$$

On constate qu'il y a une (très) forte corrélation entre prix et surface d'un appartement. De plus, les points relevés semble se répartir autour d'une droite comme le montre la Figure 3.2. On peut donc chercher à expliquer l'une des variables par l'autre au moyen d'une relation affine. Ceci est l'objet de la section suivante.

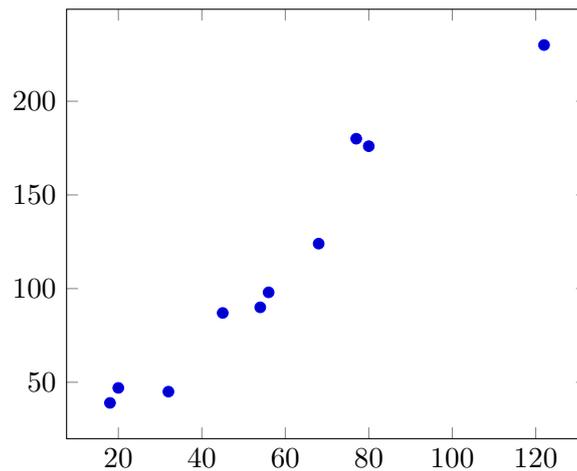


FIGURE 3.2 – Représentation des prix relevés des appartements en fonction de leurs surfaces (Exemple 3.3).

3.3 Régression

Lorsque deux variables quantitatives X et Y définies sur une même population sont fortement corrélées, il est naturel de rechercher une relation fonctionnelle entre celles-ci, c'est-à-dire chercher une fonction f telle que $Y = f(X)$. Quand on cherche à prédire Y en supposant X connue on parle de *régression* (ou d'*ajustement*) de Y en X . Quand on cherche à prédire X en supposant Y connue on parle de *régression* (ou d'*ajustement*) de X en Y . Il peut arriver que seule l'une des deux régressions ait du sens. Par exemple, il est plus sensé de vouloir prédire la glycémie d'un patient par la quantité de sucre qu'il a ingéré que l'inverse.

En toute généralité, une relation fonctionnelle entre les variables peut faire intervenir des fonctions relativement compliquées (modèles exponentiels, logarithmiques...). Nous restreindrons notre attention au cas simple où l'on cherche une relation entre X et Y de la forme

$$Y = aX + b \quad (\text{régression de } Y \text{ en } X)$$

ou

$$X = a'Y + b' \quad (\text{régression de } X \text{ en } Y).$$

Conditions d'application du modèle de régression linéaire : On effectuera une telle régression si :

- le nuage de points semble se répartir le long d'une droite,
- il est raisonnable de vouloir expliquer/prédire le caractère associé à la variable Y et par celui associé à X ou vice versa,
- le coefficient de corrélation $\text{Cor}(X, Y)$ est compris dans $[-1; -0,7] \cup [0,7; 1]$.

On cherche alors à trouver a et b (ou a' et b') tels que la droite de régression s'approche le mieux possible des données en un certain sens. Le sens retenu est celui des moindres carrés.

Définition 3.8. Soit (X, Y) une série statistique bivariée telle que X et Y sont quantitatives.

1. La droite de régression de Y en X (au sens des moindres carrés) est la droite $D_{Y|X}$ d'équation :

$$D_{Y|X} : y = ax + b,$$

avec a, b choisis de telle sorte que $E(a, b) = \sum_{i=1}^N (y_i - (ax_i + b))^2$ soit minimale.

2. La droite de régression de X en Y (au sens des moindres carrés) est la droite $D_{X|Y}$ d'équation :

$$D_{X|Y} : x = a'y + b',$$

avec a', b' choisis de telle sorte que $E(a', b') = \sum_{i=1}^N (x_i - (a'y_i + b'))^2$ soit minimale.

Remarque 3.6.

1. La droite de régression de Y en X minimise la somme des écarts verticaux aux points observés alors que la droite de régression de X en Y minimise la somme des écarts horizontaux aux points observés.
2. Les droites $D_{Y|X}$ et $D_{X|Y}$ sont confondues en cas de corrélation totale (*i.e.* si $\text{Cor}(X, Y) = \pm 1$). Dans le cas contraire, les droites $D_{Y|X}$ et $D_{X|Y}$ se coupent en $G(\bar{X}, \bar{Y})$ et forment un angle qui est d'autant plus fermé que les variables sont corrélées. Dans le cas où les variables sont indépendantes, $D_{Y|X}$ et $D_{X|Y}$ sont perpendiculaires.

Des techniques d'analyse permettent de montrer que :

Proposition 3.6. Soit (X, Y) une série statistique bivariée telle que X et Y sont quantitatives.

1. L'équation de la droite de régression de Y en X est :

$$D_{Y|X} : y = ax + b,$$

avec $a = \frac{\text{Cov}(X, Y)}{\text{V}[X]}$ et $b = \bar{Y} - a\bar{X}$.

2. L'équation de la droite de régression de X en Y est :

$$D_{X|Y} : x = a'y + b',$$

$$\text{avec } a' = \frac{\text{Cov}(X,Y)}{\text{V}[Y]} \text{ et } b' = \bar{X} - a'\bar{Y}.$$

Exemple 3.3 (suite). Reprenons l'Exemple 3.3. On a vu que $|\text{Cor}(X, Y)| \geq 0,8$ et que le nuage de points à une forme allongée ; le modèle de régression linéaire fait donc sens ici.

La droite de régression de Y en X s'écrit sous la forme :

$$y = ax + b,$$

avec

$$a = \frac{\text{Cov}(X, Y)}{\text{V}[X]} \simeq 2 \quad \text{et} \quad b = \bar{Y} - a\bar{X} \simeq -2,8.$$

Si un nouvel appartement dont la surface est de $x^* = 51\text{m}^2$ est proposé à la vente, cette droite de régression permet de prédire un prix de vente de :

$$y^* = ax^* + b \simeq 2 \times 51 - 2,8 = 99,2\text{K}\text{€}.$$

La droite de régression de X en Y s'écrit sous la forme :

$$x = a'y + b',$$

avec

$$a' = \frac{\text{Cov}(X, Y)}{\text{V}[Y]} \simeq 0,47 \quad \text{et} \quad b' = \bar{X} - a'\bar{Y} \simeq 4,748.$$

Ces droites de régressions sont représentées dans la Figure 3.3.

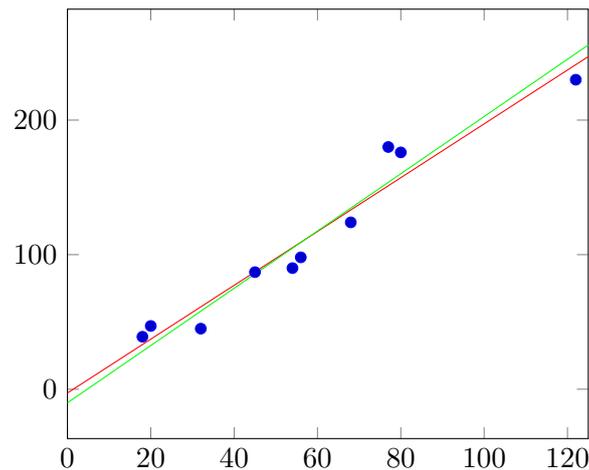


FIGURE 3.3 – Représentation des prix relevés des appartements en fonction de leurs surfaces (Exemple 3.3) et des droites de régression de Y en X (rouge) et de X en Y (vert).

3.4 Complément : test du χ^2 d'indépendance

Dans cette section, on présente la structure du *test du χ^2 d'indépendance*. Celui-ci précise en un certain sens les notions et décisions présentées dans la Section 3.2.2. Il pourra être appliqué dans des contextes divers dans plusieurs modules des Semestres 2, 3 et 4 (et même après), comme les modules de méthodes et conception d'enquêtes, marketing, mathématiques ou encore logiciels métiers et lors de projets tutorés ou stages.

On illustre ici la structure de ce test sur l'exemple suivant, facilement repéré par la couleur bleue.

Exemple 3.4. Un mois après le lancement d'une campagne de publicité, le service marketing d'une entreprise souhaite savoir si la sa campagne a été perçue aussi bien sur l'ensemble du territoire de Bourgogne/Franche-Comté ou non. Pour cela, il a interrogé un panel de 150 habitants sur leur département de résidence X et le fait Y qu'ils aient vu la publicité ou non. Les résultats ont été consignés dans le tableau suivant.

$X \backslash Y$	Oui	Non
Côte-d'Or	19	9
Doubs	20	9
Jura	8	6
Nièvre	5	6
Haute-Saône	4	8
Saône-et-Loire	5	23
Yonne	8	10
Territoire de Belfort	3	7

Étape 1 : Définition des hypothèses :

On définit les hypothèses :

H_0 (hypothèse nulle) : Les variables X et Y sont indépendantes.

H_1 (hypothèse alternative) : Les variables X et Y ne sont pas indépendantes.

Dans le contexte de l'Exemple 3.4, l'hypothèse H_0 s'interprète comme « la perception de la campagne ne dépend pas de la localisation géographique », autrement dit la campagne a été perçue de la même façon sur l'ensemble du territoire, alors que l'hypothèse H_1 s'interprète comme « la perception de la campagne dépend de la localisation géographique ». Dans ce second cas, on pourra chercher une explication au fait que cette campagne ne soit pas aussi bien perçue sur l'ensemble du territoire.

Étape 2 : Calcul des effectifs théoriques sous l'hypothèse d'indépendance H_0 :

CHAPITRE 3. STATISTIQUES BIVARIÉES

Cette étape est analogue à ce que l'on a présenté dans la Section 3.2.2. On détermine d'abord les effectifs marginaux en X et Y ($n_{i\cdot}$ et $n_{\cdot j}$), puis on dresse le tableau des effectifs théoriques en utilisant que ceux-ci sont donnés par :

$$T_{i,j} = \frac{n_{i\cdot} \cdot n_{\cdot j}}{N},$$

où N est l'effectif total.

Note importante : Si certains effectifs théoriques sont inférieurs à 5, on doit regrouper des lignes ou des colonnes pour que ce ne soit plus le cas, sans quoi l'approximation suivante par la loi du χ^2 n'est pas valable. Si des regroupements sont effectués pour les effectifs théoriques, on réalise les mêmes regroupements pour les effectifs observés.

Dans l'Exemple 3.2.2, on complète le tableau des effectifs observés avec les effectifs marginaux comme suit.

$X \backslash Y$	Oui	Non	$n_{i\cdot}$
Côte-d'Or	19	9	28
Doubs	20	9	29
Jura	8	6	14
Nièvre	5	6	11
Haute-Saône	4	8	12
Saône-et-Loire	5	23	28
Yonne	8	10	18
Territoire de Belfort	3	7	10
$n_{\cdot j}$	72	78	$N = 150$

On en déduit le tableau des effectifs théoriques sous l'hypothèse d'indépendance H_0 , donné ci-dessous.

$X \backslash Y$	Oui	Non	$n_{i\cdot}$
Côte-d'Or	$\frac{28 \times 72}{150} = 13,44$	14,56	28
Doubs	13,92	15,08	29
Jura	6,72	7,28	14
Nièvre	5,28	5,72	11
Haute-Saône	5,76	6,24	12
Saône-et-Loire	13,44	14,56	28
Yonne	8,64	9,36	18
Territoire de Belfort	4,8	5,2	10
$n_{\cdot j}$	72	78	$N = 150$

Observons que l'un des effectifs théorique est inférieur à 5 : 4,8 correspondant à la modalité (Territoire de Belfort, Oui). Il faut donc effectuer des regroupements. On choisit ici de regrouper la ligne du Territoire de Belfort avec la ligne de la Saône-et-Loire. Notons que ce choix est arbitraire, mais qu'un choix plus naturel s'impose lorsque les variables sont quantitatives.

On obtient donc dans cet exemple, le tableau de effectifs théoriques après regroupement :

$X \backslash Y$	Oui	Non	$n_{i.}$
Côte-d'Or	13,44	14,56	28
Doubs	13,92	15,08	29
Jura	6,72	7,28	14
Nièvre	5,28	5,72	11
Haute-Saône	5,76	6,24	12
Saône-et-Loire	13,44	14,56	28
Yonne et Territoire de Belfort	13,44	14,56	28
$n_{.j}$	72	78	$N = 150$

qui est à comparer au tableau des effectifs observés (après regroupement) :

$X \backslash Y$	Oui	Non	$n_{i.}$
Côte-d'Or	19	9	28
Doubs	20	9	29
Jura	8	6	14
Nièvre	5	6	11
Haute-Saône	4	8	12
Saône-et-Loire	5	23	28
Yonne et Territoire de Belfort	11	17	28
$n_{.j}$	72	78	$N = 150$

Étape 3 : Calcul de la distance du χ^2 observée :

On calcule, comme dans la Section 3.2.2, en utilisant éventuellement un tableau intermédiaire et en se basant sur les tableaux après regroupements, la distance du χ^2 observée qui est dans ce cadre la valeur observée de la statistique de test :

$$\chi_{\text{obs}}^2 = \sum_{i,j} \frac{(n_{i,j} - T_{i,j})^2}{T_{i,j}}$$

Dans le contexte de l'Exemple 3.4, on obtient :

$$\chi_{\text{obs}}^2 \simeq 22,1063.$$

Étape 4 : Degrés de liberté et loi de la statistique de test :

On détermine le nombre de degrés de liberté grâce à la formule :

$$\text{ddl} = (\text{nbc} - 1)(\text{nbl} - 1),$$

où nbc et nbl désignent respectivement le nombre de colonnes et de lignes dans le tableau après regroupements éventuels. La loi de la statistique de test est alors simplement la loi du Khi-2 à ddl degrés de liberté, notée $\chi^2(\text{ddl})$.

Dans le contexte de l'Exemple 3.4, on a après regroupement 7 lignes et 2 colonnes donc :

$$\text{ddl} = (7 - 1) \times (2 - 1) = 6,$$

et la statistique de test suit la loi $\chi^2(6)$.

Étape 5 : Risque d'erreur de première espèce et valeur critique :

On fixe un paramètre $\alpha \in [0, 1]$ appelé *risque d'erreur de première espèce*. Celui-ci est la probabilité de rejeter à tort H_0 , autrement dit :

$$\alpha = \mathbf{P}[\text{rejet de } H_0 | H_0 \text{ est vraie}]$$

est la probabilité de rejeter l'hypothèse d'indépendance H_0 à la fin du test sachant que H_0 est vraie. Ce risque est souvent imposé par l'énoncé et est un choix de l'utilisateur du test sinon.

Connaissant α et ddl, on détermine une *valeur critique* $\chi_c^2 = \chi_{c,\alpha,\text{ddl}}^2$ dépendant de α et dd telle que si χ_2 suit la loi $\text{chi}_2(\text{ddl})$, on a :

$$\mathbf{P}[\chi^2 > \chi_c^2] = \alpha.$$

Ceci est fait soit par lecture dans une table du χ^2 , soit, à l'aide d'Excel, en appelant la fonction

$$\text{KHIDEUX.INVERSE}(\alpha; \text{ddl}).$$

Dans le contexte de l'Exemple 3.4, en choisissant le risque $\alpha = 1\% = 0,01$ et en utilisant que ddl=6, on obtient :

$$\chi_c^2 \simeq 16,8119.$$

Étape 6 : Règles de décision :

- Si $\chi_{\text{obs}}^2 > \chi_c^2$, on rejette H_0 ;
- Si $\chi_{\text{obs}}^2 \leq \chi_c^2$, on **ne** rejette **pas** H_0 .

Ces règles de décisions ce comprennent comme suit. La valeur du Khi-2 observée χ_{obs}^2 mesure une distance entre le tableau des effectifs observés et celui des effectifs théoriques sous l'hypothèse H_0 d'indépendance. Ainsi, si les deux tableaux sont significativement éloignés pour cette distance, on rejettera le fait qu'il correspondent à une même situation et donc l'hypothèse d'indépendance.

Dans le contexte de l'Exemple 3.4, les règles de décision sont :

- Si $\chi_{\text{obs}}^2 > \chi_c^2 \simeq 16,8119$, on rejette H_0 ;
- Si $\chi_{\text{obs}}^2 \leq \chi_c^2 \simeq 16,8119$, on **ne** rejette **pas** H_0 .

Étape 7 : Conclusion :

On conclue en utilisant les Étapes 3, 5 et 6.

Dans le contexte de l'Exemple 3.4, on a :

$$\chi_{\text{obs}}^2 \simeq 22,1063 > \chi_c^2 \simeq 16,8119$$

et on rejette H_0 au risque d'erreur de première espèce de 1%. Ainsi, on conclue que la campagne de publicité a été perçue de façon différente sur l'ensemble du territoire de Bourgogne/Franche-Comté au risque de première espèce de 1%.

Moins formellement, il y a moins d'un pourcent de chance d'avoir rejeter l'hypothèse d'indépendance alors qu'elle était vraie, ou encore de chercher à expliquer pourquoi la campagne de publicité aurait été perçue de façon différente sur le territoire alors que ce n'est pas le cas (et d'ons de risquer de fournir un travail supplémentaire inutile!).

Chapitre 4

Analyse à une variable réelle

4.1 Premières définitions

Définition 4.1. Une fonction (réelle) f associe à tout nombre réel x appartenant à un certain ensemble D_f , appelé ensemble de définition de f , un nombre réel $f(x)$, appelé image de x par f ou valeur de la fonction f en x .

Remarque 4.1. Dans les exemples qui nous intéresseront, les domaines de définition des fonctions seront des intervalles ou des réunions (finies) d'intervalles. Lorsque la fonction f admet une formulation explicite, sous une forme close, son domaine de définition est l'ensemble des réels pour lesquels cette expression est calculable.

Exemple 4.1. La fonction f définie par $f(x) = 2x + 1$ est définie sur \mathbf{R} ($D_f = \mathbf{R}$). La fonction g définie par $g(x) = \frac{1}{x}$ admet pour domaine de définition $D_g = \mathbf{R}^* = \mathbf{R} \setminus \{0\}$ puisque l'on ne peut pas diviser par 0.

Définition 4.2. Soit f une fonction définie sur D_f et $I \subset D_f$.

On dit que f est :

1. croissante sur I si, pour tous $x, y \in I$, $x < y$ implique que $f(x) \leq f(y)$;
2. strictement croissante sur I si, pour tous $x, y \in I$, $x < y$ implique que $f(x) < f(y)$;
3. décroissante sur I si, pour tous $x, y \in I$, $x < y$ implique que $f(x) \geq f(y)$;
4. strictement décroissante sur I si, pour tous $x, y \in I$, $x < y$ implique que $f(x) > f(y)$.

Exemple 4.1 (suite). La fonction f de l'Exemple 4.1 est strictement croissante sur \mathbf{R} et la fonction g , de ce même exemple, est strictement décroissante sur $] -\infty; 0[$ et sur $]0; +\infty[$ (mais pas sur \mathbf{R}^* !).

Définition 4.3. Soient f et g deux fonctions définies sur D_f et D_g respectivement telles que f prenne ses valeurs dans un ensemble $E \subset D_g$.

On appelle composée de f par g la fonction $g \circ f$ définie par :

$$(g \circ f)(x) = g(f(x)), \quad \text{pour tout } x \in D_f.$$

Définition 4.4. Soient f et g deux fonctions définies sur D_f et D_g respectivement.

On dit que f et g sont réciproques l'une de l'autre si

$$(g \circ f)(x) = x, \quad \text{pour tout } x \in D_f \quad \text{et} \quad (f \circ g)(y) = y, \quad \text{pour tout } y \in D_g.$$

On note alors $g = f^{-1}$.

Remarque 4.2.

1. Certaines fonctions n'admettent pas de réciproque. Par exemple, la fonction $x \mapsto x^2$ n'admet pas de réciproque sur \mathbf{R} ; par contre, sa restriction à $\mathbf{R}_+ = [0; +\infty[$ admet pour réciproque la fonction racine carrée $y \mapsto \sqrt{y}$.
2. Si une fonction est strictement monotone (*i.e.* strictement croissante ou strictement décroissante) alors elle admet une réciproque.
3. Si f et g sont réciproques l'une de l'autre, alors leurs graphes sont symétriques par rapport à la première bissectrice du plan (*i.e.* la droite d'équation $y = x$).

Exemple 4.1 (suite). La fonction f de l'Exemple 4.1 admet pour réciproque la fonction h définie par $h(y) = \frac{1}{2}y - \frac{1}{2}$. En effet, ces deux fonction sont définies sur \mathbf{R} et on a :

$$(h \circ f)(x) = h(f(x)) = \frac{1}{2}f(x) - \frac{1}{2} = \frac{1}{2}(2x + 1) - \frac{1}{2} = x, \quad \text{pour tout } x \in \mathbf{R}$$

et

$$(f \circ h)(y) = f(h(y)) = 2h(y) + 1 = 2\left(\frac{1}{2}y - \frac{1}{2}\right) + 1 = y, \quad \text{pour tout } y \in \mathbf{R}.$$

La fonction g (fonction *inverse*) de ce même exemple sa propre réciproque. Le vérifier en exercice.

4.2 Fonctions usuelles

4.2.1 Fonctions linéaires et affines

Définition 4.5. On appelle fonction linéaire toute fonction f s'écrivant sous la forme

$$f(x) = ax, \quad \text{pour un certain } a \in \mathbf{R}.$$

On appelle fonction affine toute fonction f s'écrivant sous la forme

$$f(x) = ax + b, \quad \text{pour certains } a, b \in \mathbf{R}.$$

Remarque 4.3.

1. Si une fonction est linéaire, elle est en particulier affine.
2. Une fonction affine est strictement croissante si, et seulement si, $a > 0$ et est strictement décroissante si, et seulement si, $a < 0$. Elle est constante si $a = 0$.
3. Le graphe d'une fonction affine est une droite dont le coefficient directeur est a et l'ordonnée à l'origine b .

4.2.2 Fonctions quadratiques

Définition 4.6. On appelle fonction quadratique toute fonction f s'exprimant sous la forme :

$$f(x) = ax^2 + bx + c, \quad \text{avec } a \in \mathbf{R}^* \text{ et } b, c \in \mathbf{R}.$$

Remarque 4.4.

1. Les fonctions quadratiques sont définies sur \mathbf{R} .
2. Une fonction quadratique n'est autre qu'un polynôme de degré 2.

Définition 4.7. On appelle discriminant du polynôme de degré 2 $ax^2 + bx + c$ la quantité :

$$\Delta = b^2 - 4ac.$$

La proposition suivante permet de déterminer les éventuelles racines de la fonction quadratique $f(x) = ax^2 + bx + c$, c'est-à-dire les éventuelles solutions de l'équation $ax^2 + bx + c = 0$.

Proposition 4.1. Soit $f(x) = ax^2 + bx + c$ un polynôme de degré 2.

1. Si $\Delta > 0$, alors f admet deux racines simples :

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

2. Si $\Delta = 0$, alors f admet une racine double :

$$x_1 = \frac{-b}{2a}.$$

3. Si $\Delta < 0$, alors f n'admet pas de racine.

Preuve : On a :

$$\begin{aligned} ax^2 + bx + c = 0 &\iff x^2 + \frac{b}{a}x + \frac{c}{a} = 0 \\ &\iff x^2 + 2\frac{b}{2a}x + \frac{c}{a} = 0 \\ &\iff x^2 + 2\frac{b}{2a}x + \frac{b^2}{4a^2} + \frac{c}{a} - \frac{b^2}{4a^2} = 0 \\ &\iff \left(x + \frac{b}{2a}\right)^2 + \frac{4ac - b^2}{4a^2} = 0. \end{aligned}$$

En posant $X = x + \frac{b}{2a}$, cette dernière équation est équivalente à $X^2 = \frac{\Delta}{4a^2}$.

Si $\Delta < 0$, il est clair que cette équation n'admet pas de solution puisque le carré d'un nombre réel est toujours positif. Ceci prouve le point 3..

Si $\Delta = 0$, l'équation $X^2 = \frac{\Delta}{4a^2} = 0$ admet pour unique solution $X = 0$ ce qui est équivalent à $x = \frac{-b}{2a}$. Ceci prouve le point 2..

Si $\Delta > 0$, l'équation $X^2 = \frac{\Delta}{4a^2} = 0$ admet pour solutions $X_1 = -\frac{\sqrt{\Delta}}{2a}$ et $X_2 = \frac{\sqrt{\Delta}}{2a}$. En utilisant que le changement de variable $X = x + \frac{b}{2a}$ est équivalent à $x = X - \frac{b}{2a}$, on obtient

que l'équation $ax^2 + bx + c$ admet pour solutions $x_1 = \frac{-b-\sqrt{\Delta}}{2a}$ et $x_2 = \frac{-b+\sqrt{\Delta}}{2a}$. Ceci prouve le point 1.. \square

La courbe représentative d'une fonction quadratique est une parabole dont les branches sont orientées vers le haut si $a > 0$ et vers le bas si $a < 0$. Il est symétrique par rapport à la droite d'équation $x = \frac{-b}{2a}$. La fonction $f(x) = ax^2 + bx + c$ admet un *minimum global* (resp. *maximum global*) en $x = \frac{-b}{2a}$ si $a > 0$ (resp. $a < 0$).

Exemple 4.2. Soit $f(x) = 3x^2 - 9x - 12$. Le coefficient du monôme de degré 2 est $a = 3 > 0$. La courbe représentative de f est donc une parabole dont les branches sont orientées vers le haut. Cette parabole est symétrique par rapport à la droite d'équation $x = \frac{-b}{2a} = \frac{-(-9)}{2 \times 3} = \frac{3}{2}$. La fonction f admet un minimum global en $x = \frac{-b}{2a} = \frac{3}{2}$. Son discriminant est :

$$\Delta = b^2 - 4ac = (-9)^2 - 4 \times 3 \times (-12) = 81 + 144 = 225 = 15^2 > 0.$$

Ses racines sont donc

$$x_1 = \frac{-b - \sqrt{\Delta}}{2a} = \frac{-(-9) - 15}{2 \times 3} = -1 \quad \text{et} \quad x_2 = \frac{-b + \sqrt{\Delta}}{2a} = \frac{-(-9) + 15}{2 \times 3} = 4.$$

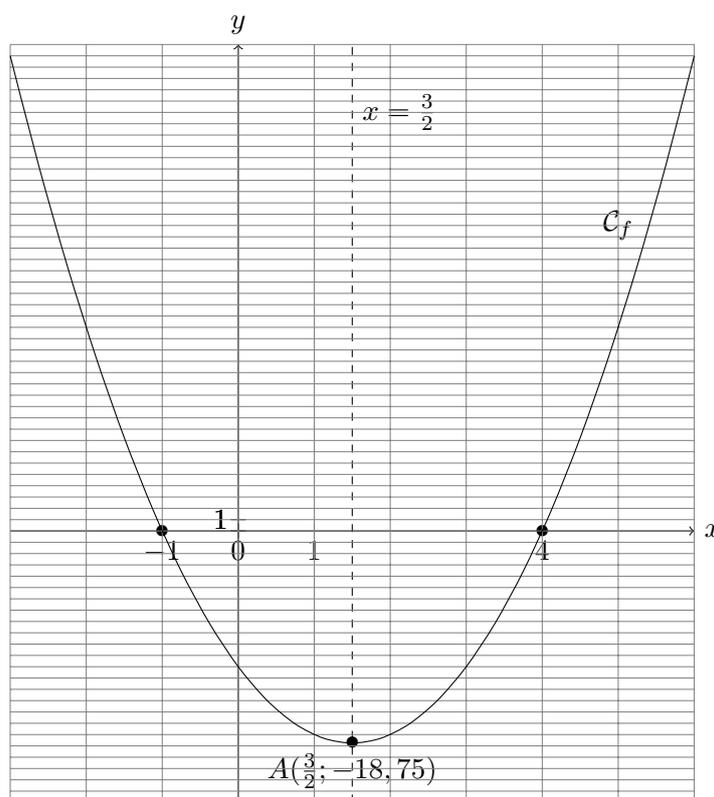


FIGURE 4.1 – Courbe représentative C_f de la fonction $f(x) = 3x^2 - 9x - 12$ et son axe de symétrie (pointillés). Cette fonction atteint son minimum $-18,75$ en $\frac{3}{2}$.

4.2.3 Fonctions polynomiales

Définition 4.8. On appelle fonction polynomiale, ou simplement polynôme, de degré $n \in \mathbf{N}^*$ toute fonction f s'écrivant sous la forme :

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,$$

avec $a_n \in \mathbf{R}^*$ et $a_{n-1}, \dots, a_0 \in \mathbf{R}$.

Remarque 4.5.

1. Les fonctions polynomiales sont définies sur \mathbf{R} .
2. Toute fonction polynomiale de degré n admet au plus n racines.

4.2.4 Fonction inverse

Définition 4.9. On appelle fonction inverse la fonction :

$$\begin{aligned} f : \mathbf{R}^* &\longrightarrow \mathbf{R}^*. \\ x &\longmapsto \frac{1}{x} \end{aligned}$$

La fonction inverse est décroissante sur chacun des intervalles $] -\infty, 0[$ et $]0, +\infty[$ et est sa propre réciproque. Son graphe est représenté dans la Figure 4.2.

4.2.5 Fonctions rationnelles

Définition 4.10. On appelle fonction rationnelle toute fonction f s'écrivant sous la forme :

$$f(x) = \frac{P(x)}{Q(x)}$$

où P et Q sont deux polynômes. Le domaine de définition d'une telle fonction est :

$$D_f = \{x \in \mathbf{R} : Q(x) \neq 0\}.$$

Exemple 4.3. Soit f la fonction définie par :

$$f(x) = \frac{x}{x^2 - 2x + 1}.$$

Cette fonction est une fonction rationnelle et son domaine de définition est :

$$\begin{aligned} D_f &= \{x \in \mathbf{R} : x^2 - 2x + 1 \neq 0\} \\ &= \{x \in \mathbf{R} : (x - 1)^2 \neq 0\} \\ &= \mathbf{R} \setminus \{1\}. \end{aligned}$$

Son graphe est donné dans la Figure 4.3.

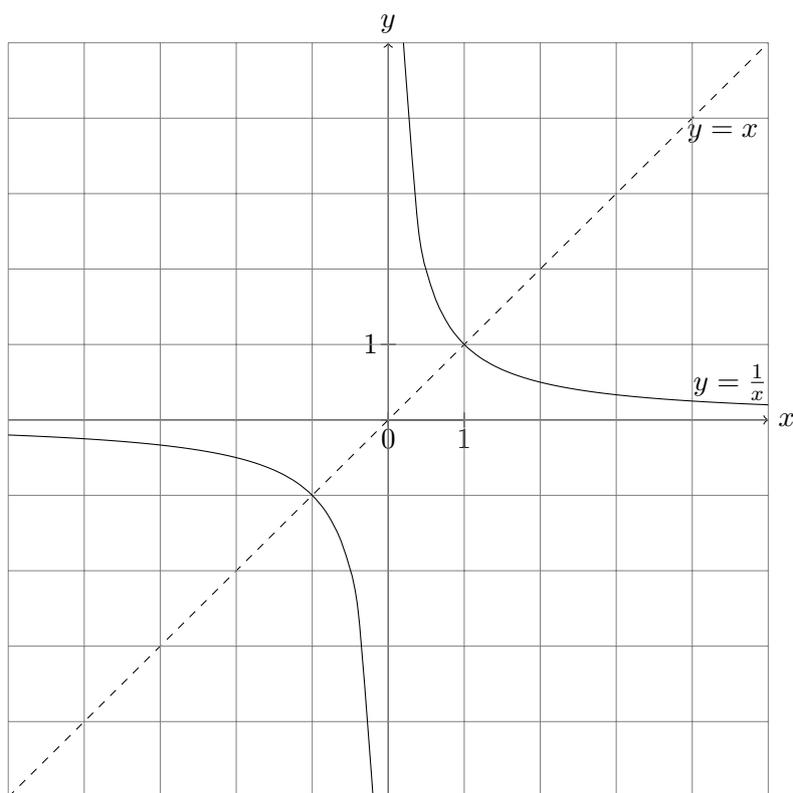


FIGURE 4.2 – Courbe représentative de la fonction inverse. Cette fonction étant sa propre réciproque, sa courbe représentative est symétrique par rapport à la première bissectrice du plan (pointillés).

4.2.6 Fonctions exponentielles

Définition[-Théorème] 4.11. Soit $a > 0$. Il existe une unique fonction $f : \mathbf{R} \mapsto \mathbf{R}_+^*$ continue en au moins un point, valant a en 1 et transformant la somme en produit, c'est-à-dire vérifiant :

$$f(x + y) = f(x) \times f(y), \quad \text{pour tous } x, y \in \mathbf{R}.$$

Cette fonction est appelée exponentielle de base a et est notée $\exp_a(\cdot)$ ou a^{\cdot} .

Proposition 4.2 (Admise). Soit $a > 0$.

1. $\exp_a(0) = 1$.
2. Si $0 < a < 1$, la fonction \exp_a est strictement décroissante sur \mathbf{R} ; si $a = 1$, \exp_1 est constante égale à 1 ; si $a > 1$, la fonction \exp_a est strictement croissante sur \mathbf{R} .
3. On a, pour tous $x, y \in \mathbf{R}$:

$$\exp_a(0) = a^0 = 1, \quad \exp_a(-x) = a^{-x} = \frac{1}{a^x} = \frac{1}{\exp_a(x)} \quad \text{et} \quad a^{xy} = (a^x)^y.$$

Parmi les fonction exponentielles, on distingue la fonction exponentielle de base e , où e est le réel vérifiant :

$$\lim_{h \rightarrow 0} \frac{e^h - 1}{h} = 1.$$

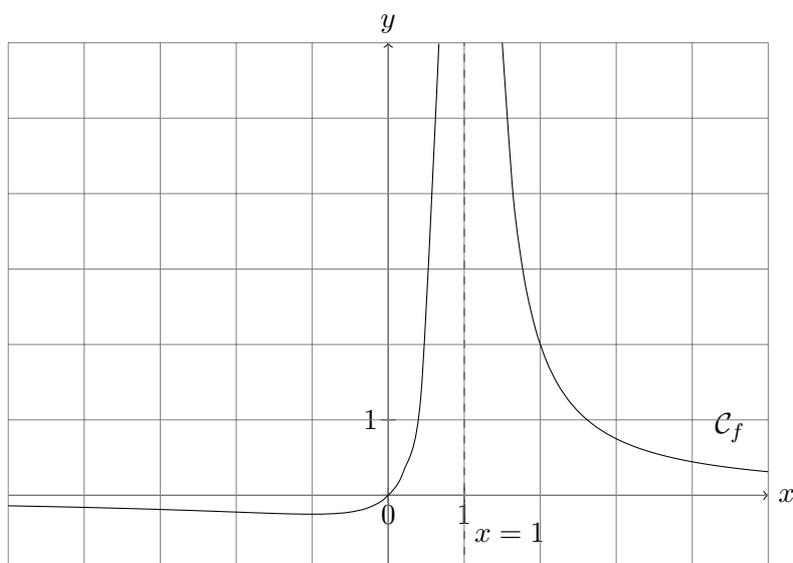


FIGURE 4.3 – Courbe représentative \mathcal{C}_f de la fonction $f(x) = \frac{x}{x^2 - 2x + 1}$. Celle-ci admet une asymptote verticale d'équation $x = 1$.

Cette fonction est appelée *fonction exponentielle* et est notée $\exp(\cdot)$ ou e^\cdot . Dans toute la suite, lorsque la base de la fonction exponentielle n'est pas précisée, il s'agit implicitement de la fonction exponentielle de base e .

Proposition 4.3 (Admise et rappel de la caractérisation). 1. La fonction \exp est strictement croissante sur \mathbf{R} .

2. L'image de \mathbf{R} par \exp est \mathbf{R}_+^* ; en particulier \exp est à valeurs strictement positives.

3. On a, pour tous $x, y \in \mathbf{R}$:

$$\exp(0) = 1, \quad \exp(-x) = \frac{1}{\exp(x)} \quad \text{et} \quad \exp(xy) = (\exp(x))^y.$$

et par définition $x, y \in \mathbf{R}$:

$$\exp(x + y) = \exp(x) \exp(y) \quad \text{d'où} \quad \exp(x - y) = \frac{\exp(x)}{\exp(y)}.$$

Remarque 4.6.

1. On a défini ici la fonction exponentielle par ces propriétés algébriques. Une définition alternative de cette fonction est la suivante. La fonction exponentielle (de base e) est l'unique fonction f de \mathbf{R} dans \mathbf{R} , valant 1 en 0, dérivable sur \mathbf{R} et étant sa propre dérivée (voir Section 4.5), c'est-à-dire vérifiant :

$$f'(x) = f(x), \quad \text{pour tout } x \in \mathbf{R}.$$

2. Le nombre e est irrationnel et vaut approximativement :

$$e \simeq 2,71828182845.$$

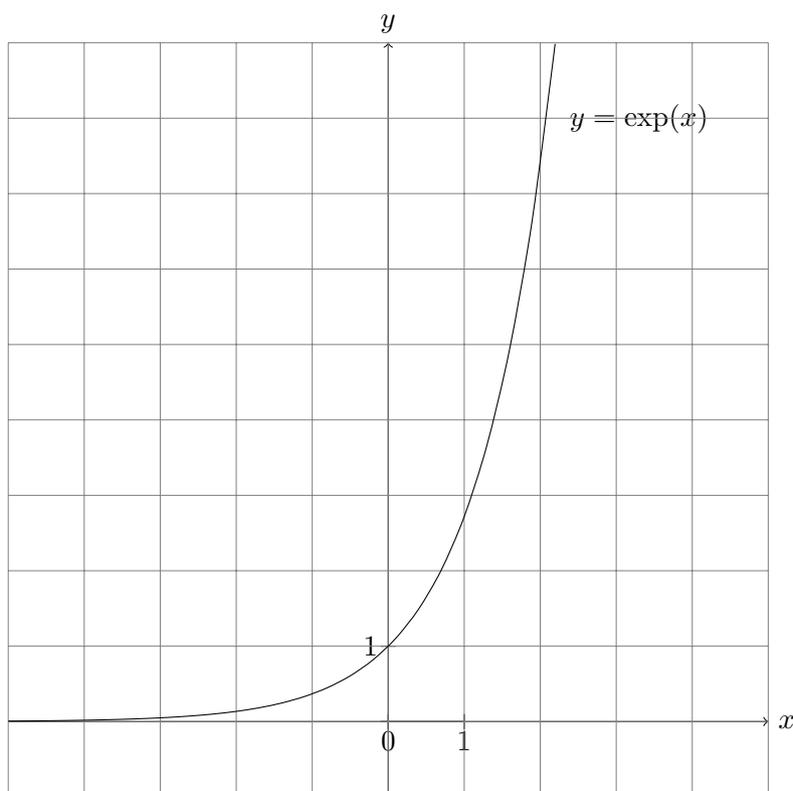


FIGURE 4.4 – Courbe représentative de la fonction exponentielle (de base e).

4.2.7 Fonctions logarithmes

Logarithme népérien

Définition[-Théorème] 4.12. *La fonction \exp admet une fonction réciproque définie sur \mathbf{R}_+^* et à valeurs dans \mathbf{R} . Cette fonction est appelée logarithme népérien et est notée $\ln(\cdot)$*

Remarque 4.7. En d'autres termes, \ln est la fonction telle que pour tout $x \in \mathbf{R}$ et tout $y \in \mathbf{R}_+^*$:

$$y = \exp(x) \iff x = \ln(y).$$

En particulier, $\ln(x) = 0$ si, et seulement si, $x = 1$.

Proposition 4.4. 1. *La fonction \ln est strictement croissante sur \mathbf{R}_+^* .*

2. *Pour tous $x, y \in \mathbf{R}_+^*$, on a :*

$$\ln(xy) = \ln(x) + \ln(y).$$

3. *Pour tous $x, y \in \mathbf{R}_+^*$, on a :*

$$\ln\left(\frac{x}{y}\right) = \ln(x) - \ln(y).$$

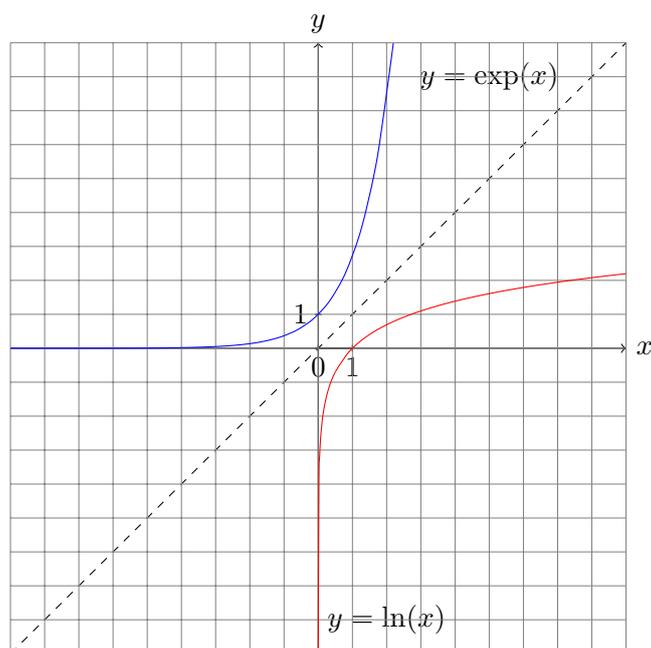


FIGURE 4.5 – Courbe représentative de la fonction logarithme népérien (rouge). Cette courbe est la symétrique par rapport à la première bissectrice du plan (pointillés) de celle de la fonction exponentielle (bleu).

4. Pour tout $x \in \mathbf{R}_+^*$ et tout $p \in \mathbf{R}$, on a :

$$\ln(x^p) = p \ln(x).$$

Preuve : Pour montrer que la fonction \ln est strictement croissante sur \mathbf{R}_+^* , il suffit de voir que si $0 < x < y$ alors $\ln(x) < \ln(y)$. Or puisque la fonction \exp est strictement croissante sur \mathbf{R} , pour $x, y \in \mathbf{R}_+^*$, on a :

$$\ln(x) < \ln(y) \iff \exp(\ln(x)) < \exp(\ln(y)) \iff x < y.$$

Ceci prouve le point 1..

On a :

$$\begin{aligned} z = \ln(x) + \ln(y) &\iff \exp(z) = \exp(\ln(x) + \ln(y)) \\ &\iff \exp(z) = \exp(\ln(x)) \exp(\ln(y)) = xy \\ &\iff z = \ln(xy). \end{aligned}$$

Ceci prouve le point 2. ; la preuve du point 3. est similaire. Passons à la preuve du point 4..
On a :

$$\begin{aligned} z = p \ln(x) &\iff \exp(z) = \exp(p \ln(x)) \\ &\iff \exp(z) = \exp(\ln(x))^p = x^p \\ &\iff z = \ln(x^p). \end{aligned}$$

□

Logarithme à base $a > 0, a \neq 1$

Définition 4.13. Soit $a > 0, a \neq 1$. On appelle logarithme à base a la fonction notée $\log_a(\cdot)$ et définie sur \mathbf{R}_+^* par :

$$\log_a(x) = \frac{\ln(x)}{\ln(a)}.$$

Proposition 4.5 (Admise). Soit $a > 0, a \neq 1$.

1. La fonction \log_a est la réciproque de la fonction exponentielle de base a . Elle est telle que, pour tout $x \in \mathbf{R}$ et tout $y \in \mathbf{R}_+^*$, on a :

$$y = \exp_a(x) \iff x = \log_a(y).$$

2. Pour tout $x \in \mathbf{R}$, on a $a^x = \exp_a(x) = \exp(x \ln(a))$.

3. Pour tous $x, y \in \mathbf{R}_+^*$, on a :

$$\log_a(xy) = \log_a(x) + \log_a(y).$$

4. Pour tous $x, y \in \mathbf{R}_+^*$, on a :

$$\log_a\left(\frac{x}{y}\right) = \log_a(x) - \log_a(y).$$

5. Pour tout $x \in \mathbf{R}_+^*$, on a :

$$\log_a(x^p) = p \log_a(x).$$

Exercice 4.1. Résoudre dans \mathbf{R} les équations suivantes :

1. $2 \ln(x+1) - \ln(2x) = 1$;
2. $2 \exp(2x) + \exp(x) = 3$.

Exercice 4.2. Résoudre dans \mathbf{R}^2 les systèmes suivants :

$$(S_1) \begin{cases} \exp(y-1) \exp(x-4) = 1 \\ \ln(3x) + \ln(y/3) = 1 \end{cases} \quad \text{et} \quad (S_2) \begin{cases} x + y = 10 \\ 2^x - 5^y = 0 \end{cases}.$$

4.2.8 Fonctions puissances et racines

Pour $k \in \mathbf{N}^*$ pair, la fonction *puissance* k^e $x \mapsto x^k$ est strictement croissante sur \mathbf{R}_+ et l'image de \mathbf{R}_+ par cette fonction est \mathbf{R}_+ lui-même. Sa réciproque sur \mathbf{R}_+ est appelée *racine* k^e et est notée $\sqrt[k]{\cdot}$. De la même façon, pour $k \in \mathbf{N}^*$ impair, la fonction *puissance* k^e $x \mapsto x^k$ est strictement croissante sur \mathbf{R} et l'image de \mathbf{R} par cette fonction est \mathbf{R} lui-même. Sa réciproque sur \mathbf{R} est appelée *racine* k^e et est notée $\sqrt[k]{\cdot}$.

Proposition 4.6. Soit $k \in \mathbf{N}^*$ et $x \in \mathbf{R}_+$.

On a :

$$\sqrt[k]{x} = x^{\frac{1}{k}}.$$

Preuve : Il suffit d'observer que $x \mapsto x^{\frac{1}{k}}$ est la réciproque de $x \mapsto x^k$. □

On a donc pu définir les fonctions puissance pour des exposants entiers naturels ou inverses d'entiers naturel. On peut aisément étendre ces définitions au cas d'exposants rationnels. Pour le cas des exposants arbitraires, on a recours aux fonctions exponentielle et logarithme.

Définition 4.14. Soit $\alpha \in \mathbf{R}$ et $x \in \mathbf{R}_+^*$. On définit la puissance α de x par :

$$x^\alpha = \exp(\alpha \ln(x)).$$

Remarque 4.8.

1. On peut vérifier que cette définition est cohérente avec les règles de calcul usuelles :

$$x^\alpha x^\beta = x^{\alpha+\beta}, \quad (x^\alpha)^\beta = x^{\alpha\beta} \quad \text{et} \quad x^{-\alpha} = \frac{1}{x^\alpha}.$$

2. La fonction $x \mapsto x^\alpha$ a un comportement très différent selon que $\alpha < 0$, $0 < \alpha < 1$ ou $\alpha > 1$ comme l'illustre la Figure 4.6.

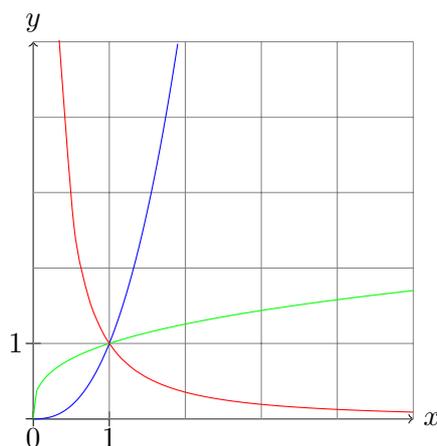


FIGURE 4.6 – Graphe de la fonction $x \mapsto x^\alpha$ pour $\alpha = -1,5 < 0$ (rouge), $\alpha = \frac{1}{3} \in]0; 1[$ (vert) et $\alpha = 2,5 > 1$ (bleu).

4.2.9 Fonctions indicatrices

Définition 4.15. Soit $A \subset \mathbf{R}$. On appelle fonction indicatrice de l'ensemble A la fonction notée $\mathbf{1}_A(\cdot)$ et définie sur \mathbf{R} par :

$$\mathbf{1}_A(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{sinon} \end{cases}.$$

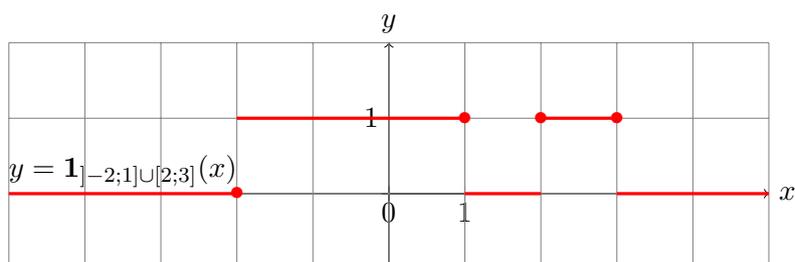


FIGURE 4.7 – Graphe de la fonction indicatrice de l'ensemble $A =]-2; 1] \cup [2; 3]$.

4.2.10 Fonction valeur absolue

Définition 4.16. On appelle (fonction) valeur absolue la fonction notée $|\cdot|$ et définie sur \mathbf{R} par :

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

Il est immédiat de vérifier que l'on a l'inégalité triangulaire suivante :

Proposition 4.7 (Inégalité triangulaire). Pour tous $x, y \in \mathbf{R}$, on a :

$$|x + y| \leq |x| + |y|.$$

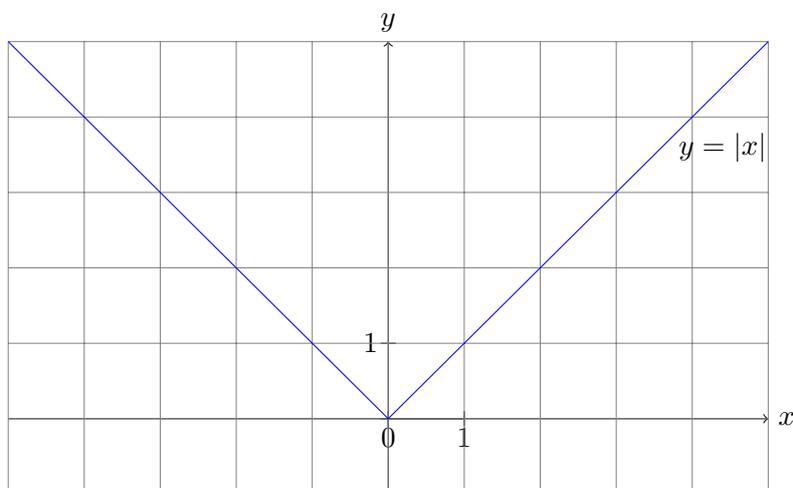


FIGURE 4.8 – Graphe de la fonction valeur absolue.

4.3 Limites

4.3.1 Définitions formelles

L'objectif de cette sous-section est de définir rigoureusement des expressions comme « $f(x)$ tend vers l (ou admet pour limite l) lorsque x tend vers a ». Cette sous-section peut être omise

en première lecture.

Limites infinies en l'infini

Supposons que f soit bien définie pour tout x suffisamment grand. On veut définir l'affirmation « $f(x)$ tend vers $\pm\infty$ lorsque x tend vers a ». Ceci se traduit intuitivement par le fait que pour tout seuil, grand dans les positifs (resp. les négatifs) $A > 0$ (resp. $-A$), les valeurs de $f(x)$ sont plus grandes que A (resp. plus petites que $-A$) dès que x est assez grand. Ceci conduit à la définition rigoureuse suivante.

Définition 4.17. *On dit que f tend vers $+\infty$ (resp. $-\infty$) en $+\infty$ si pour tout $A > 0$, il existe $B > 0$ tel que :*

$$x \geq B \implies f(x) \geq A \quad (\text{resp. } f(x) \leq -A).$$

On note alors :

$$\lim_{x \rightarrow +\infty} f(x) = +\infty \quad (\text{resp. } \lim_{x \rightarrow +\infty} f(x) = -\infty).$$

On définit de manière analogue les limites infinies en $-\infty$.

Définition 4.18. *On dit que f tend vers $+\infty$ (resp. $-\infty$) en $-\infty$ si pour tout $A > 0$, il existe $B > 0$ tel que :*

$$x \leq -B \implies f(x) \geq A \quad (\text{resp. } f(x) \leq -A).$$

On note alors :

$$\lim_{x \rightarrow -\infty} f(x) = +\infty \quad (\text{resp. } \lim_{x \rightarrow -\infty} f(x) = -\infty).$$

Exemple 4.4. Vérifions que $\lim_{x \rightarrow -\infty} x^4 = +\infty$. Soit $A > 0$. Posons $B = A^{\frac{1}{4}}$. On a :

$$\begin{aligned} x \leq -B &\implies x \leq -A^{\frac{1}{4}} \\ &\implies x \geq A \end{aligned}$$

Ainsi,

$$\lim_{x \rightarrow -\infty} x^4 = +\infty.$$

Limites finies en l'infini

Supposons que f soit bien définie pour tout x suffisamment grand. On veut définir proprement l'affirmation « $f(x)$ tend vers un réel l lorsque x tend vers $+\infty$ ». Ceci se traduit intuitivement par le fait que pour toute (petite) marge $\varepsilon > 0$, les valeurs de $f(x)$ sont au plus à une distance ε de l dès que x est assez grand. Ceci conduit à la définition rigoureuse suivante.

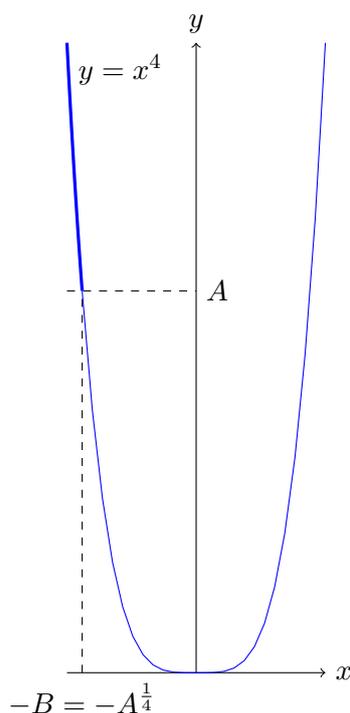


FIGURE 4.9 – Illustration du fait que $\lim_{x \rightarrow -\infty} x^4 = +\infty$. Pour tout $A > 0$, pour tout $x \leq -B = -A^{\frac{1}{4}}$, on a $x^4 \geq A$.

Définition 4.19. On dit que f admet pour limite l en $+\infty$ si pour tout $\varepsilon > 0$, il existe $A > 0$ tel que :

$$x \geq A \implies |f(x) - l| \leq \varepsilon.$$

On note alors :

$$\lim_{x \rightarrow +\infty} f(x) = l.$$

On définit de manière analogue les limites infinies en $-\infty$.

Définition 4.20. On dit que f admet pour limite l en $-\infty$ si pour tout $\varepsilon > 0$, il existe $A > 0$ tel que :

$$x \leq -A \implies |f(x) - l| \leq \varepsilon.$$

On note alors :

$$\lim_{x \rightarrow -\infty} f(x) = l.$$

Exemple 4.5. Vérifions que $\lim_{x \rightarrow +\infty} \frac{1}{\sqrt{x}} = 0$. Soit $\varepsilon > 0$. Posons $A = \varepsilon^{-2}$. On a :

$$\begin{aligned} x \geq A &\implies x \geq \varepsilon^{-2} \\ &\implies \sqrt{x} \geq \varepsilon^{-1} \\ &\implies \left| \frac{1}{\sqrt{x}} - 0 \right| = \frac{1}{\sqrt{x}} \leq \varepsilon. \end{aligned}$$

Ainsi,

$$\lim_{x \rightarrow +\infty} \frac{1}{\sqrt{x}} = 0.$$

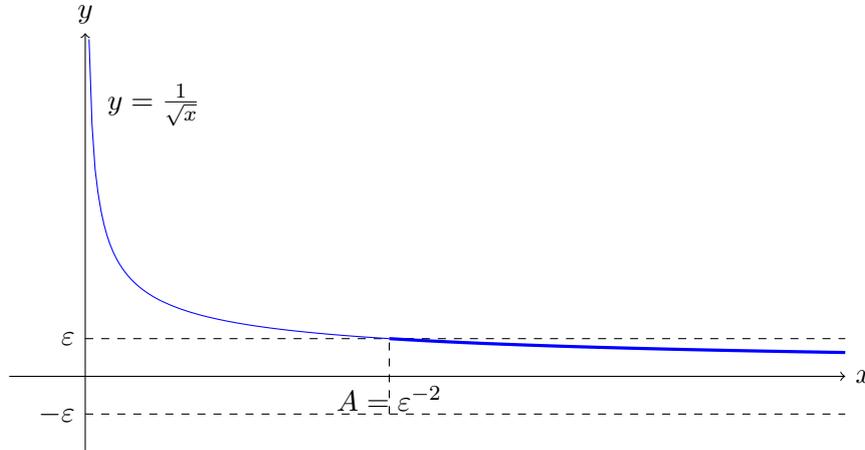


FIGURE 4.10 – Illustration du fait que $\lim_{x \rightarrow +\infty} \frac{1}{\sqrt{x}} = 0$. Pour tout $\varepsilon > 0$, pour tout $x \geq A = \varepsilon^{-2}$, on a $\left| \frac{1}{\sqrt{x}} - 0 \right| = \frac{1}{\sqrt{x}} \leq \varepsilon$.

Limites infinies en un point fini

Supposons que f soit bien définie pour tout x suffisamment proche de $a \in \mathbf{R}$, mais pas nécessairement en a . On veut définir proprement l'affirmation « $f(x)$ tend vers $\pm\infty$ lorsque x tend vers a ». Ceci se traduit intuitivement par le fait que pour tout seuil, grand dans les positifs (resp. les négatifs) $A > 0$ (resp. $-A$), les valeurs de $f(x)$ sont plus grandes que A (resp. plus petites que $-A$) dès que x est assez proche de a . Ceci conduit à la définition rigoureuse suivante.

Définition 4.21. On dit que f tend vers $+\infty$ (resp. $-\infty$) en $a \in \mathbf{R}$ si pour tout $A > 0$, il existe $\delta > 0$ tel que :

$$|x - a| \leq \delta \implies f(x) \geq A \quad (\text{resp. } f(x) \leq -A).$$

On note alors :

$$\lim_{x \rightarrow a} f(x) = +\infty \quad (\text{resp. } \lim_{x \rightarrow a} f(x) = -\infty).$$

Exemple 4.6. Vérifions que $\lim_{x \rightarrow 1} \frac{1}{x^2 - 2x + 1} = +\infty$. Soit $A > 0$. Posons $\varepsilon = \frac{1}{\sqrt{A}}$. On a :

$$\begin{aligned} |x - 1| \leq \varepsilon = \frac{1}{\sqrt{A}} &\implies 0 \leq (x - 1)^2 = x^2 - 2x + 1 \leq \frac{1}{A} \\ &\implies \frac{1}{x^2 - 2x + 1} \geq A \end{aligned}$$

Ainsi,

$$\lim_{x \rightarrow 1} \frac{1}{x^2 - 2x + 1} = +\infty.$$

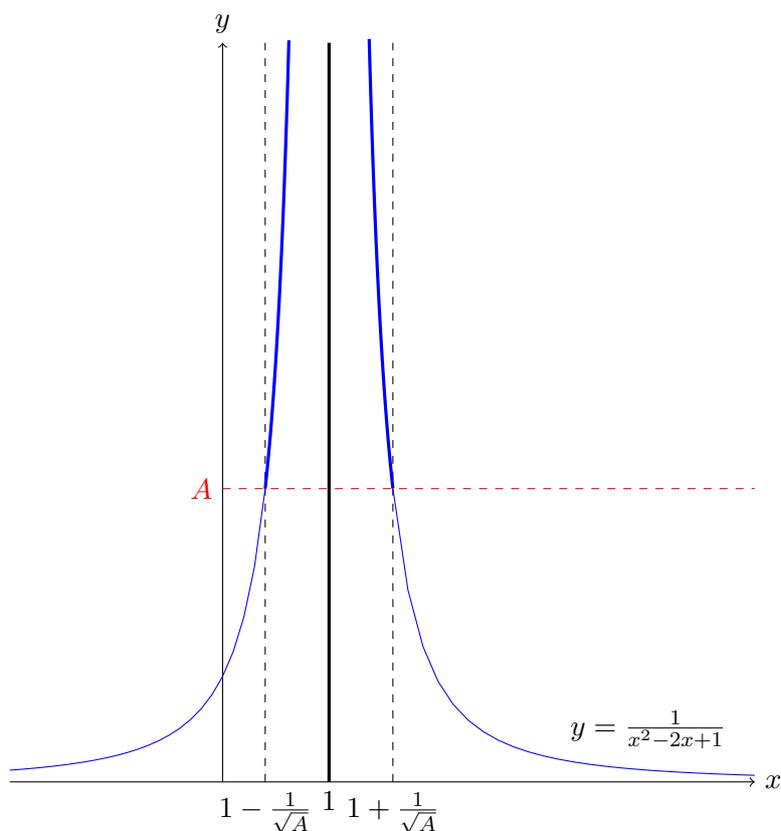


FIGURE 4.11 – Illustration du fait que $\lim_{x \rightarrow 1} \frac{1}{x^2 - 2x + 1} = +\infty$. Pour tout $A > 0$, pour tout x vérifiant $|x - 1| \leq \frac{1}{\sqrt{A}}$, on a $\frac{1}{x^2 - 2x + 1} \geq A$.

Limites finies en un point fini

Supposons que f soit bien définie pour tout x suffisamment proche de $a \in \mathbf{R}$, mais pas nécessairement en a . On veut définir proprement l'affirmation « $f(x)$ tend vers $l \in \mathbf{R}$ lorsque x tend vers a ». Ceci se traduit intuitivement par le fait que pour toute (petite) marge $\varepsilon > 0$, les valeurs de $f(x)$ sont au plus à une distance ε de l dès que x est assez proche de a . Ceci conduit à la définition rigoureuse suivante.

Définition 4.22. On dit que f admet pour limite l en a si pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que :

$$|x - a| \leq \delta \implies |f(x) - l| \leq \varepsilon.$$

On note alors :

$$\lim_{x \rightarrow a} f(x) = l.$$

Exercice 4.3. Vérifier, en utilisant la définition, que $\lim_{x \rightarrow 3} \frac{-x^2 + 6x - 9}{(x-5)(x-3)^2} = \frac{1}{2}$.

Limites à droite et à gauche

Certaines fonctions présentent des comportements différents à droite et à gauche de certains points comme l'illustre le cas de la fonction inverse au voisinage de 0 (voir Figure 4.2). Ce constat conduit à la définition des limites à droite et à gauche de fonction en un point.

Définition 4.23. On dit que f admet $l_+ \in \mathbf{R}$ pour limite à droite (resp. $l_- \in \mathbf{R}$ pour limite à gauche) en $a \in \mathbf{R}$ si pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que :

$$a < x \leq a + \delta \implies |f(x) - l_+| \leq \varepsilon \quad (\text{resp. } a - \delta \leq x < a \implies |f(x) - l_-| \leq \varepsilon).$$

On note alors :

$$\lim_{\substack{x \rightarrow a \\ x > a}} f(x) = l_+ \quad (\text{resp. } \lim_{\substack{x \rightarrow a \\ x < a}} f(x) = l_-).$$

Définition 4.24. On dit que f admet $+\infty$ pour limite à droite (resp. pour limite à gauche) en $a \in \mathbf{R}$ si pour tout $A > 0$, il existe $\delta > 0$ tel que :

$$a < x \leq a + \delta \implies f(x) \geq A \quad (\text{resp. } a - \delta \leq x < a \implies f(x) \geq A).$$

On note alors :

$$\lim_{\substack{x \rightarrow a \\ x > a}} f(x) = +\infty \quad (\text{resp. } \lim_{\substack{x \rightarrow a \\ x < a}} f(x) = +\infty).$$

Définition 4.25. On dit que f admet $-\infty$ pour limite à droite (resp. pour limite à gauche) en $a \in \mathbf{R}$ si pour tout $A > 0$, il existe $\delta > 0$ tel que :

$$a < x \leq a + \delta \implies f(x) \leq -A \quad (\text{resp. } a - \delta \leq x < a \implies f(x) \leq -A).$$

On note alors :

$$\lim_{\substack{x \rightarrow a \\ x > a}} f(x) = -\infty \quad (\text{resp. } \lim_{\substack{x \rightarrow a \\ x < a}} f(x) = -\infty).$$

Exemple 4.7. On peut voir que :

$$\lim_{\substack{x \rightarrow 0 \\ x > 0}} \frac{1}{x} = +\infty \quad \text{et} \quad \lim_{\substack{x \rightarrow 0 \\ x < 0}} \frac{1}{x} = -\infty$$

4.3.2 Limites usuelles

Dans cette sous-section sont collectées les limites de référence qu'il est indispensable de connaître.

Puissances entières et leurs inverses Pour $n \in \mathbf{N}^*$, on a :

$$\lim_{x \rightarrow +\infty} x^n = +\infty, \quad \lim_{x \rightarrow -\infty} x^n = \begin{cases} +\infty & \text{si } n \text{ est pair} \\ -\infty & \text{si } n \text{ est impair} \end{cases}, \quad \lim_{x \rightarrow \pm\infty} \frac{1}{x^n} = 0,$$

$$\lim_{x \rightarrow 0} \frac{1}{x^n} = +\infty \quad \text{si } n \text{ est pair,}$$

$$\lim_{\substack{x \rightarrow 0 \\ x > 0}} \frac{1}{x^n} = +\infty \quad \text{et} \quad \lim_{\substack{x \rightarrow 0 \\ x < 0}} \frac{1}{x^n} = -\infty \quad \text{si } n \text{ est impair.}$$

Puissances arbitraires Si $\alpha > 0$, alors

$$\lim_{x \rightarrow +\infty} x^\alpha = +\infty.$$

Si $\alpha < 0$, alors

$$\lim_{x \rightarrow +\infty} x^\alpha = 0 \quad \text{et} \quad \lim_{\substack{x \rightarrow 0 \\ x > 0}} x^\alpha = +\infty.$$

Exponentielles et logarithmes

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

$$\lim_{\substack{x \rightarrow 0 \\ x > 0}} \ln(x) = -\infty \quad \text{et} \quad \lim_{x \rightarrow +\infty} \ln(x) = +\infty.$$

4.3.3 Opérations sur les limites

La proposition suivante donne les règles de calcul permettant de déduire les limites de fonctions plus complexes à partir de limites de base.

Proposition 4.8. Soit $a, b \in \mathbf{R} \cup \{-\infty; +\infty\}$.

1. Si $\alpha \in \mathbf{R}$ et $\lim_{x \rightarrow a} f(x) = l \in \mathbf{R}$, alors $\lim_{x \rightarrow a} \alpha f(x) = \alpha l$.
2. Si $\alpha > 0$ et $\lim_{x \rightarrow a} f(x) = +\infty$ (resp. $-\infty$), alors $\lim_{x \rightarrow a} \alpha f(x) = +\infty$ (resp. $-\infty$).
3. Si $\alpha < 0$ et $\lim_{x \rightarrow a} f(x) = +\infty$ (resp. $-\infty$), alors $\lim_{x \rightarrow a} \alpha f(x) = -\infty$ (resp. $+\infty$).
4. Si $\lim_{x \rightarrow a} f(x) = l_1 \in \mathbf{R} \cup \{-\infty; +\infty\}$ et $\lim_{x \rightarrow a} g(x) = l_2 \in \mathbf{R} \cup \{-\infty; +\infty\}$, alors $\lim_{x \rightarrow a} f(x) + g(x) = l_1 + l_2$ avec $+\infty + \infty = +\infty$, $-\infty - \infty = -\infty$ et, pour $l \in \mathbf{R}$, $l + \infty = +\infty$ $l - \infty = -\infty$.
5. Si $\lim_{x \rightarrow a} f(x) = l_1 \in \mathbf{R}$ et $\lim_{x \rightarrow a} g(x) = l_2 \in \mathbf{R}$, alors $\lim_{x \rightarrow a} f(x)g(x) = l_1 l_2$.
6. Si $\lim_{x \rightarrow a} f(x) > 0$ et $\lim_{x \rightarrow a} g(x) = +\infty$ (resp. $-\infty$), alors $\lim_{x \rightarrow a} f(x)g(x) = +\infty$ (resp. $-\infty$).
7. Si $\lim_{x \rightarrow a} f(x) < 0$ et $\lim_{x \rightarrow a} g(x) = +\infty$ (resp. $-\infty$), alors $\lim_{x \rightarrow a} f(x)g(x) = -\infty$ (resp. $+\infty$).
8. Si $\lim_{x \rightarrow a} f(x) = l_1 \in \mathbf{R}$ et $\lim_{x \rightarrow a} g(x) = l_2 \in \mathbf{R}^*$, alors $\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \frac{l_1}{l_2}$.
9. Si $\lim_{x \rightarrow a} f(x) = l \in \mathbf{R}$ et $\lim_{x \rightarrow a} g(x) = \pm\infty$, alors $\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 0$.
10. Si $\lim_{x \rightarrow a} f(x) > 0$, $\lim_{x \rightarrow a} g(x) = 0$ et g est positive au voisinage de a (resp. négative au voisinage de a), alors $\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = +\infty$ (resp. $-\infty$).
11. Si $\lim_{x \rightarrow a} f(x) < 0$, $\lim_{x \rightarrow a} g(x) = 0$ et g est positive au voisinage de a (resp. négative au voisinage de a), alors $\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = -\infty$ (resp. $+\infty$).
12. Si $\lim_{x \rightarrow a} f(x) = b \in \mathbf{R} \cup \{-\infty; +\infty\}$ et $\lim_{x \rightarrow b} g(x) = l \in \mathbf{R} \cup \{-\infty; +\infty\}$, alors $\lim_{x \rightarrow a} (g \circ f)(x) = \lim_{x \rightarrow a} g(f(x)) = l$.

Remarque 4.9. On peut observer que, par exemple, on ne peut pas déterminer la limite de $f + g$ dans le point 4. lorsque $\lim_{x \rightarrow a} f(x) = -\infty$ et $\lim_{x \rightarrow a} g(x) = +\infty$. On parle dans

ce cas de forme indéterminée. D'autres formes indéterminées sont du type « $\frac{0}{0}$ », « $\frac{\pm\infty}{\pm\infty}$ » ou « $\pm\infty \times 0$ ».

Exemple 4.8. Déterminons la limite en $+\infty$ de la fonction exponentielle de base $a > 0$. Rappelons que :

$$\exp_a(x) = \exp(x \ln(a)).$$

Puisque $\ln(a) < 0$ si $0 < a < 1$ et $\ln(a) > 0$ si $a > 1$, on a :

$$\lim_{x \rightarrow +\infty} x \ln(a) = -\infty \quad \text{si } 0 < a < 1$$

et

$$\lim_{x \rightarrow +\infty} x \ln(a) = +\infty \quad \text{si } a > 1.$$

Or, $\lim_{y \rightarrow -\infty} \exp(y) = 0$ et $\lim_{y \rightarrow +\infty} \exp(y) = +\infty$. On en déduit, par composition, que :

$$\lim_{x \rightarrow +\infty} \exp_a(x) = \lim_{x \rightarrow +\infty} \exp(x \ln(a)) = 0 \quad \text{si } 0 < a < 1$$

et

$$\lim_{x \rightarrow +\infty} \exp_a(x) = \lim_{x \rightarrow +\infty} \exp(x \ln(a)) = +\infty \quad \text{si } a > 1.$$

La limite de la fonction $x \mapsto \exp(x) - x^2$ en $+\infty$ est une forme indéterminée. Nous verrons dans la Section 4.3.5 comment lever de telles indéterminations.

Exercice 4.4. Déterminer, pour $a > 0$ la limite en $+\infty$ de la fonction logarithme à base a .

4.3.4 Comparaison de limites

Le théorème de comparaison de limites suivant se déduit directement des définitions.

Théorème 4.1. Soit $a \in \mathbf{R} \cup \{-\infty; +\infty\}$.

1. Si $f(x) \leq g(x)$ dans un voisinage de a et $\lim_{x \rightarrow a} f(x) = +\infty$, alors $\lim_{x \rightarrow a} g(x) = +\infty$.
2. Si $f(x) \leq g(x)$ dans un voisinage de a et $\lim_{x \rightarrow a} g(x) = -\infty$, alors $\lim_{x \rightarrow a} f(x) = -\infty$.
3. Si $f(x) \leq g(x) \leq h(x)$ dans un voisinage de a et $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} h(x) = l \in \mathbf{R}$, alors $\lim_{x \rightarrow a} g(x) = l$.

Preuve : Exercice. □

Exemple 4.9. Rappelons que pour tout réel x , $-1 \leq \cos(x) \leq 1$. On a donc, pour tout réel $x > 0$ que :

$$-\frac{1}{x} \leq \frac{\cos(x)}{x} \leq \frac{1}{x}.$$

On déduit du point 3. et du fait que $\lim_{x \rightarrow +\infty} -\frac{1}{x} = \lim_{x \rightarrow +\infty} \frac{1}{x} = 0$ que :

$$\lim_{x \rightarrow +\infty} \frac{\cos(x)}{x} = 0.$$

4.3.5 Formes indéterminées

Cette sous-section a pour but de fournir des outils pour lever les indéterminations dans le calcul de limites. Nous commençons par comparer les croissances de fonctions usuelles.

Croissances comparées

Proposition 4.9 (Admise). *On a :*

$$\lim_{x \rightarrow +\infty} \frac{\exp(x)}{x} = +\infty, \quad \lim_{x \rightarrow -\infty} x \exp(x) = 0,$$

$$\lim_{x \rightarrow +\infty} \frac{\ln(x)}{x} = 0 \quad \text{et} \quad \lim_{\substack{x \rightarrow 0 \\ x > 0}} x \ln(x) = 0.$$

Plus généralement, on a, pour tout $\alpha > 0$ et tout $\beta \in \mathbf{R}$:

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

$$\lim_{x \rightarrow +\infty} \frac{(\ln(x))^\beta}{x^\alpha} = 0 \quad \text{et} \quad \lim_{\substack{x \rightarrow 0 \\ x > 0}} x^\alpha |\ln(x)|^\beta = 0.$$

Remarque 4.10. Informellement, il faut retenir que :

1. quand x tend vers $+\infty$, $\exp(x)$ tend plus vite vers l'infini que n'importe quelle puissance positive de x et $\ln(x)$ tend moins vite vers l'infini que n'importe quelle puissance positive de x ;
2. quand x tend vers $-\infty$, $\exp(x)$ tend plus vite vers 0 que n'importe quelle puissance négative de x ;
3. quand x tend vers 0 par valeurs positives, $|\ln(x)|$ tend moins vite vers $+\infty$ que n'importe quelle puissance négative de x .

Méthodes

Indéterminations du type « $+\infty - \infty$ » Il convient dans ce cas de factoriser l'expression par le terme ayant la croissance la plus rapide vers l'infini.

Exemple 4.10. Déterminons la limite en $+\infty$ de la fonction $x \mapsto \exp(x) - x^3$. Le terme dominant est $\exp(x)$. On a donc :

$$\lim_{x \rightarrow +\infty} \exp(x) - x^3 = \lim_{x \rightarrow +\infty} \exp(x) \left(1 - \frac{x^3}{\exp(x)} \right).$$

Or, $\lim_{x \rightarrow +\infty} \frac{x^3}{\exp(x)} = 0$ donc $\lim_{x \rightarrow +\infty} 1 - \frac{x^3}{\exp(x)} = 1$ et $\lim_{x \rightarrow +\infty} \exp(x) = +\infty$. On en déduit que :

$$\lim_{x \rightarrow +\infty} \exp(x) - x^3 = \lim_{x \rightarrow +\infty} \exp(x) \left(1 - \frac{x^3}{\exp(x)} \right) = +\infty.$$

Cette méthode permet d'affirmer que la limite en $\pm\infty$ de tout polynôme est celle de son monôme de plus haut degré.

Proposition 4.10. Soit $a_n \in \mathbf{R}^*$ et $a_{n-1}, \dots, a_0 \in \mathbf{R}$.

On a :

$$\lim_{x \rightarrow \pm\infty} a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 = \lim_{x \rightarrow \pm\infty} a_n x^n.$$

Preuve : On a :

$$a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 = a_n x^n \left(1 + \frac{a_{n-1}}{a_n x} + \dots + \frac{a_1}{a_n x^{n-1}} + \frac{a_0}{a_n x^n} \right).$$

Or, chacun des termes $\frac{a_{n-1}}{a_n x}, \dots, \frac{a_0}{a_n x^n}$ tend vers 0 quand x tend vers $\pm\infty$. D'où le résultat. \square

Indéterminations du type « $\frac{\pm\infty}{\pm\infty}$ » Il convient dans ce cas de factoriser le numérateur et le dénominateur de l'expression par le terme ayant la croissance la plus rapide vers l'infini et d'utiliser les résultats de croissances comparées.

Exemple 4.11. Déterminons la limite en $-\infty$ de la fonction $x \mapsto \frac{3x^3 - 2x + 7}{2x^2 + 5x - 1}$. Au numérateur, le terme dominant est celui du 3^e degré alors qu'au dénominateur, il s'agit de celui du 2nd degré. On peut donc écrire que :

$$\frac{3x^3 - 2x + 7}{2x^2 + 5x - 1} = \frac{x^3 \left(3 - \frac{2}{x^2} + \frac{7}{x^3} \right)}{x^2 \left(2 + \frac{5}{x} - \frac{1}{x^2} \right)} = x \frac{3 - \frac{2}{x^2} + \frac{7}{x^3}}{2 + \frac{5}{x} - \frac{1}{x^2}}.$$

On en déduit que :

$$\begin{aligned} \lim_{x \rightarrow -\infty} \frac{3x^3 - 2x + 7}{2x^2 + 5x - 1} &= \lim_{x \rightarrow -\infty} \left(x \frac{3 - \frac{2}{x^2} + \frac{7}{x^3}}{2 + \frac{5}{x} - \frac{1}{x^2}} \right) \\ &= \lim_{x \rightarrow -\infty} x \times \lim_{x \rightarrow -\infty} \frac{3 - \frac{2}{x^2} + \frac{7}{x^3}}{2 + \frac{5}{x} - \frac{1}{x^2}} = -\infty \times \frac{3}{2} = -\infty. \end{aligned}$$

Remarque 4.11. Ceci permet entre autre de déterminer les limites en $\pm\infty$ de toute fonction rationnelle.

Indéterminations du type « $\frac{0}{0}$ » ou « $0 \times (\pm\infty)$ » En général, l'idée est de remarquer que ce type d'indéterminations est équivalent à celles du type « $\frac{\pm\infty}{\pm\infty}$ » à un jeu d'écriture près. En effet, si la limite de $f \times g$ conduit à une indétermination du type « $0 \times \pm\infty$ », elle est la même que celle de $\frac{g}{(1/f)}$ que nous savons déjà lever, puisqu'elle est du type « $\frac{\pm\infty}{\pm\infty}$ ». De même, une indétermination du type « $\frac{0}{0}$ » se ramène aisément à une indétermination du type « $\frac{\pm\infty}{\pm\infty}$ » par passage à l'inverse au numérateur et au dénominateur.

Exemple 4.12. Déterminons la limite de $x \mapsto \exp(-2x) (2x^5 + 5x^2 - 9)$ lorsque x tend

vers $+\infty$. On a :

$$\begin{aligned} \lim_{x \rightarrow +\infty} \exp(-2x) (2x^5 + 5x^2 - 9) &= \lim_{x \rightarrow +\infty} \frac{1}{\exp(2x)} (2x^5 + 5x^2 - 9) \\ &= \lim_{x \rightarrow +\infty} \frac{2x^5 + 5x^2 - 9}{\exp(2x)} \\ &= \lim_{x \rightarrow +\infty} \frac{x^5}{\exp(2x)} \left(2 + \frac{5}{x^3} - \frac{9}{x^5} \right) \\ &= 2 \lim_{x \rightarrow +\infty} \frac{x^5}{\exp(2x)} \\ &= 0. \end{aligned}$$

4.4 Continuité

Définition 4.26. On dit que la fonction f définie au voisinage de x_0 est continue en x_0 si, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que :

$$|x - x_0| \leq \delta \implies |f(x) - f(x_0)| \leq \varepsilon.$$

On dit que f est continue sur $I \subset \mathbf{R}$ si elle est continue en tout point de I .

Remarque 4.12. Intuitivement, dire que f est continue sur I revient à dire que l'on peut tracer le graphe de sa restriction à I « sans lever le crayon ».

Exemple 4.13. Les fonctions polynomiales, rationnelles, exponentielles, logarithmes et valeur absolue sont continues sur leurs ensembles de définition respectifs. Par contre, les fonctions indicatrices ne sont, en général, pas continues sur leur ensemble de définition (\mathbf{R}). Par exemple, $\mathbf{1}_{[a;b]}$ présente des discontinuités en a et b .

Proposition 4.11. La somme, le produit, le quotient et la composée de fonctions continues sont continues sur leurs ensembles de définition.

Les théorèmes suivants sont des théorèmes fondamentaux portant sur les fonctions continues sur des intervalles fermés.

Théorème 4.2 (Théorème de Bolzano). Soit $f : [a; b] \longrightarrow \mathbf{R}$ une fonction continue sur $[a; b]$.

Si $f(a) < 0$ et $f(b) > 0$, alors l'équation $f(x) = 0$ admet au moins une solution dans l'intervalle $[a; b]$.

Si de plus, f est strictement croissante sur $[a; b]$, cette solution est unique.

Théorème 4.3 (Théorème des valeurs intermédiaires). Soit $f : [a; b] \longrightarrow \mathbf{R}$ une fonction continue sur $[a; b]$.

1. Pour tout γ tel que $f(a) < \gamma < f(b)$, l'équation $f(x) = \gamma$ admet au moins une solution c dans l'intervalle $[a; b]$.

Si de plus, f est strictement croissante sur $[a; b]$, cette solution est unique.

2. Pour tout γ tel que $f(b) < \gamma < f(a)$, l'équation $f(x) = \gamma$ admet au moins une solution c dans l'intervalle $[a; b]$.

Si de plus, f est strictement décroissante sur $[a; b]$, cette solution est unique.

Exercice 4.5. Vérifier que l'équation $\exp(2x + 1) = 2$ admet une unique solution dans l'intervalle $[-1; 3]$.

4.5 Dérivabilité

4.5.1 Définitions

Définition 4.27. Soit f une fonction continue au voisinage de x_0 .

On appelle taux de variation de f en x_0 la fonction définie, pour $|h|$ suffisamment petit, par :

$$(\tau_f(x_0))(h) = \frac{f(x_0 + h) - f(x_0)}{h}.$$

Définition 4.28. Soit f une fonction continue au voisinage de x_0 .

Si, $\lim_{h \rightarrow 0} (\tau_f(x_0))(h)$ existe et est finie, on appelle nombre dérivé de f en x_0 la quantité :

$$f'(x_0) = \lim_{h \rightarrow 0} (\tau_f(x_0))(h).$$

On dit dans ce cas que f est dérivable en x_0 .

Définition 4.29. Soit f une fonction définie sur D_f et $I \subset D_f$.

On dit que f est dérivable sur I si f est dérivable en tout point de I et on appelle ensemble de dérivabilité de f le plus grand sous-ensemble de D_f sur lequel f est dérivable.

Remarque 4.13. Dans certains contextes, une notation s'impose plus naturellement que x pour la variable et on peut étudier une fonction f d'une variable s, t, y ou autre. Lorsque l'on veut insister sur la variable par rapport à laquelle on dérive, disons s , on utilise la notation :

$$\frac{d}{ds} f(s)$$

au lieu de $f'(s)$ pour la dérivée de f en s . Si aucune ambiguïté n'est possible, on utilise la notation habituelle $f'(s)$. Dans tous les cas, les aspects techniques restent inchangés. Ceci est illustré par l'exemple suivant.

Exemple 4.14. En économie, la consommation C d'un ménage est donnée par :

$$C = C_0 + c(Y - T),$$

avec C_0 la consommation incompressible, $c \in [0, 1]$ la propension marginale à consommer (part du revenu disponible consacré à la consommation), Y le revenu (Yield en anglais) du ménage et $T = tY$ le montant des taxes où $t \in [0, 1[$ désigne le taux d'imposition.

Ainsi, on peut voir la consommation du ménage comme une fonction C de son revenu Y et étudier la fonction :

$$C : Y \mapsto C(Y) = C_0 + c(Y - T) = C_0 + c(Y - tY) = C_0 + c(1 - t)Y.$$

Insistons sur le fait qu'ici la seule variable est Y , les lettres c, T et t désignent des constantes dans ce modèle. En utilisant les notions rappelées ci-dessous mais déjà connues du lecteur, on obtient que :

$$\frac{d}{dY} C(Y) = C'(Y) = c(1 - t) \geq 0.$$

On en déduit que la consommation du ménage croît avec son revenu ; plus précisément, on en déduit que si le revenu du ménage croît d'une unité, sa consommation croît de $c(1-t)$. Notons que cette conclusion donne la valeur exacte de l'accroissement de la consommation engendrée par la hausse du revenu d'une unité puisque C est une fonction *affine* du revenu.

Si C n'était pas affine en le revenu, on pourrait obtenir une approximation de cette augmentation en utilisant, par exemple, la tangente à la courbe de C en le point adéquat (voir Section 4.5.4). Une alternative parfois utilisée en économie est de considérer la notion de *dérivée discrète* ; pour cela on calcule :

$$\frac{\Delta C(Y)}{\Delta Y} := \frac{C(Y+1) - C(Y)}{Y+1 - Y} = C(Y+1) - C(Y).$$

La valeur de l'accroissement de la consommation engendrée par la hausse du revenu d'une unité est alors exacte, que C soit une fonction affine ou non. Il convient d'être vigilants à ne pas confondre les notions de dérivée et de dérivée discrète, proches mais distinctes.

4.5.2 Dérivées usuelles

Le tableau suivant récapitule les dérivées usuelles et domaines de dérivabilité des fonctions.

Fonction $f(x)$	Domaine de définition	Domaine de dérivabilité	Dérivée $f'(x)$
k constante	\mathbf{R}	\mathbf{R}	0
$x^n, n \in \mathbf{N}$	\mathbf{R}	\mathbf{R}	nx^{n-1}
$\frac{1}{x^n}, n \in \mathbf{N}$	\mathbf{R}^*	\mathbf{R}^*	$-\frac{n}{x^{n+1}}$
$x^\alpha, \alpha > 0$	\mathbf{R}_+	\mathbf{R}_+^*	$\alpha x^{\alpha-1}$
$x^\alpha, \alpha < 0$	\mathbf{R}_+^*	\mathbf{R}_+^*	$\alpha x^{\alpha-1}$
$\exp(x)$	\mathbf{R}	\mathbf{R}	$\exp(x)$
$\ln(x)$	\mathbf{R}_+^*	\mathbf{R}_+^*	$\frac{1}{x}$
$\ln x $	\mathbf{R}^*	\mathbf{R}^*	$\frac{1}{x}$

4.5.3 Opérations sur les dérivées

Proposition 4.12 (Admise). 1. Soient $\alpha \in \mathbf{R}$ et u dérivable en x . Alors, la fonction αu est dérivable en x et on a :

$$(\alpha u)'(x) = \alpha u'(x).$$

2. Soient u et v dérivables en x . Alors, les fonctions $u+v$ et $u \times v$ sont dérivables en x et on a :

$$(u+v)'(x) = u'(x) + v'(x) \quad \text{et} \quad (u \times v)'(x) = u'(x)v(x) + u(x)v'(x).$$

3. Soient u et v dérivables en x et v telle que $v(x) \neq 0$. Alors, la fonction $\frac{u}{v}$ est dérivable en x et on a :

$$\left(\frac{u}{v}\right)'(x) = \frac{u'(x)v(x) - u(x)v'(x)}{(v(x))^2}.$$

En particulier,

$$\left(\frac{1}{v}\right)'(x) = -\frac{v'(x)}{(v(x))^2}.$$

4. Soient u dérivable en x et v dérivable en $u(x)$. Alors, la fonction $v \circ u$ est dérivable en x et on a :

$$(v \circ u)'(x) = u'(x)v'(u(x)).$$

En particulier, si u est dérivable en x ,

$$(\exp \circ u)'(x) = (\exp(u(x)))' = u'(x) \exp(u(x)),$$

et si de plus $u(x) > 0$,

$$(\ln \circ u)'(x) = (\ln(u(x)))' = \frac{u'(x)}{u(x)}.$$

Exemple 4.15.

1. La fonction $f : x \mapsto x^2 \exp(x)$ est dérivable sur \mathbf{R} . En posant $u(x) = x^2$ et $v(x) = \exp(x)$, on obtient que $u'(x) = 2x$ et $v'(x) = \exp(x)$. On en déduit que, pour tout $x \in \mathbf{R}$:

$$f'(x) = u'(x)v(x) + u(x)v'(x) = 2x \exp(x) + x^2 \exp(x) = \exp(x)x(x+2).$$

2. La fonction rationnelle $g : x \mapsto \frac{x^3-1}{x-2}$ est dérivable sur $\mathbf{R} \setminus \{2\}$. En posant $u(x) = x^3 - 1$ et $v(x) = x - 2$, on obtient que $u'(x) = 3x^2$ et $v'(x) = 1$, puis que :

$$g'(x) = \frac{u'(x)v(x) - u(x)v'(x)}{(v(x))^2} = \frac{3x^2(x-2) - 1 \times (x^3-1)}{(x-2)^2} = \frac{2x^3 - 6x^2 + 1}{(x-2)^2}.$$

3. La fonction $h : x \mapsto \ln(x^2 - 1)$ est dérivable sur $] - \infty; -1[\cup]1; +\infty[$. En posant $u(x) = x^2 - 1$ et $v(x) = \ln(x)$, on obtient que $u'(x) = 2x$ et $v'(x) = \frac{1}{x}$. On en déduit que :

$$h'(x) = u'(x)v'(u(x)) = 2x \times \frac{1}{x^2-1} = \frac{2x}{x^2-1}.$$

4.5.4 Tangente à une courbe en un point

Définition 4.30. Soit f une fonction dérivable en x_0 .

La tangente à la courbe représentative de f en x_0 est la droite d'équation :

$$T_{f,x_0} : y = f'(x_0)(x - x_0) + f(x_0).$$

Remarque 4.14. Il s'agit de la meilleure approximation locale de la fonction f par une fonction affine au voisinage de x_0 .

Exemple 4.16. Reprenons la fonction $g : x \mapsto \frac{x^3-1}{x-2}$ de l'Exemple 4.15. On a vu que :

$$g'(x) = \frac{2x^3 - 6x^2 + 1}{(x-2)^2}, \quad \text{pour tout } x \in \mathbf{R} \setminus \{2\}.$$

En particulier, $g'(0) = \frac{2 \times 0^3 - 6 \times 0^2 + 1}{(0-2)^2} = \frac{1}{4}$ et $g(0) = \frac{0^3-1}{0-2} = \frac{1}{2}$. L'équation de la tangente à la courbe représentative de g (voir Figure 4.12) est donc :

$$T_{g,0} : y = g'(0)(x - 0) + g(0) = \frac{1}{4}x + \frac{1}{2}.$$

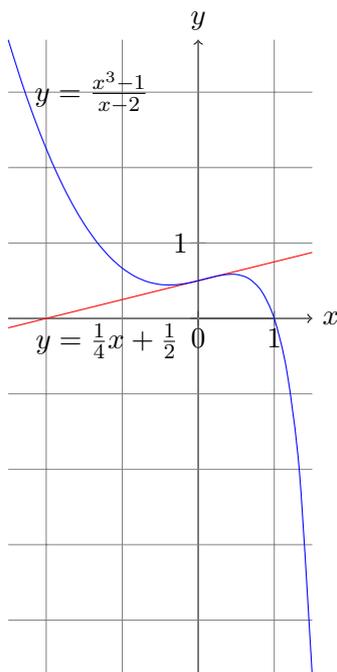


FIGURE 4.12 – Courbe représentative de la fonction $g : x \mapsto \frac{x^3-1}{x-2}$ (bleu) et sa tangente en 0, d'équation $y = \frac{1}{4}x + \frac{1}{2}$ (rouge). Cette dernière est la meilleure approximation locale de la fonction g par une fonction affine au voisinage de 0.

4.5.5 Applications à l'étude des variations et à la recherche d'extrema

Définition 4.31. Soit $E \subset \mathbf{R}$. On dit que x_0 est un point intérieur de E s'il existe un intervalle de la forme $[x_0 - \alpha; x_0 + \alpha]$, $\alpha > 0$, inclus dans E .

Définition 4.32. Soit f une fonction et D_f son ensemble de définition.

1. On dit qu'une fonction f admet un maximum local (resp. minimum local) en un point intérieur x_0 de D_f s'il existe un intervalle $I = [x_0 - \beta; x_0 + \beta] \subset D_f$, $\beta > 0$, tel que pour tout $x \in I$, $f(x_0) \geq f(x)$ (resp. $x \in I$, $f(x_0) \leq f(x)$).
2. On dit qu'une fonction f admet un maximum local (resp. minimum local) en un point x_0 sur le bord droit de D_f s'il existe un intervalle $I = [x_0; x_0 + \beta] \subset D_f$, $\beta > 0$, tel que pour tout $x \in I$, $f(x_0) \geq f(x)$ (resp. $x \in I$, $f(x_0) \leq f(x)$).
3. On dit qu'une fonction f admet un maximum local (resp. minimum local) en un point x_0 sur le bord gauche de D_f s'il existe un intervalle $I = [x_0 - \beta; x_0] \subset D_f$, $\beta > 0$, tel que pour tout $x \in I$, $f(x_0) \geq f(x)$ (resp. $x \in I$, $f(x_0) \leq f(x)$).

Remarque 4.15.

1. Dans tous les cas on parle d'*extremum local*.
2. Si l'inégalité est stricte pour tout $x \in I \setminus \{x_0\}$, on parle d'*extremum local strict*.

La proposition suivante donne une condition nécessaire pour qu'une fonction dérivable admette un extremum en x_0 .

Proposition 4.13. *Soit f une fonction dérivable au voisinage d'un point intérieur x_0 de D_f . Pour que f admette un extremum local en x_0 , il est nécessaire que $f'(x_0) = 0$.*

Preuve : Il faut voir que si f admet un extremum local en x_0 alors $f'(x_0) = 0$. Quitte à remplacer f par $-f$, on peut supposer que f admet un maximum local en x_0 . D'après les hypothèses, il existe $\alpha > 0$ tel que f est dérivable sur $[x_0 - \alpha, x_0 + \alpha]$. Puisque f admet un maximum local en x_0 , il existe $\beta > 0$ tel que $f(x_0) \geq f(x)$ pour tout $x \in [x_0 - \beta; x_0 + \beta]$. Toutes ces propriétés sont vérifiées sur l'intervalle $[x_0 - \gamma; x_0 + \gamma]$ où $\gamma = \min(\alpha; \beta)$.

Pour $x \in [x_0 - \gamma; x_0[$, on a $x - x_0 < 0$ et $f(x) - f(x_0) \leq 0$ puisque f admet un maximum local en x_0 . Ainsi, pour $x \in [x_0 - \gamma; x_0[$, on a :

$$\frac{f(x) - f(x_0)}{x - x_0} \geq 0.$$

En faisant tendre x vers x_0 , on en déduit que nécessairement, $f'(x_0) \geq 0$.

Pour $x \in]x_0; x_0 + \gamma]$, on a $x - x_0 > 0$ et $f(x) - f(x_0) \leq 0$ puisque f admet un maximum local en x_0 . Ainsi, $x \in]x_0; x_0 + \gamma]$, on a :

$$\frac{f(x) - f(x_0)}{x - x_0} \leq 0.$$

En faisant tendre x vers x_0 , on en déduit que nécessairement, $f'(x_0) \leq 0$.

Puisque, nécessairement, $f'(x_0) \geq 0$ et $f'(x_0) \leq 0$, on en déduit que $f'(x_0) = 0$. □

Remarque 4.16. Le fait que, pour un point intérieur x_0 de D_f , $f'(x_0) = 0$ ne suffit pas pour décider si f admet un minimum local en x_0 , un maximum local en x_0 ou aucun des deux. Dans le dernier cas, on dit que f présente un *point selle* en x_0 . Si f est deux fois dérivable en x_0 , le signe de la dérivée seconde f'' en x_0 (i.e. de la dérivée de la dérivée évaluée en x_0) peut permettre de conclure. Plus précisément, si $f'(x_0) = 0$ et $f''(x_0) > 0$ (resp. $f''(x_0) < 0$) alors f admet un minimum local en x_0 (resp. un maximum local en x_0).

La notion de dérivée permet également d'étudier les variations d'une fonction.

Proposition 4.14. *Soit f une fonction dérivable sur un intervalle I .*

Si $f'(x) \geq 0$ (resp. $f'(x) \leq 0$) pour tout $x \in I$, alors f est croissante (resp. décroissante) sur I .

Si l'inégalité est stricte sauf en un nombre fini de points, alors f est strictement croissante (resp. strictement décroissante) sur I .

Remarque 4.17. Cette proposition est intuitive si l'on se souvient que le nombre $f'(x_0)$ est le coefficient directeur de la tangente à la courbe de f en x_0 . Sa preuve n'est pourtant pas évidente et s'appuie sur des résultats avancés d'analyse (Théorème des accroissements finis, ...). Elle sera discutée dans la Section 4.7.

On résume généralement toute l'information sur une fonction dans un *tableau de variation*. Un tel tableau fait apparaître :

- les bornes du domaine de définition de la fonction (une double barre est placée sous les valeurs interdites),
- le signe de la dérivée de cette fonction,

— les variations de la fonction, au moyen de flèches, et les valeurs des extrema locaux et limites au bord du domaine de définition de la fonction.

Détaillons l'utilisation de ces techniques sur un exemple partant d'une situation concrète.

Exemple 4.17. Un industriel est seul à fabriquer un produit spécifique et est assuré de vendre toute sa production. Le coût de fabrication x unités en une semaine est de $5x^2 + 5x + 100$ € et le prix de vente d'une unité est de 85€. Il veut déterminer la quantité de produit à fabriquer pour maximiser son bénéfice hebdomadaire sur ce produit.

Le bénéfice réalisé sur la vente de $x \geq 0$ produits en une semaine est donné par le montant obtenu par les ventes auquel on soustrait le coût de production, c'est-à-dire par :

$$f(x) = 85x - (5x^2 + 5x + 100) = -5x^2 + 80x - 100, \quad x \geq 0.$$

Remarque : On se restreint aux $x \geq 0$ puisque le nombre de produits fabriqués ne peut pas être négatif.

La fonction f est dérivable sur \mathbf{R}_+ et sa dérivée est :

$$f'(x) = -10x + 8.$$

Celle-ci est strictement positive pour $x < 8$, strictement négative pour $x > 8$ et nulle en $x = 8$. La fonction f est donc croissante sur $[0; 8]$ et décroissante sur $[8; +\infty]$. Son maximum est donc atteint en 8 et vaut $f(8) = 220$. La quantité optimale à produire est donc de 8 unités et le bénéfice maximal est de 220€.

Avant de dresser le tableau de variation de cette fonction notons que :

$$f(0) = -100 \quad \text{et} \quad \lim_{x \rightarrow +\infty} f(x) = -\infty.$$

Le tableau de variation de f prend la forme suivante :

x	0	8	∞
$f'(x)$	+	0	-
f	0	220	$-\infty$

4.6 Applications en économie

4.6.1 Coût marginal

Supposons que le coût de fabrication de q unités d'un produit soit donné par $CT(q)$ où CT une fonction dérivable en q . Le coût de production d'une unité supplémentaire est approché par le *coût marginal* $Cm(q) = CT'(q)$.

Exemple 4.18. Supposons que le coût de fabrication de q unités d'un produit soit donné par $CT(q) = 10 + \sqrt{200q}$. Produire 50 unités coûte alors $CT(50) = 110$ €. Le coût (réel) de

production de la 51^e unité est $CT(51) - CT(50) \simeq 0,995$. Il est (convenablement) approché par son coût marginal :

$$Cm(50) = CT'(50) = \frac{100}{\sqrt{200 \times 50}} = 1.$$

4.6.2 Coût moyen, coût marginal et optimum technique

Si le coût total de production de $q > 0$ objets est donné par $CT(q)$, le *coût moyen* de production d'une unité est donné par :

$$CM(q) = \frac{CT(q)}{q}$$

lorsque l'on produit q unités.

On a donc que :

$$CM'(q) = \frac{CT'(q)q - CT(q)}{q^2}$$

est du signe de

$$CT'(q)q - CT(q)$$

ou encore de

$$CT'(q) - \frac{CT(q)}{q}$$

c'est-à-dire de

$$Cm(q) - CM(q).$$

Ainsi, le coût moyen est croissant aux points pour lesquels le coût marginal est supérieur au coût moyen et décroissant aux points pour lesquels le coût marginal est inférieur au coût moyen.

En économie, l'*optimum technique* q^* est défini comme la quantité à produire afin de minimiser le coût moyen de production. Lorsqu'il existe (ce qui est le cas dans la plupart des cas pratiques), il vérifie nécessairement :

$$CM'(q^*) = \frac{CT'(q^*)q^* - CT(q^*)}{(q^*)^2} = 0$$

c'est-à-dire

$$Cm(q^*) = CM(q^*).$$

Graphiquement, les courbes représentatives des coûts moyen et marginal se coupe en l'optimum technique.

4.6.3 Revenu marginal

Supposons que le revenu total engendré par la fabrication de x unités d'un produit soit donné par $f(x)$ où f une fonction dérivable en x . De la même façon que dans la sous-section précédente, le revenu total engendré par la fabrication de $x + 1$ unités est (convenablement) approché par le *revenu marginal* $Rm = f'(x)$.

4.6.4 Élasticité

Définition 4.33. Soit f une fonction dérivable en x et telle que $f(x) \neq 0$.

On appelle élasticité (instantanée) de f en x le nombre :

$$\mathcal{E}_f(x) = \frac{xf'(x)}{f(x)}.$$

Remarque 4.18. En utilisant la définition de la dérivée, on peut réécrire l'élasticité de f en x sous la forme :

$$\mathcal{E}_f(x) = \lim_{h \rightarrow 0} \frac{\frac{f(x+h)-f(x)}{f(x)}}{\frac{h}{x}}.$$

Intuitivement, dans la limite précédente, le numérateur $\frac{f(x+h)-f(x)}{f(x)}$ mesure de combien de pourcent $f(x+h)$ est plus grand que $f(x)$ alors que le dénominateur $\frac{h}{x}$ mesure de combien de pourcent $x+h$ est plus grand que x . En pratique, on utilise que si x augmente de $\varepsilon\%$, avec ε petit, $f(x)$ augmente approximativement de $(\mathcal{E}_f(x) \times \varepsilon)\%$.

Exemple 4.19. Supposons que la demande d'un produit sur le marché soit donnée par $f(x) = 100 - 4x$ où x est le prix de vente. On a $f'(x) = -4$, pour tout x et donc $\mathcal{E}_f(x) = -\frac{4x}{100-4x}$. Ainsi, si le prix de vente, initialement x , augmente de 0,5%, la demande baissera de $\frac{2x}{100-4x}\%$. En particulier, si le prix de vente initial est de $x = 20\text{€}$ et augmente de 0,5%, la demande baissera de 2%.

Typologie des biens à partir de l'élasticité revenu

Considérons un bien dont la fonction demande est donnée par $D(Y, p)$ où Y désigne le revenu (Yield en anglais) du consommateur et p le prix de vente du bien. Dans ce paragraphe, le prix p est fixé et le revenu Y est la variable. L'élasticité revenu est l'élasticité de D vue comme une fonction de Y :

$$\mathcal{E}_D^{\text{revenu}}(Y) = \frac{YD'(Y)}{D(Y)}.$$

En économie, on définit alors les types de biens suivants grâce à l'élasticité revenu pour un consommateur ayant un revenu Y_0 fixé :

- Un bien *inférieur* est un bien dont la consommation diminue lorsque le revenu augmente c'est-à-dire tel que :

$$\mathcal{E}_D^{\text{revenu}}(Y_0) < 0.$$

- Un bien *normal* est un bien dont la consommation augmente lorsque le revenu augmente. c'est-à-dire tel que :

$$\mathcal{E}_D^{\text{revenu}}(Y_0) > 0.$$

On considère qu'il est normal que la demande augmente lorsque le revenu augmente. Parmi les biens normaux, on peut distinguer les biens de *lux* et les biens de *première nécessité*.

- Un bien de *première nécessité* est un bien dont la consommation augmente, en pourcentage, moins que le revenu :

$$0 < \mathcal{E}_D^{\text{revenu}}(Y_0) < 1.$$

Si le revenu de l'individu augmente de 1%, sa consommation va augmenter de moins de 1%.

- Un bien de *luxe* est un bien dont la consommation augmente, en pourcentage, plus que le revenu :

$$\mathcal{E}_D^{\text{revenu}}(Y_0) > 1.$$

Si le revenu de l'individu augmente de 1%, sa consommation va augmenter de plus de 1%.

Typologie des biens à partir de l'élasticité prix directe

Considérons un bien dont la fonction demande est donnée par $D(Y, p)$ où Y désigne le revenu (Yield en anglais) du consommateur et p le prix de vente du bien. Dans ce paragraphe, le revenu Y du consommateur est fixé et le prix p du bien est la variable. L'élasticité prix directe est l'élasticité de D vue comme une fonction de p :

$$\mathcal{E}_D^{\text{prix}}(p) = \frac{pD'(p)}{D(p)}.$$

En économie, on définit alors les types de biens suivants grâce à l'élasticité prix directe pour un prix p_0 fixé :

- Un bien *ordinaire* (ou *normal*) est un bien dont la consommation diminue lorsque son prix augmente c'est-à-dire tel que :

$$\mathcal{E}_D^{\text{prix}}(p_0) < 0.$$

On considère qu'il est normal que la demande diminue lorsque le prix augmente.

- Pour compléter la classification des bien « anormaux » (du point de vue de l'élasticité prix) pour lesquels la demande augmente lorsque le prix augmente :

$$\mathcal{E}_D^{\text{prix}}(p_0) > 0,$$

il faut s'intéresser, dans un second temps, à leur élasticité revenu. Ainsi, on dit qu'un bien est :

- de *Veblen* si :

$$\mathcal{E}_D^{\text{prix}}(p_0) > 0$$

et

$$\mathcal{E}_D^{\text{revenu}}(Y_0) > 0;$$

Note : Ces biens portent le nom de l'effet découvert par l'économiste et sociologue américain Thorstein Veblen (1857 – 1929). C'est le cas des biens de luxe.

- de *Giffen* si :

$$\mathcal{E}_D^{\text{prix}}(p_0) > 0$$

et

$$\mathcal{E}_D^{\text{revenu}}(Y_0) < 0 \quad (\text{bien inférieur}).$$

Note : Ces biens sont très rares et portent le nom de l'économiste écossais Robert Giffen (1837 – 1910) qui en a découvert l'existence en étudiant la consommation de pommes de terre en Irlande.

4.7 Complément : autour du lien entre variations d'une fonction et signe de sa dérivée

Dans cette section, nous développons les outils d'analyse permettant de comprendre précisément le lien entre variations d'une fonction et signe de sa dérivée. Cette section est technique et peut être largement omise.

Les autres cas étant similaires, nous nous restreindrons à la preuve de la proposition suivante.

Proposition 4.15. *Si f est dérivable sur un intervalle I et $f'(x) \geq 0$ pour tout $x \in I$, alors f est croissante sur I .*

Nous débuterons l'argument par l'établissement du Théorème de Rolle.

Théorème 4.4 (Théorème de Rolle). *Soit f une fonction continue sur $[a; b]$ et dérivable sur $]a; b[$.*

Si $f(a) = f(b)$, alors il existe un réel $c \in]a; b[$ tel que :

$$f'(c) = 0.$$

Preuve : L'intervalle $[a; b]$ est un intervalle fermé et borné (on dit que c'est un *compact*). Or si une fonction est continue sur un intervalle fermé borné alors elle est bornée sur cet intervalle et atteint ses bornes dans cet intervalle (ce que l'on admettra).

Notons m le minimum de f sur $[a; b]$, M son maximum sur $[a; b]$. Soit $c, d \in [a; b]$ tel que $f(c) = m$ et $f(d) = M$.

Distinguons deux cas.

Premier cas : $M = m$. Puisque $m = M$, f est constante sur $[a; b]$ et sa dérivée est constante égale à 0. Il n'y a rien à faire et tout choix de $c \in]a; b[$ convient.

Second cas : $M \neq m$. Puisque $f(a) = f(b)$, si M (resp. m) est atteint en a , alors m (resp. M) n'est atteint ni en a , ni en b . Ainsi, au moins un des deux extrema ne peut être atteint qu'à l'intérieur de l'intervalle $]a; b[$. Soit $c \in]a; b[$, un point où f atteint cet extremum.

Puisque f est dérivable sur l'intervalle $]a; b[$ et que f admet un extremum en $c \in]a; b[$, la Proposition 4.13 assure que $f'(c) = 0$. Le Théorème de Rolle est donc établi. \square

Le théorème suivant est essentiellement une conséquence du Théorème de Rolle et sera l'ingrédient principal de la preuve de la Proposition 4.15.

Théorème 4.5 (Théorème des accroissements finis). *Soit f une fonction continue l'intervalle $[a; b]$ et dérivable sur $]a; b[$.*

Il existe un réel $c \in]a; b[$ tel que :

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Preuve : Définissons sur l'intervalle $[a; b]$ la fonction g par :

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a}x.$$

Les fonctions f et $x \mapsto \frac{f(b)-f(a)}{b-a}x$ étant continues sur $[a; b]$ et dérivables sur $]a; b[$, la fonction g est continue sur $[a; b]$ et dérivable sur $]a; b[$ comme différence de deux telles fonctions. De plus on a :

$$\begin{aligned} g(a) &= f(a) - \frac{f(b) - f(a)}{b - a}a \\ &= \frac{f(a)(b - a) - a(f(b) - f(a))}{b - a} \\ &= \frac{bf(a) - af(a) - af(b) + af(a)}{b - a} \\ &= \frac{bf(a) - af(b)}{b - a} \\ &= \frac{bf(b) - af(b) - bf(b) + bf(a)}{b - a} \\ &= \frac{f(b)(b - a) - b(f(b) - f(a))}{b - a} \\ &= f(b) - \frac{f(b) - f(a)}{b - a}b \\ &= g(b) \end{aligned}$$

Ainsi, par le Théorème de Rolle (Théorème 4.4) appliqué à la fonction g , on obtient qu'il existe $c \in]a; b[$ tel que $g'(c) = 0$.

Or,

$$g'(x) = f'(x) - \frac{f(b) - f(a)}{b - a}.$$

Il s'ensuit que :

$$f'(c) - \frac{f(b) - f(a)}{b - a} = g'(c) = 0$$

et donc

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

□

Preuve de la Proposition 4.15 : Soient $a, b \in I$ tels que $a < b$. Nous devons montrer que $f(a) \leq f(b)$.

Puisque l'on a supposé la fonction f dérivable sur l'intervalle I , elle l'est aussi sur le sous-intervalle $[a; b]$.

En appliquant le Théorème des accroissements finis (Théorème 4.5), on obtient l'existence d'un réel $c \in]a; b[$ tel que :

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Puisque $c \in]a; b[\subset I$, on a $f'(c) \geq 0$. Ainsi,

$$\frac{f(b) - f(a)}{b - a} \geq 0.$$

Or, on a choisi $a < b$ donc $b - a > 0$. Il s'ensuit que $f(b) - f(a) \geq 0$, c'est-à-dire $f(b) \geq f(a)$.

Nous venons de voir que pour tout $a, b \in I$, si $a < b$ alors $f(b) \geq f(a)$. La fonction f est donc croissante sur I . \square

Chapitre 5

Suites numériques

5.1 Premières définitions

Définition 5.1. On appelle suite de nombres réels toute famille infinie d'éléments de \mathbf{R} , appelés termes, indexée par \mathbf{N} . On note classiquement une suite $(u_n)_{n \in \mathbf{N}}$.

Définition 5.2. On dit que la suite $(u_n)_{n \in \mathbf{N}}$ est croissante (resp. strictement croissante, décroissante, strictement décroissante) si, pour tout $n \in \mathbf{N}$, $u_{n+1} \geq u_n$ (resp. $u_{n+1} > u_n$, $u_{n+1} \leq u_n$, $u_{n+1} < u_n$).

Il est parfois pratique d'étudier la différence ou le quotient (quand ceci est possible) de deux termes consécutifs d'une suite pour obtenir son sens de variation. La proposition suivante s'obtient par des calculs immédiats.

Proposition 5.1. Soit $(u_n)_{n \in \mathbf{N}}$ une suite.

1. La suite $(u_n)_{n \in \mathbf{N}}$ est croissante (resp. strictement croissante, décroissante, strictement décroissante) si, et seulement si, pour tout $n \in \mathbf{N}$, $u_{n+1} - u_n \geq 0$ (resp. $u_{n+1} - u_n > 0$, $u_{n+1} - u_n \leq 0$, $u_{n+1} - u_n < 0$).
2. Supposons que tous les termes de la suite $(u_n)_{n \in \mathbf{N}}$ soient strictement positifs. Alors, la suite $(u_n)_{n \in \mathbf{N}}$ est croissante (resp. strictement croissante, décroissante, strictement décroissante) si, et seulement si, pour tout $n \in \mathbf{N}$, $\frac{u_{n+1}}{u_n} \geq 1$ (resp. $\frac{u_{n+1}}{u_n} > 1$, $\frac{u_{n+1}}{u_n} \leq 1$, $\frac{u_{n+1}}{u_n} < 1$).

Exemple 5.1. Soit $(u_n)_{n \in \mathbf{N}}$ la suite définie par :

$$u_n = \frac{n!}{n^n}, \quad \text{pour tout } n \in \mathbf{N},$$

avec $n! = 1 \times 2 \times \cdots \times n$ et par convention $0! = 1$.

Pour tout $n \in \mathbf{N}$, on a :

$$\begin{aligned} \frac{u_{n+1}}{u_n} &= \frac{\frac{(n+1)!}{(n+1)^{n+1}}}{\frac{n!}{n^n}} \\ &= \frac{(n+1)!}{(n+1)^{n+1}} \times \frac{n^n}{n!} \\ &= \frac{n^n}{(n+1)^n} = \left(1 - \frac{1}{n+1}\right)^n \\ &\leq 1. \end{aligned}$$

La suite $(u_n)_{n \in \mathbf{N}}$ est donc décroissante.

Définition 5.3. On dit que la suite $(u_n)_{n \in \mathbf{N}}$ est majorée (resp. minorée) s'il existe $M \in \mathbf{R}$ (resp. $m \in \mathbf{R}$) tel que pour tout $n \in \mathbf{N}$, $u_n \leq M$ (resp. $u_n \geq m$). On dit alors que M (resp. m) est un majorant (resp. minorant) de $(u_n)_{n \in \mathbf{N}}$.

Exemple 5.2. La suite définie par $u_n = \frac{1}{n^2+1}$ est majorée par $\frac{1}{2}$ et minorée par 0.

5.2 Convergence et limites de suites

Définition 5.4. On dit que la suite $(u_n)_{n \in \mathbf{N}}$ converge vers $l \in \mathbf{R}$ si, pour tout $\varepsilon > 0$, il existe $N \in \mathbf{N}$ tel que pour tout $n \in \mathbf{N}$:

$$n \geq N \implies |u_n - l| \leq \varepsilon.$$

On note alors :

$$\lim_{n \rightarrow +\infty} u_n = l.$$

Définition 5.5. On dit que la suite $(u_n)_{n \in \mathbf{N}}$ diverge vers $+\infty$ (resp. $-\infty$) si, pour tout $A > 0$, il existe $N \in \mathbf{N}$ tel que pour tout $n \in \mathbf{N}$:

$$n \geq N \implies u_n \geq A \quad (\text{resp. } u_n \leq -A).$$

On note alors :

$$\lim_{n \rightarrow +\infty} u_n = +\infty \quad (\text{resp. } -\infty).$$

Proposition 5.2. 1. Si $(u_n)_{n \in \mathbf{N}}$ converge vers l_1 et $(v_n)_{n \in \mathbf{N}}$ converge vers l_2 , alors $(u_n + v_n)_{n \in \mathbf{N}}$ converge vers $l_1 + l_2$, $(u_n \times v_n)_{n \in \mathbf{N}}$ converge vers $l_1 l_2$, et si $l_2 \neq 0$, $\left(\frac{u_n}{v_n}\right)_{n \in \mathbf{N}}$ converge vers $\frac{l_1}{l_2}$.

2. S'il existe N tel que pour tout $n \geq N$, $u_n \geq v_n$, alors

$$\lim_{n \rightarrow +\infty} u_n \geq \lim_{n \rightarrow +\infty} v_n.$$

En particulier, s'il existe N tel que pour tout $n \geq N$, $u_n \geq v_n$ et $(v_n)_{n \in \mathbf{N}}$ diverge vers $+\infty$ (resp. $(u_n)_{n \in \mathbf{N}}$ diverge vers $-\infty$), alors $(u_n)_{n \in \mathbf{N}}$ diverge vers $+\infty$ (resp. $(v_n)_{n \in \mathbf{N}}$ diverge vers $-\infty$).

3. Si $(u_n)_{n \in \mathbf{N}}$ et $(v_n)_{n \in \mathbf{N}}$ convergent vers un même réel l , et s'il existe N tel que pour tout $n \geq N$, $u_n \leq w_n \leq v_n$, alors $(w_n)_{n \in \mathbf{N}}$ converge aussi vers l .

Il est parfois utile de s'assurer de la convergence d'une suite sans pour autant déterminer sa limite. La proposition suivante permet de faire ceci.

Proposition 5.3. *Si $(u_n)_{n \in \mathbf{N}}$ est croissante et majorée (resp. décroissante et minorée), alors elle est convergente.*

Preuve : Quitte à remplacer la suite $(u_n)_{n \in \mathbf{N}}$ par la suite $(-u_n)_{n \in \mathbf{N}}$, il suffit de montrer que si $(u_n)_{n \in \mathbf{N}}$ est croissante et majorée, alors elle est convergente.

Soit l le plus petit majorant de la suite $(u_n)_{n \in \mathbf{N}}$. Alors, pour tout $\varepsilon > 0$, $l - \varepsilon$ n'est pas un majorant. Donc il existe N tel que $l - \varepsilon \leq u_{n_0} \leq l$. Or, $(u_n)_{n \in \mathbf{N}}$ est croissante, donc pour tout $n \geq N$, $l - \varepsilon \leq u_N \leq u_n \leq l \leq l + \varepsilon$. Ainsi, $(u_n)_{n \in \mathbf{N}}$ converge vers l . \square

Remarque 5.1. Si $(u_n)_{n \in \mathbf{N}}$ est croissante et n'est pas majorée (resp. décroissante et n'est pas minorée), alors elle diverge vers $+\infty$ (resp. $-\infty$).

Définition 5.6. *On dit que deux suites $(u_n)_{n \in \mathbf{N}}$ et $(v_n)_{n \in \mathbf{N}}$ sont adjacentes si l'une est croissante, l'autre décroissante et*

$$\lim_{n \rightarrow +\infty} u_n - v_n = 0.$$

Théorème 5.1. *Si $(u_n)_{n \in \mathbf{N}}$ et $(v_n)_{n \in \mathbf{N}}$ sont deux suites adjacentes, alors elles sont convergentes et*

$$\lim_{n \rightarrow +\infty} u_n = \lim_{n \rightarrow +\infty} v_n.$$

Preuve : Sans perte de généralité, on peut supposer que la suite $(u_n)_{n \in \mathbf{N}}$ est croissante et $(v_n)_{n \in \mathbf{N}}$ décroissante. Remarquons que $(u_n)_{n \in \mathbf{N}}$ est majorée par v_0 et $(v_n)_{n \in \mathbf{N}}$ est minorée par u_0 . Par la Proposition 5.3, ces deux suites sont convergentes.

On a de plus :

$$0 = \lim_{n \rightarrow +\infty} (u_n - v_n) = \lim_{n \rightarrow +\infty} u_n - \lim_{n \rightarrow +\infty} v_n.$$

D'où

$$\lim_{n \rightarrow +\infty} u_n = \lim_{n \rightarrow +\infty} v_n.$$

\square

Exercice 5.1. Montrer que les suite définies par :

$$u_n = \frac{n-1}{n} \quad \text{et} \quad v_n = 1 + \frac{1}{n^2}$$

sont adjacentes et déterminer leur limite.

5.3 Suites particulières

5.3.1 Suites définies par récurrence

Définition 5.7. On dit qu'une suite $(u_n)_{n \in \mathbf{N}}$ est récurrente d'ordre p si ses p premiers termes u_0, \dots, u_{p-1} sont donnés et pour tout $n \in \mathbf{N}$, u_{n+p} s'exprime en fonction de $u_n, u_{n+1}, \dots, u_{n+p-1}$. En d'autres termes, $(u_n)_{n \in \mathbf{N}}$ est récurrente d'ordre p s'il existe une fonction f de p variables telle que :

$$\begin{cases} u_{n+p} = f(u_n, u_{n+1}, \dots, u_{n+p-1}) \\ u_0, \dots, u_{p-1} \text{ donnés} \end{cases} .$$

Remarque 5.2. Si $(u_n)_{n \in \mathbf{N}}$ est définie par :

$$\begin{cases} u_{n+p} = \alpha_1 u_n + \alpha_2 u_{n+1} + \dots + \alpha_p u_{n+p-1} \\ u_0, \dots, u_{p-1} \text{ donnés} \end{cases} ,$$

on parle de suite récurrente *linéaire* d'ordre p .

Si $(u_n)_{n \in \mathbf{N}}$ est définie par :

$$\begin{cases} u_{n+p} = \alpha_0 + \alpha_1 u_n + \alpha_2 u_{n+1} + \dots + \alpha_p u_{n+p-1} \\ u_0, \dots, u_{p-1} \text{ donnés} \end{cases} ,$$

on parle de suite récurrente *affine* d'ordre p .

Exemple 5.3. La suite de Fibonacci est la suite récurrente linéaire d'ordre 2 définie par :

$$\begin{cases} u_{n+2} = u_n + u_{n+1} \\ u_0 = u_1 = 1 \end{cases} .$$

5.3.2 Suites arithmétiques

Définition 5.8. On appelle suite arithmétique de raison r et de premier terme u_0 la suite (récurrente affine d'ordre 1) définie par :

$$\begin{cases} u_{n+1} = u_n + r \\ u_0 \text{ donné} \end{cases} .$$

Remarque 5.3. Pour montrer qu'une suite est arithmétique, il suffit de vérifier que la différence de deux termes consécutifs $u_{n+1} - u_n$ est constante. Cette constante est alors sa raison r . Il s'ensuit qu'une suite arithmétique est strictement croissante si $r > 0$, strictement décroissante si $r < 0$ et constante si $r = 0$.

La proposition suivante donne l'expression du terme général d'une suite arithmétique.

Proposition 5.4. Soit $(u_n)_{n \in \mathbf{N}}$ la suite arithmétique de raison r et de premier terme u_0 . Pour tout $n \in \mathbf{N}$, on a :

$$u_n = u_0 + nr. \tag{5.3.1}$$

Remarque 5.4. En fait, une suite $(u_n)_{n \in \mathbf{N}}$ est arithmétique de raison r et de premier terme u_0 si, et seulement si, son terme général est donné par (5.3.1).

Preuve : Nous allons démontrer ce résultat par un *raisonnement par récurrence*. Le principe du raisonnement par récurrence est le suivant. Pour montrer qu'une propriété est vraie pour tout $n \in \mathbf{N}$ (ou pour tout n plus grand qu'un certain N), on peut :

- montrer qu'elle est vraie au rang initial $n = 0$ (ou $n = N$; on parle d'*initialisation*),
- montrer qu'elle est héréditaire à partir de ce rang, c'est-à-dire que si l'on suppose la propriété vraie au rang n (ou jusqu'à un rang n), on est alors capable de montrer qu'elle l'est aussi au rang $n + 1$,
- conclure par le principe du raisonnement par récurrence.

Définissons l'hypothèse de récurrence :

$$(H_n) \quad \begin{array}{l} \text{le terme d'ordre } n \text{ de la suite arithmétique de raison } r \\ \text{et de premier terme } u_0 \text{ est donné par } u_n = u_0 + nr. \end{array}$$

Initialisation : au rang $n = 0$, c'est évident ! En effet, $u_0 = u_0 + 0 \times r$.

Hérédité : Supposons que, pour un certain entier naturel n , (H_n) soit vraie et montrons que (H_{n+1}) est alors vraie.

On a :

$$\begin{aligned} u_{n+1} &= u_n + r && \text{(par définition)} \\ &= u_0 + nr + r && \text{(d'après l'hypothèse de récurrence } (H_n)) \\ &= u_0 + (n + 1)r. \end{aligned}$$

Ainsi, la propriété (H_n) est héréditaire à partir du rang $n = 0$.

Conclusion : La propriété (H_n) étant vraie au rang $n = 0$ et héréditaire à partir de ce rang, le principe du raisonnement par récurrence nous permet de conclure que le terme général de la suite arithmétique de raison r et de premier terme u_0 est donné par :

$$u_n = u_0 + nr, \quad \text{pour tout } n \in \mathbf{N}.$$

□

Corollaire 5.1. Soit $(u_n)_{n \in \mathbf{N}}$ la suite arithmétique de raison r et de premier terme u_0 .

Pour tous $n, k \in \mathbf{N}$, on a :

$$u_{n+k} = u_n + kr.$$

Preuve : D'après la Proposition 5.4, on a :

$$u_{n+k} = u_0 + (n + k)r = u_0 + nr + kr = u_n + kr.$$

□

Corollaire 5.2. Soit $(u_n)_{n \in \mathbf{N}}$ la suite arithmétique de raison $r \neq 0$ et de premier terme u_0 .

Alors, $(u_n)_{n \in \mathbf{N}}$ diverge vers $+\infty$ si $r > 0$ et vers $-\infty$ si $r < 0$.

Preuve : D'après la Proposition 5.4, on a :

$$\lim_{n \rightarrow +\infty} u_n = \lim_{n \rightarrow +\infty} (u_0 + nr) = u_0 + r \lim_{n \rightarrow +\infty} n = \begin{cases} +\infty & \text{si } r > 0 \\ -\infty & \text{si } r < 0 \end{cases} .$$

□

Méthode pour déterminer la raison et le terme initial d'une suite arithmétique

Supposons que $(u_n)_{n \in \mathbf{N}}$ soit une suite arithmétique dont on ne connaît ni le premier terme ni la raison mais que l'on connaisse deux termes de cette suite $u_k = a$ et $u_n = b$. Pour déterminer sa raison r et son premier terme u_0 , il suffit d'utiliser la Proposition 5.4 pour obtenir que r et u_0 satisfont le système :

$$\begin{cases} u_0 + kr = u_k = a \\ u_0 + nr = u_n = b \end{cases} .$$

On résout ensuite ce système en u_0 et r .

Remarque 5.5.

1. Si u_0 ou r est connu, il n'y a qu'une équation à considérer.
2. Alternativement, on peut déterminer la raison r de la suite en calculant de deux façons différentes le reste de la différence $u_n - u_k$ (qui vaut $(n - k)r$ puis déduire la valeur de u_0 en utilisant que $u_n = u_0 + nr$ donc $u_0 = u_n - nr$. De cette méthode, on peut extraire une méthode pour déterminer la raison d'une suite arithmétique à partir de deux de ses termes et une méthode pour déterminer le terme initial d'une suite arithmétique à partir de sa raison et un de ses termes.

Exemple 5.4. Déterminons la raison r et le premier terme u_0 de la suite arithmétique $(u_n)_{n \in \mathbf{N}}$ telle que $u_3 = 9$ et $u_7 = 17$. Puisque $(u_n)_{n \in \mathbf{N}}$ est arithmétique de raison r et de premier terme u_0 , la Proposition 5.4 assure que $9 = u_3 = u_0 + 3r$ et $17 = u_7 = u_0 + 7r$. Ainsi, r et u_0 satisfont le système :

$$\begin{cases} u_0 + 3r = 9 \\ u_0 + 7r = 17 \end{cases} .$$

Or, on a :

$$\begin{aligned} \begin{cases} u_0 + 3r = 9 \\ u_0 + 7r = 17 \end{cases} &\iff \begin{cases} u_0 + 3r = 9 \\ 4r = 8 \end{cases} \\ &\iff \begin{cases} u_0 = 9 - 3r = 9 - 3 \times 2 = 3 \\ r = 2 \end{cases} . \end{aligned}$$

Ainsi, $(u_n)_{n \in \mathbf{N}}$ est arithmétique de raison $r = 2$ et de premier terme $u_0 = 3$.

Sommes partielles

Définition 5.9. Soit $(u_n)_{n \in \mathbf{N}}$ une suite.

On appelle somme partielle d'ordre n de la suite $(u_n)_{n \in \mathbf{N}}$ la somme :

$$S_n = \sum_{k=0}^{n-1} u_k = u_0 + u_1 + \dots + u_{n-1} + u_{n-1}.$$

Proposition 5.5. Soit $(u_n)_{n \in \mathbf{N}}$ la suite arithmétique de raison r et de premier terme u_0 .

Alors, pour tout $n \in \mathbf{N}^*$, on a :

$$S_n = \sum_{k=0}^{n-1} u_k = u_0 + \dots + u_{n-1} = \frac{n(u_0 + u_{n-1})}{2}.$$

Remarque 5.6. Pour éviter toute erreur, il est conseillé de se souvenir de la formule précédente comme suit :

$$S_n = \frac{\text{« nombre de termes »} \times \text{« premier terme + dernier terme »}}{2}.$$

On peut alors écrire des formules analogues pour d'autres sommes partielles de termes consécutifs d'une suite arithmétique (par nécessairement les n premiers termes), comme dans la proposition suivante.

Proposition 5.6. Soit $(u_n)_{n \in \mathbf{N}}$ la suite arithmétique de raison r et de premier terme u_0 .

Alors, pour tout $n \in \mathbf{N}^*$, on a :

$$\begin{aligned} S_n &= \sum_{k=0}^{n-1} u_k = u_0 + \dots + u_{n-1} = \frac{n(u_0 + u_{n-1})}{2}, \\ \tilde{S}_n &= \sum_{k=1}^n u_k = u_1 + \dots + u_n = \frac{n(u_1 + u_n)}{2}, \\ \bar{S}_n &= \sum_{k=0}^n u_k = u_0 + \dots + u_n = \frac{(n+1)(u_0 + u_n)}{2}, \\ \check{S}_n &= \sum_{k=1}^{n-1} u_k = u_1 + \dots + u_{n-1} = \frac{(n-1)(u_1 + u_{n-1})}{2}. \end{aligned}$$

Preuve : Nous allons donner deux preuves de ce résultat.

Preuve 1 [« à la Gauss »] : Commençons par écrire que, puisque $(u_n)_{n \in \mathbf{N}}$ est la suite arithmétique de raison r et de premier terme u_0 ,

$$\begin{aligned} S_n &= u_0 + u_1 + \dots + u_{n-2} + u_{n-1} \\ &= u_0 + u_0 + r + \dots + u_0 + (n-2)r + u_0 + (n-1)r. \end{aligned}$$

L'astuce est alors d'écrire l'identité précédente de gauche à droite et de droite à gauche et de sommer terme à terme :

$$\begin{array}{r} S_n = u_0 + u_0 + r + \dots + u_0 + (n-2)r + u_0 + (n-1)r \\ + S_n = u_0 + (n-1)r + u_0 + (n-2)r + \dots + u_0 + r + u_0 \\ \hline 2S_n = 2u_0 + (n-1)r + 2u_0 + (n-1)r + \dots + 2u_0 + (n-1)r + 2u_0 + (n-1)r \end{array}.$$

En observant qu'il y a $(n + 1)$ termes dans le membre de droite de la dernière identité, on obtient que :

$$2S_n = n(2u_0 + (n - 1)r) = n(u_0 + u_{n-1}),$$

d'où

$$S_n = \frac{n(u_0 + u_{n-1})}{2}.$$

Preuve 2 [par récurrence] : Montrons le résultat par récurrence sur $n \in \mathbf{N}^*$.

L'hypothèse de récurrence est, pour $n \in \mathbf{N}^*$:

$$(H_n) \quad S_n = \frac{n(u_0 + u_{n-1})}{2}$$

Initialisation : au rang $n = 1$, c'est évident ! En effet, $S_1 = u_1 = \frac{1 \times (u_1 + u_1)}{2}$.

Hérédité : Supposons que, pour un certain entier naturel n , (H_n) soit vraie et montrons que (H_{n+1}) est alors vraie.

On a :

$$\begin{aligned} S_{n+1} &= S_n + u_n && \text{(par définition de } S_{n+1} \text{ et } S_n) \\ &= \frac{n(u_0 + u_{n-1})}{2} + u_n && \text{(d'après l'hypothèse de récurrence } (H_n)) \\ &= \frac{n(u_0 + u_0 + (n - 1)r)}{2} + u_0 + nr && \text{(en utilisant la Proposition 5.4)} \\ &= \frac{2nu_0 + n(n - 1)r + 2u_0 + 2nr}{2} \\ &= \frac{2(n + 1)u_0 + (n + 1)nr}{2} \\ &= \frac{(n + 1)(2u_0 + nr)}{2} \\ &= \frac{(n + 1)(u_0 + u_n)}{2} && \text{(en utilisant la Proposition 5.4)} \end{aligned}$$

Ainsi, la propriété (H_n) est héréditaire à partir du rang $n = 0$.

Conclusion : La propriété (H_n) étant vraie au rang $n = 1$ et héréditaire à partir de ce rang, le principe du raisonnement par récurrence nous permet de conclure que :

$$S_n = \frac{n(u_0 + u_{n-1})}{2}, \quad \text{pour tout } n \in \mathbf{N}^*.$$

□

Exemple 5.5. Soit $(u_n)_{n \in \mathbf{N}}$ la suite arithmétique de raison $r = 2$ et de premier terme $u_0 = 1$. Alors,

$$\bar{S}_n = u_0 + u_1 + \cdots + u_n = \frac{(n + 1)(1 + 1 + 2n)}{2} = (n + 1)^2.$$

Exercice 5.2. Déterminer, pour tout $n \in \mathbf{N}$, la somme des entiers compris entre 1 et n .

5.3.3 Suites géométriques

Définition 5.10. On appelle suite géométrique de raison q et de premier terme u_0 la suite (récurrente linéaire d'ordre 1) définie par :

$$\begin{cases} u_{n+1} = qu_n \\ u_0 \text{ donné} \end{cases} .$$

Remarque 5.7.

1. Pour montrer qu'une suite est géométrique, il suffit de vérifier que le quotient de deux termes consécutifs $\frac{u_{n+1}}{u_n}$ est constant. Cette constante est alors sa raison q .
2. Les suites géométriques joueront un rôle important en *mathématiques financières* (voir Chapitre 6).

Proposition 5.7. Soit $(u_n)_{n \in \mathbf{N}}$ la suite géométrique de raison q et de premier terme u_0 .

1. Si $0 < q < 1$ et $u_0 > 0$ (resp. $u_0 < 0$), la suite $(u_n)_{n \in \mathbf{N}}$ est strictement décroissante (resp. strictement croissante).
2. Si $q > 1$ et $u_0 > 0$ (resp. $u_0 < 0$), la suite $(u_n)_{n \in \mathbf{N}}$ est strictement croissante (resp. strictement décroissante).
3. Si $q = 1$ ou $u_0 = 0$, la suite $(u_n)_{n \in \mathbf{N}}$ est constante.
4. Si $q = 0$ et $u_0 > 0$ (resp. $u_0 < 0$), la suite $(u_n)_{n \in \mathbf{N}}$ est (resp. croissante); elle est en fait constante égale à 0 à partir du rang 1.

Preuve : Exercice! □

Remarque 5.8. Si $q < 0$, chaque terme est du signe opposé à celui de son prédécesseur ; la suite est alors oscillante.

La proposition suivante donne l'expression du terme général d'une suite géométrique.

Proposition 5.8. Soit $(u_n)_{n \in \mathbf{N}}$ la suite géométrique de raison q et de premier terme u_0 .

Pour tout $n \in \mathbf{N}$, on a :

$$u_n = q^n u_0. \tag{5.3.2}$$

Remarque 5.9. En fait, une suite $(u_n)_{n \in \mathbf{N}}$ est géométrique de raison q et de premier terme u_0 si, et seulement si, son terme général est donné par (5.3.2).

Preuve : Nous allons montrer ce résultat par récurrence sur $n \in \mathbf{N}$. Définissons l'hypothèse de récurrence :

$$(H_n) \quad \begin{array}{l} \text{le terme d'ordre } n \text{ de la suite géométrique de raison } q \\ \text{et de premier terme } u_0 \text{ est donné par } u_n = q^n u_0. \end{array}$$

Initialisation : au rang $n = 0$, c'est évident ! En effet, $u_0 = q^0 u_0$.

Hérédité : Supposons que, pour un certain entier naturel n , (H_n) soit vraie et montrons que (H_{n+1}) est alors vraie.

On a :

$$\begin{aligned} u_{n+1} &= qu_n && \text{(par définition)} \\ &= qq^n u_0 && \text{(d'après l'hypothèse de récurrence (H}_n\text{))} \\ &= q^{n+1} u_0. \end{aligned}$$

Ainsi, la propriété (H_n) est héréditaire à partir du rang $n = 0$.

Conclusion : La propriété (H_n) étant vraie au rang $n = 0$ et héréditaire à partir de ce rang, le principe du raisonnement par récurrence nous permet de conclure que le terme général de la suite géométrique de raison q et de premier terme u_0 est donné par :

$$u_n = q^n u_0, \quad \text{pour tout } n \in \mathbf{N}.$$

□

Corollaire 5.3. Soit $(u_n)_{n \in \mathbf{N}}$ la suite géométrique de raison q et de premier terme u_0 .

Pour tous $n, k \in \mathbf{N}$, on a :

$$u_{n+k} = q^k u_n.$$

Preuve : D'après la Proposition 5.8, on a :

$$u_{n+k} = q^{n+k} u_0 = q^k q^n u_0 = q^k u_n.$$

□

Corollaire 5.4. Soit $(u_n)_{n \in \mathbf{N}}$ la suite géométrique de raison q et de premier terme u_0 .

1. Si $q \leq -1$, la suite $(u_n)_{n \in \mathbf{N}}$ n'admet pas de limite.
2. Si $-1 < q < 1$, la suite $(u_n)_{n \in \mathbf{N}}$ converge vers 0.
3. Si $q = 1$, la suite $(u_n)_{n \in \mathbf{N}}$ converge vers u_0 .
4. Si $q > 1$, la suite diverge vers $\text{sgn}(u_0)\infty$.

Preuve : Exercice !

□

Méthode pour déterminer la raison et le terme initial d'une suite géométrique

Supposons que $(u_n)_{n \in \mathbf{N}}$ soit une suite géométrique dont on ne connaît ni le premier terme ni la raison mais que l'on connaisse deux termes de cette suite $u_k = a$ et $u_n = b$. Pour déterminer sa raison q et son premier terme u_0 , il suffit d'utiliser la Proposition 5.8 pour obtenir que q et u_0 satisfont le système :

$$\begin{cases} q^k u_0 = u_k = a \\ q^n u_0 = u_n = b \end{cases}.$$

On résout ensuite ce système en u_0 et q .

Remarque 5.10.

1. Si k et n ont la même parité, la solution n'est pas unique.
2. Alternativement, on peut déterminer la raison q de la suite en calculant de deux façons différentes le quotient $\frac{u_n}{u_k}$ (qui vaut q^{n-k} puis déduire la valeur de u_0 en utilisant que $u_n = u_0 q^n$ donc $u_0 = \frac{u_n}{q^n}$. De cette méthode, on peut extraire une méthode pour déterminer la raison d'une suite géométrique à partir de deux de ses termes et une méthode pour déterminer le terme initial d'une suite géométrique à partir de sa raison et un de ses termes.

Exemple 5.6. Déterminons la raison q et le premier terme u_0 de la suite géométrique $(u_n)_{n \in \mathbf{N}}$ telle que $u_3 = 80$ et $u_6 = 640$. Puisque $(u_n)_{n \in \mathbf{N}}$ est géométrique de raison q et de premier terme u_0 , la Proposition 5.8 assure que $80 = u_3 = q^3 u_0$ et $640 = u_6 = q^6 u_0$. Ainsi, q et u_0 satisfont le système :

$$\begin{cases} q^3 u_0 = 80 \\ q^6 u_0 = 640 \end{cases} .$$

Or, on a :

$$\begin{aligned} \begin{cases} q^3 u_0 = 80 \\ q^6 u_0 = 640 \end{cases} &\iff \begin{cases} q^3 u_0 = 80 \\ q^3 = \frac{q^6 u_0}{q^3 u_0} = \frac{640}{80} = 8 \end{cases} \\ &\iff \begin{cases} u_0 = \frac{80}{q^3} = \frac{80}{2^3} = 10 \\ q = \sqrt[3]{8} = 2 \end{cases} \end{aligned}$$

Ainsi, $(u_n)_{n \in \mathbf{N}}$ est géométrique de raison $q = 2$ et de premier terme $u_0 = 10$.

Méthode pour déterminer à partir de quel rang une suite géométrique dépasse une valeur prescrite

Soit $A > 0$ et soit $(u_n)_{n \in \mathbf{N}}$ la suite géométrique de raison $q > 1$ et de premier terme $u_0 > 0$. Pour déterminer à partir de quel rang N la suite $(u_n)_{n \in \mathbf{N}}$ dépasse A , il suffit d'écrire que :

$$\begin{aligned} u_n \geq A &\iff q^n u_0 \geq A && \text{(d'après la Proposition 5.8)} \\ &\iff q^n \geq \frac{A}{u_0} && \text{(car } u_0 > 0) \\ &\iff \ln(q^n) \geq \ln\left(\frac{A}{u_0}\right) && \text{(car la fonction } \ln \text{ est croissante)} \\ &\iff n \ln(q) \geq \ln(A) - \ln(u_0) && \text{(par les propriétés de } \ln) \\ &\iff n \geq \frac{\ln(A) - \ln(u_0)}{\ln(q)} && \text{(car } q > 1 \text{ et donc } \ln(q) > 0). \end{aligned}$$

Ainsi, le choix de N comme le plus petit entier supérieur à $\frac{\ln(A) - \ln(u_0)}{\ln(q)}$ convient.

Remarque 5.11. La même méthode permet de déterminer, pour $\varepsilon > 0$, à partir de quel rang N une suite géométrique de raison $q \in]-1; 1[$ à ses termes compris entre $-\varepsilon$ et ε .

Exemple 5.7. Déterminons à partir de quel rang N la suite géométrique $(u_n)_{n \in \mathbf{N}}$ de raison

$q = 7$ et de premier terme $u_0 = 5$ dépasse 1000000. On a :

$$\begin{aligned}
 u_n \geq 1000000 &\iff u_0 q^n = 5 \times 7^n \geq 1000000 && \text{(d'après la Proposition 5.8)} \\
 &\iff 7^n \geq \frac{1000000}{5} = 200000 \\
 &\iff \ln(7^n) \geq \ln(200000) && \text{(car la fonction } \ln \text{ est croissante)} \\
 &\iff n \ln(7) \geq \ln(200000) && \text{(par les propriétés de } \ln \text{)} \\
 &\iff n \geq \frac{\ln(200000)}{\ln(7)} \simeq 6,3
 \end{aligned}$$

Ainsi, le choix de $N = 7$ convient.

Sommes partielles

Proposition 5.9. Soit $(u_n)_{n \in \mathbf{N}}$ la suite géométrique de raison $q \neq 1$ et de premier terme u_0 . Alors, pour tout $n \in \mathbf{N}^*$, on a :

$$S_n = \sum_{k=0}^{n-1} u_k = u_0 + \dots + u_{n-1} = u_0 \frac{1 - q^n}{1 - q}.$$

Remarque 5.12. Pour éviter toute erreur, il est conseillé de se souvenir de la formule précédente comme suit :

$$S_n = \text{« premier terme »} \frac{1 - \text{« raison puissance nombre de termes »}}{1 - \text{« raison »}}.$$

On peut alors écrire des formules analogues pour d'autres sommes partielles de termes consécutifs d'une suite géométrique (par nécessairement les n premiers termes), comme dans la proposition suivante.

Proposition 5.10. Soit $(u_n)_{n \in \mathbf{N}}$ la suite géométrique de raison q et de premier terme u_0 . Alors, on a :

$$\begin{aligned}
 \tilde{S}_n &= \sum_{k=1}^n u_k = u_1 + \dots + u_n = u_1 \frac{1 - q^n}{1 - q}, && n \geq 1 \\
 \bar{S}_n &= \sum_{k=0}^n u_k = u_0 + \dots + u_n = u_0 \frac{1 - q^{n+1}}{1 - q}, && n \geq 0 \\
 \check{S}_n &= \sum_{k=1}^{n-1} u_k = u_1 + \dots + u_{n-1} = u_1 \frac{1 - q^{n-1}}{1 - q}, && n \geq 2.
 \end{aligned}$$

Preuve : Plusieurs preuves de ce résultat existent.

Preuve 1 [« à la Gauss »] : Commençons par écrire que, puisque $(u_n)_{n \in \mathbf{N}}$ est la suite géométrique de raison q et de premier terme u_0 ,

$$\begin{aligned}
 S_n &= u_0 + u_1 + \dots + u_{n-2} + u_{n-1} \\
 &= u_0 + qu_0 + \dots + q^{n-2}u_0 + q^{n-1}u_0.
 \end{aligned}$$

L'astuce est alors de retrancher qS_n à S_n :

$$\begin{aligned}(1-q)S_n &= S_n - qS_n \\ &= u_0 + qu_0 + \cdots + q^{n-2}u_0 + q^{n-1}u_0 - q(u_0 + qu_0 + \cdots + q^{n-2}u_0 + q^{n-1}u_0) \\ &= u_0 + qu_0 + \cdots + q^{n-2}u_0 + q^{n-1}u_0 - qu_0 - \cdots - q^{n-2}u_0 - q^{n-1}u_0 - q^n u_0 \\ &= u_0(1 - q^n).\end{aligned}$$

En simplifiant par $1 - q \neq 0$, on obtient que :

$$S_n = u_0 \frac{1 - q^n}{1 - q}.$$

Preuve 2 [par récurrence] : Exercice! □

Corollaire 5.5. Soit $(u_n)_{n \in \mathbf{N}}$ la suite géométrique de raison $q \in]-1; 1[$ et de premier terme u_0 .

Alors, pour tout $n \in \mathbf{N}$, on a :

$$S_\infty = \lim_{n \rightarrow +\infty} S_n = u_0 \frac{1}{1 - q}.$$

Preuve : Puisque $|q| < 1$, on a $\lim_{n \rightarrow +\infty} q^{n+1} = 0$. On en déduit le résultat à l'aide de la Proposition 5.9. □

Exemple 5.8. Soit $(u_n)_{n \in \mathbf{N}}$ la suite géométrique de raison $q = \frac{1}{4}$ et de premier terme $u_0 = 3$. Alors, pour tout $n \in \mathbf{N}$:

$$S_n = u_0 + u_1 + \cdots + u_n = 3 \times \frac{1 - \frac{1}{4^{n+1}}}{1 - \frac{1}{4}} = 3 \times \frac{1 - \frac{1}{4^{n+1}}}{\frac{3}{4}} = 4 - \frac{1}{4^n}$$

et

$$S_\infty = \lim_{n \rightarrow +\infty} S_n = 4.$$

5.3.4 Suites arithmético-géométriques

Définition 5.11. On appelle suite arithmético-géométrique toute suite (récurrente affine d'ordre 1) définie par :

$$\begin{cases} u_{n+1} = au_n + b \\ u_0 \text{ donné} \end{cases}, \quad a, b \in \mathbf{R}.$$

Remarque 5.13. Dans le cas particulier où $a = 1$, on retrouve une suite arithmétique et dans le cas particulier où $b = 0$, on retrouve une suite géométrique. Dans le cas $a \neq 1$, l'étude de la suite $(u_n)_{n \in \mathbf{N}}$ se réalise au moyen de celle de la suite $v_n = u_n - \frac{b}{1-a}$.

Proposition 5.11. *Le terme général de la suite arithmético-géométrique définie par :*

$$\begin{cases} u_{n+1} = au_n + b \\ u_0 \text{ donné} \end{cases}, \quad a, b \in \mathbf{R}, a \neq 1,$$

est donné, pour tout $n \in \mathbf{N}$, par :

$$u_n = a^n \left(u_0 - \frac{b}{1-a} \right) + \frac{b}{1-a}.$$

Preuve : Posons pour tout $n \in \mathbf{N}$:

$$v_n = u_n - \frac{b}{1-a}.$$

On a :

$$v_{n+1} = u_{n+1} - \frac{b}{1-a} = au_n + b - \frac{b}{1-a} = au_n + \frac{b(1-a) - b}{1-a} = a \left(u_n - \frac{b}{1-a} \right) = av_n.$$

Ainsi, la suite $(v_n)_{n \in \mathbf{N}}$ est géométrique de raison a et on a, pour tout $n \in \mathbf{N}$:

$$u_n - \frac{b}{1-a} = v_n = a^n v_0 = a^n \left(u_0 - \frac{b}{1-a} \right)$$

d'où le résultat. □

Exercice 5.3. Étudier la suite arithmético-géométrique définie par :

$$\begin{cases} u_{n+1} = 5u_n - 8 \\ u_0 = 3 \end{cases}.$$

5.4 Complément : la méthode de Newton

La *méthode de Newton* est une méthode numérique de résolution d'équations de la forme $f(x) = 0$ où f est une fonction dérivable.

L'idée est la suivante. Supposons que f soit une fonction dérivable sur un intervalle $]a; b[$ et s'annulant une seule fois sur cet intervalle en r . Rappelons que l'équation de la tangente à la courbe représentative de f en x_0 est $y = f'(x_0)(x - x_0) + f(x_0)$. Si x_0 est assez proche de la racine r , la tangente en x_0 fournit une bonne approximation de la fonction f sur un intervalle contenant r .

On peut donc écrire que :

$$0 = f(r) \simeq f'(x_0)(r - x_0) + f(x_0).$$

On en déduit que, si $f'(x_0) \neq 0$:

$$r \simeq x_0 + \frac{f(x_0)}{f'(x_0)}$$

donc que, si x_0 est proche de r , $x_0 + \frac{f(x_0)}{f'(x_0)}$ l'est encore plus (voir Figure Newt).

On considère donc la suite récurrente définie par :

$$\begin{cases} u_{n+1} = u_n - \frac{f(u_n)}{f'(u_n)} \\ u_0 \text{ donné proche de } r \end{cases} .$$

On montre que, sous certaines conditions, cette suite converge vers r .

Théorème 5.2 (Admis). *Soit f une fonction deux fois continument dérivable sur un intervalle I , s'annulant en un point r intérieur à I et telle que $f'(r) \neq 0$.*

Alors, il existe $\varepsilon > 0$ tel que la suite définie par :

$$\begin{cases} u_{n+1} = u_n - \frac{f(u_n)}{f'(u_n)} \\ u_0 \in]r - \varepsilon; r + \varepsilon[\text{ donné} \end{cases}$$

converge vers r .

S'il existe $\eta > 0$ tel que, pour tout $x \in [r - \eta; r + \eta] \subset I$, on a $|f''(x)| \leq M$ et $|1/f'(x)| \leq m$ alors tout ε vérifiant $0 < \varepsilon < \min(2/(mM); \eta)$ convient.

Remarque 5.14. On peut également avoir des informations sur la vitesse à laquelle la suite $(u_n)_{n \in \mathbf{N}}$ converge vers r . Par exemple, si $f'(r) > 0$ et $f''(x) \geq 0$ pour tout $x \in [r, b]$, alors, pour tout $n \in \mathbf{N}$:

$$0 \leq u_n - r \leq \frac{f(u_n)}{f'(r)}.$$

Exemple 5.9. La méthode de Newton permet par exemple de déterminer une valeur approchée avec une précision arbitraire de $\sqrt{2}$. Pour cela considérons l'équation $f(x) := x^2 - 2 = 0$ avec $x \in [1, 2]$. L'unique solution de cette équation sur l'intervalle $[1, 2]$ est clairement $\sqrt{2}$. On a :

$$f'(x) = 2x \quad \text{et} \quad f''(x) = 2.$$

En particulier, $f'(\sqrt{2}) \neq 0$, $|f''(x)| \leq 2$ et $|1/f'(x)| \leq \frac{1}{2}$. Ainsi, la suite définie par :

$$\begin{cases} u_{n+1} = u_n - \frac{u_n^2 - 2}{2u_n} \\ u_0 = 1,9 \end{cases}$$

converge vers r (Exercice : le montrer en vérifiant que cette suite est décroissante et minorée). On a de plus, pour tout $n \in \mathbf{N}$:

$$0 \leq u_n - r\sqrt{2} \leq \frac{u_n^2 - 2}{2\sqrt{2}} \leq \frac{u_n^2 - 2}{2}.$$

Le tableau suivant donne les premiers termes de la suite et la borne sur la différence $u_n - \sqrt{2}$; on observe que l'on obtient très vite une excellente précision.

Terme de la suite	Borne sur la précision
$u_0 = 1,9$	$\frac{u_0^2 - 2}{2} = 0,805$
$u_1 \simeq 1,4763157895$	$\frac{u_1^2 - 2}{2} \simeq 0,09$
$u_2 \simeq 1,4155197486$	$\frac{u_2^2 - 2}{2} \simeq 0,0018$
$u_3 \simeq 1,414214165$	$\frac{u_3^2 - 2}{2} \simeq 9 \times 10^{-7}$
$u_4 \simeq 1,4142135624$	$\frac{u_4^2 - 2}{2} \simeq 9 \times 10^{-13}$

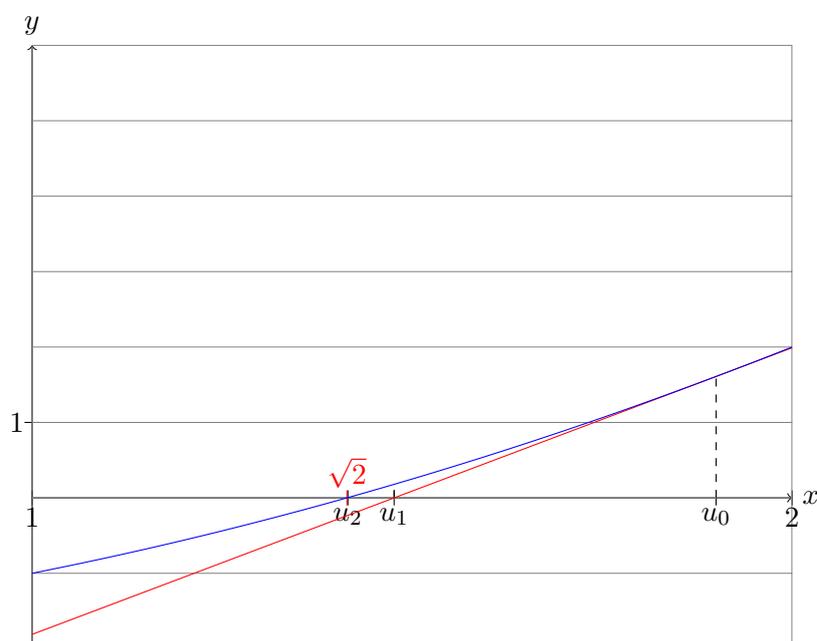


FIGURE 5.1 – Courbe représentative de la fonction $f : x \mapsto x^2 - 2$ sur l'intervalle $[1; 2]$ (bleu), sa tangente en 1,9 (rouge) et les premiers termes de la suite utilisée dans la méthode de Newton dans ce cas.

Chapitre 6

Mathématiques financières

Dans ce chapitre, on s'intéresse à quelques notions de mathématiques financières, notamment les *intérêts*, la *capitalisation*, l'*actualisation* et les *emprunts*.

6.1 Intérêts, capitalisation et actualisation

Les *intérêts* d'un placement ou d'un prêt forment la rémunération de celui-ci. Leur calcul dépend du *taux de rémunération*, de la durée, du capital investi et du fait que l'on considère un placement à intérêts *simples* ou *composés*.

Notations 6.1. *On notera :*

- C_0 le capital investi initialement,
- τ le taux du placement ou du prêt,
- n le nombre de périodes,
- I_n les intérêts de la n^e période,
- $V_n = V_{n-1} + I_n$ la valeur acquise au terme de la n^e période avec $V_0 = C_0$ le capital investi initialement.

6.1.1 Intérêts simples

Définition 6.1. *On dit qu'un capital C_0 est placé à intérêts simples et au taux τ par période si les intérêts de chaque période sont donnés par :*

$$I = C_0\tau.$$

La valeur acquise au bout de n périodes par le placement d'un capital C_0 à intérêts simples, au taux τ est alors :

$$V_n = C_0 + nI.$$

Remarque 6.1.

1. On reconnaît ici que, dans le cas d'un seul versement, la suite $(V_n)_{n \in \mathbf{N}}$ des valeurs acquises est la suite arithmétique de premier terme C_0 et de raison $I = C_0\tau$. Ceci permet de déduire facilement le capital investi, la valeur acquise, le taux ou le nombre de périodes d'un placement à partir des trois autres paramètres. On peut également profiter de ce fait pour déterminer le temps nécessaire à l'obtention d'une valeur acquise prescrite

connaissant le taux et le capital investi ou encore le montant du capital à investir pour obtenir une valeur acquise prescrite connaissant le taux et la durée souhaitée.

2. Les intérêts ne portent, dans ce cas, que sur le capital investi.
3. Dans la pratique, les intérêts simples ne sont que rarement utilisés. Ils le sont généralement pour des placement à court terme (moins d'un an).
4. Il est d'usage de considérer qu'un mois comporte 30 jours et donc une année 360 jours répartis en 25 quinzaines (de 14,4 jours) ou 50 semaines. Les jours sont comptés de manière exacte, en jours pleins. Sur les livrets usuels les intérêts sont calculés par quinzaine.

Exemple 6.1. Si l'on place 500€ à intérêts simples au taux annuel de 0,75%, les intérêts de chaque période s'élèveront à $I = 3,75$ €. Après 20 années, le capital acquis sera $V_{20} = C_0 + 20I = 500 + 20 \times 3,75 = 575$ €.

Exemple 6.2. Si 20000€ placés à intérêts simples ont permis d'obtenir une valeur acquise de 20400€ au bout de 8 mois, la relation $V_n = C_0 + nI$ entraîne que les intérêts de chaque période s'élèvent à $I = \frac{1}{n}(V_n - C_0) = \frac{1}{8}(20400 - 20000) = 50$ €. On en déduit que le taux mensuel du placement est de $\tau = \frac{I}{C_0} = \frac{50}{20000} = 0,25\%$.

Exemple 6.3. On place 1000€ sur un livret à intérêts simples rémunéré au taux annuel de 7%. Déterminons dans combien d'années au minimum, on aura atteint une valeur acquise de 1500€. On veut trouver n tel que :

$$V_n \geq 1500.$$

On obtient successivement :

$$\begin{aligned} V_n \geq 1500 &\iff C_0 + nI = C_0 + nC_0\tau \geq 1500 \\ &\iff 1000 + n \times 1000 \times 0,07 \geq 1500 \\ &\iff 70n \geq 500 \\ &\iff n \geq \frac{500}{70} \simeq 7,1. \end{aligned}$$

Il faudra attendre au moins 8 ans pour une valeur acquise supérieure ou égale à 1500€.

Exercice 6.1. Considérons un placement à intérêts simples rémunéré au taux annuel de 3%. Déterminer le capital à investir pour obtenir, au bout de 5 ans, une valeur acquise de 12000€.

Définition 6.2. Deux taux d'intérêts simples τ_1 et τ_2 , définis sur des périodes de durées P_1 et P_2 respectivement (exprimées en la même unité de temps), sont dit proportionnels si pour tout capital C_0 les valeurs acquises au bout de la même durée sont identiques, c'est-à-dire :

$$C_0\tau_1 \times \frac{n}{P_1} = C_0\tau_2 \times \frac{n}{P_2}, \quad \text{pour tout } C_0, \text{ pour tout } n \in \mathbf{N}.$$

Il est immédiat de vérifier que :

Proposition 6.1. Deux taux d'intérêts simples τ_1 et τ_2 , définis sur des périodes P_1 et P_2 respectivement, sont proportionnels si, et seulement si,

$$\frac{\tau_1}{P_1} = \frac{\tau_2}{P_2}.$$

En particulier, le taux mensuel τ_m^{prop} proportionnel à un taux annuel τ_a est donné par

$$\tau_m^{\text{prop}} = \frac{\tau_a}{12}$$

alors que le taux annuel τ_a^{prop} proportionnel à un taux mensuel τ_m est donné par

$$\tau_a^{\text{prop}} = 12\tau_m.$$

Exemple 6.4. Déterminons le taux trimestriel τ_t^{prop} proportionnel au taux annuel $\tau_a = 2\%$. Un trimestre comporte 3 mois (soit 90 jours) et une année 12 mois (soit 360 jours). Le taux trimestriel τ_t^{prop} proportionnel à τ_a vérifie donc :

$$\frac{\tau_t^{\text{prop}}}{90} = \frac{\tau_a}{360},$$

soit

$$\tau_t^{\text{prop}} = \frac{90}{360}\tau_a = \frac{1}{4} \times 2\% = 0,5\%.$$

6.1.2 Intérêts composés

Définition 6.3. On dit qu'un capital C_0 est placé à intérêts composés et au taux τ par période si les intérêts de la n^e période sont calculés à partir de la valeur acquise V_{n-1} au début de cette période par la formule :

$$I_n = V_{n-1}\tau.$$

On remarque que, pour tout $n \in \mathbf{N}$:

$$V_{n+1} = V_n + I_{n+1} = V_n + \tau V_n = (1 + \tau) V_n,$$

avec $V_0 = C_0$. Ainsi, la suite $(V_n)_{n \in \mathbf{N}}$ des valeurs acquises est la suite géométrique de premier terme C_0 et de raison $1 + \tau$ ce qui prouve la proposition suivante.

Proposition 6.2. Si un capital C_0 est placé à intérêts composés et au taux τ par période la valeur acquise V_n au terme de la n^e période est donnée par :

$$V_n = C_0 (1 + \tau)^n.$$

Remarque 6.2.

1. On a vu que, dans le cas d'un seul versement, la suite $(V_n)_{n \in \mathbf{N}}$ des valeurs acquises est la suite géométrique de premier terme C_0 et de raison $1 + \tau$. Ceci permet de déduire facilement le capital investi, la valeur acquise, le taux ou le nombre de périodes d'un placement à partir des trois autres paramètres. On peut également profiter de ce fait pour déterminer le temps nécessaire à l'obtention d'une valeur acquise prescrite connaissant le taux et le capital investi ou encore le montant du capital à investir pour obtenir une valeur acquise prescrite connaissant le taux et la durée souhaitée.
2. Les intérêts d'une période portent ici sur la valeur acquise au début de la période.

3. Les placement à intérêts composés sont les plus fréquents. Lorsqu'il n'est pas précisé si un placement est rémunéré par des intérêts simples ou composés, on considère par défaut qu'il s'agit d'intérêts composés.

Exemple 6.5. Si l'on place 200€ à intérêts composés au taux annuel de 1,25%, la valeur acquise après 10 années sera $V_{10} = C_0(1 + \tau)^{10} = 200 \times 1,0125^{10} \simeq 226,45$ €. Le montant total des intérêts perçus en 10 ans s'élève donc à $226,45 - 200 = 26,45$ €.

Exemple 6.6. Si 3000€ placés à intérêts composés on permis d'obtenir une valeur acquise de 4000€ au bout de 20 ans, la relation $V_n = C_0(1 + \tau)^n$ permet de déterminer le taux annuel du placement comme suit. On a :

$$\begin{aligned} V_{20} = 4000 &\iff C_0(1 + \tau)^{20} = 4000 \\ &\iff 3000(1 + \tau)^{20} = 4000 \\ &\iff (1 + \tau)^{20} = \frac{4}{3} \\ &\iff 1 + \tau = \left(\frac{4}{3}\right)^{\frac{1}{20}} \\ &\iff \tau = \left(\frac{4}{3}\right)^{\frac{1}{20}} - 1 \end{aligned}$$

Ainsi, le taux annuel du placement est de $\tau = \left(\frac{4}{3}\right)^{\frac{1}{20}} - 1 = 1,45\%$.

Exemple 6.7. Un produit financier est actuellement rémunéré au taux annuel de 0,75%. Déterminons combien d'années sont nécessaires avant de doubler un capital placé si aucun autre versement n'est fait. On veut trouver n tel que :

$$V_n \geq 2C_0.$$

On obtient successivement :

$$\begin{aligned} V_n \geq 2C_0 &\iff C_0 \times 1,0075^n \geq 2C_0 \\ &\iff 1,0075^n \geq 2 \\ &\iff n \ln(1,0075) = \ln(1,0075^n) \geq \ln(2) \\ &\iff n \geq \frac{\ln(2)}{\ln(1,0075)} \simeq 92,77 \end{aligned}$$

Il faudra être patient et attendre au moins 93 ans pour doubler la mise !

Exercice 6.2. Considérons un placement à intérêts composés rémunéré au taux annuel de 4%. Déterminer le capital à investir pour obtenir, au bout de 3 ans, une valeur acquise de 100000€.

Définition 6.4. Deux taux d'intérêts composés τ_1 et τ_2 , définis sur des périodes de durées P_1 et P_2 respectivement (exprimées en la même unité de temps), sont dit équivalents si pour tout capital C_0 les valeurs acquises au bout de la même durée sont identiques, c'est-à-dire :

$$C_0(1 + \tau_1)^{\frac{n}{P_1}} = C_0(1 + \tau_2)^{\frac{n}{P_2}}, \quad \text{pour tout } C_0, \text{ pour tout } n \in \mathbf{N}.$$

Il est immédiat de vérifier que :

Proposition 6.3. *Deux taux d'intérêts composés τ_1 et τ_2 , définis sur des périodes P_1 et P_2 respectivement, sont équivalents si, et seulement si,*

$$(1 + \tau_1)^{\frac{1}{P_1}} = (1 + \tau_2)^{\frac{1}{P_2}}.$$

En particulier, le taux mensuel $\tau_m^{\text{équiv}}$ équivalent à un taux annuel τ_a est donné par

$$\tau_m^{\text{équiv}} = (1 + \tau_a)^{\frac{1}{12}} - 1$$

alors que le taux annuel $\tau_a^{\text{équiv}}$ équivalent à un taux mensuel τ_m est donné par

$$\tau_a^{\text{équiv}} = (1 + \tau_m)^{12} - 1.$$

Exemple 6.8. Rappelons qu'il est d'usage de considérer 25 quinzaines par an. Le taux équivalent par quinzaine au taux annuel de rémunération $\tau_a = 0,75\%$ est donc donné par :

$$\tau_q^{\text{équiv}} = (1 + \tau_a)^{\frac{1}{25}} - 1 = 1,0075^{\frac{1}{25}} - 1 \simeq 0,0000299\%.$$

6.1.3 Taux nominaux, périodiques et effectifs

Définition 6.5. *Soit un produit financier dont les capitalisations ont lieu n fois par an au taux périodique $\tau_{\text{pér}}$.*

On appelle taux nominal du produit :

$$\tau_{\text{nom}} = n\tau_{\text{pér}}.$$

Dans le cadre d'intérêts simples, il n'y a aucun problème et le taux nominal est bien le taux annuel proportionnel au taux périodique. Par contre, dans le cadre d'intérêts composés le taux nominal n'est pas le taux annuel équivalent au taux périodique. Ceci conduit à la définition du *taux annuel effectif* correspondant à un taux nominal.

Définition 6.6. *Soit un produit financier dont les capitalisations ont lieu n fois par an et affiché au taux nominal τ_{nom} .*

On appelle taux d'intérêt annuel effectif du produit le taux annuel équivalent au taux périodique correspondant, c'est-à-dire le taux :

$$\tau_{\text{eff}} = (1 + \tau_{\text{pér}})^n - 1 = \left(1 + \frac{\tau_{\text{nom}}}{n}\right)^n - 1.$$

Exemple 6.9. Le livret A est un produit financier rémunéré depuis le 1^{er} février 2020 à son taux nominal plancher de 0,5%. Les intérêts sont capitalisés au 1^e et 16 de chaque mois, soit sur 24 périodes. Son taux périodique est donc $\tau_{\text{pér}} = \frac{0,005}{24} \simeq 0,02083\%$. Le taux effectif de ce livret est :

$$\tau_{\text{eff}} = (1 + \tau_{\text{pér}})^{24} - 1 = \left(1 + \frac{0,005}{24}\right)^{24} - 1 \simeq 0,5012\%.$$

Le même raisonnement que dans l'Exemple 6.7, montre que, si sont taux n'évolue pas et sans autre versement, le Livret A permet de doubler un capital placé au bout de 139 ans.

Les intérêts considérés ci-dessous sont simples ou composés et les taux effectifs.

6.1.4 Capitalisation

Définition 6.7. La capitalisation est le calcul d'une somme future à partir d'une valeur actuelle. En d'autre terme, il s'agit du procédé permettant de passer d'un capital investi C_0 à la valeur acquise V_n à partir de ce capital à l'issue de n périodes.

Remarque 6.3. Nous avons vu, dans la Sous-section 6.1.1, que V_n s'obtient, dans le cas d'intérêts simples au taux τ par période, grâce à la formule :

$$V_n = C_0 + nI = C_0 + nC_0\tau.$$

Dans la Sous-section 6.1.2, nous avons vu que V_n s'obtient, dans le cas d'intérêts composé au taux τ par période, grâce à la formule :

$$V_n = C_0(1 + \tau)^n.$$

6.1.5 Actualisation

Définition 6.8. L'actualisation est le calcul d'une somme présente à partir d'une valeur future. En d'autre terme, il s'agit du procédé permettant de passer de la valeur acquise V_n à partir d'un capital investi C_0 à l'issue de n périodes au capital investi C_0 lui même. On appelle $V_0 = C_0$ la valeur actuelle de la valeur future V_n .

Remarque 6.4. Il s'agit exactement de faire ce qui a été demandé dans les Exercices 6.1 et 6.2!

Si les intérêts considérés sont **simples** au taux τ par période, V_n s'obtient grâce à la formule :

$$V_n = V_0 + nV_0\tau = V_0(1 + n\tau).$$

On en déduit que dans ce cas, la valeur actuelle de V_n est donnée, dans ce cas, par :

$$V_0 = \frac{V_n}{1 + n\tau}.$$

Si les intérêts considérés sont **composés** au taux τ par période, V_n s'obtient grâce à la formule :

$$V_n = V_0(1 + \tau)^n.$$

On en déduit que dans ce cas, la valeur actuelle de V_n est donnée, dans ce cas, par :

$$V_0 = V_n(1 + \tau)^{-n}.$$

Exemple 6.10. [Solution de l'Exercice 6.1] Considérons un placement à intérêts simples rémunéré au taux annuel de 3%. Déterminer le capital à investir pour obtenir, au bout de 5 ans, une valeur acquise de 12000€ revient à trouver V_0 tel que $V_5 = 12000$ €. Puisque les intérêts considérés sont simples, on écrit successivement que :

$$\begin{aligned} V_5 = 12000 &\iff V_0 + 5V_0\tau = 12000 \\ &\iff V_0(1 + 5\tau) = V_0 \times 1,15 = 12000 \\ &\iff V_0 = \frac{12000}{1,15} \simeq 10434,78. \end{aligned}$$

Pour ce placement, la valeur actuelle de la somme future de 12000€ dans 5 ans est 10434,78€.

Exemple 6.11. [Solution de l'Exercice 6.2] Considérons un placement à intérêts composés rémunéré au taux annuel de 4%. Déterminer le capital à investir pour obtenir, au bout de 3 ans, une valeur acquise de 100000€ revient à trouver V_0 tel que $V_3 = 100000$ €. Puisque les intérêts considérés sont simples, on écrit successivement que :

$$\begin{aligned} V_3 = 100000 &\iff V_0 (1 + \tau)^3 = 100000 \\ &\iff V_0 = 100000 (1 + \tau)^{-3} = 100000 \times 1,04^{-3} \simeq 88899,66. \end{aligned}$$

Pour ce placement, la valeur actuelle de la somme future de 100000€ dans 3 ans est 88899,66€.

Un schéma intuitif pour éviter les erreurs de compte de périodes

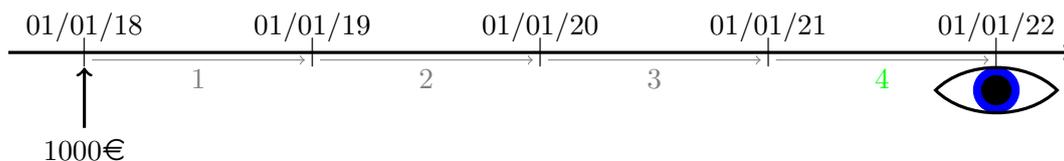
Pour calculer la valeur actuelle ou capitalisée, on peut utiliser un schéma contenant :

- un axe horizontal représentant le temps et les périodes,
- une ou plusieurs flèches représentant le(s) versement(s) ou retrait(s) selon le sens de la flèche,
- les montants versés ou retirés,
- et un œil indiquant l'instant auquel on « regarde » la valeur des sommes versées et retirées.

On sera attentif au sens dans lequel on parcourt les périodes en allant de la (ou des) flèche(s) vers l'œil (puisque'il influe sur le signe de l'exposant ou du multiplicande). Bien qu'ils soient utilisables quelque soit le type d'intérêt considéré (simple ou composé), on donne ci-dessous quelques exemples dans le cas d'intérêts composés d'usage plus courant. Il faut, bien entendu, être vigilant à adapter les formules au type d'intérêts utilisé.

Exemple 6.12. On a placé 1000 euros le 01/01/2018 sur un livret rémunéré au taux composé annuel τ de 3%. Quelle sera la valeur capitalisée de cette somme le 01/01/2022 ?

Le schéma prend la forme suivante.

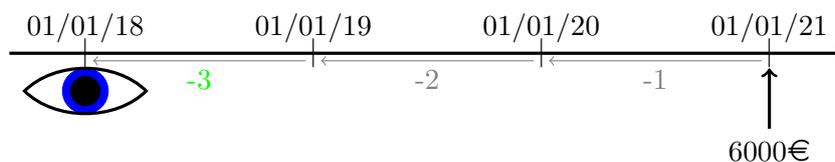


Ainsi, la valeur capitalisée au 01/01/2022 est :

$$1000 \times (1 + \tau)^4 = 1000 \times 1,03^4 \simeq 1125,51\text{€}.$$

Exemple 6.13. On a placé une somme le 01/01/2018 sur un livret rémunéré au taux composé annuel τ de 2%. Sachant qu'il n'y a pas eu d'autre retrait ou versement et qu'il y avait 6000€ sur le livret le 01/01/2021, quelle somme a été placée le 01/01/2018 ?

Le schéma prend la forme suivante.

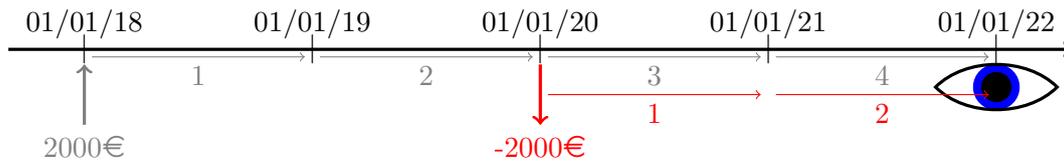


Ainsi, la valeur de la somme placée est la valeur actualisée au 01/01/2018 :

$$6000 \times 1,02^{-3} \simeq 5653,93\text{€}.$$

Exemple 6.14. On a placé 2000 euros le 01/01/2018 sur un livret rémunéré au taux composé annuel τ de 1% puis retiré 2000 euros le 01/01/2020. S'il n'y a pas d'autre versement ou retrait, de combien disposera-t-on le 01/01/2022 ?

Le schéma prend la forme suivante.



Ainsi, la somme dont on disposera au 01/01/2022 est :

$$2000 \times 1,01^4 + (-2000) \times 1,01^2 \simeq 41,01\text{€}.$$

6.2 Versements périodiques : les annuités

Les placements et emprunts à intérêts simples n'étant utilisés que très marginalement, dans toute la suite nous nous restreindrons aux cas des placements ou emprunts à intérêts composés. Les méthodes, concepts et notions présentés ci-dessous peuvent toute fois être transposés dans le cadre d'intérêts simples.

Un placement ou un prêt est rarement effectué en une seule opération. Des versements peuvent être effectués à intervalle de temps réguliers (mois, années, ...). Ceci nous conduit à la notion d'*annuité*.

Définition 6.9. Des annuités sont une suite (finie) de flux financier réalisés à intervalle de temps réguliers.

Remarque 6.5. Les annuités dépendent de la date de premier flux, de la période, du nombre de flux et de leurs montants. Le fait qu'elles soient versées en début ou fin de période joue également un rôle important. Un versement correspond à une annuité positive, un retrait à une annuité négative ; certaines annuités peuvent être nulles.

Notation 6.2. On note généralement les annuités a_1, a_2, \dots

6.2.1 Annuités de début de période

On parle d'annuités de *début de période* lorsque les flux monétaires ont lieu au début de chaque période : a_1 est versée au début de la première période, a_2 au début de la deuxième période, ... Dans ce cas a_k , compte dans le calcul des intérêts de la k^e période.

Calcul de la valeur acquise

Pour déterminer la valeur acquise, à l'issue de n périodes, par le versement d'annuités a_1, \dots, a_n en début de période au taux τ par période, on peut dresser un tableau faisant apparaître montant des annuités et valeur capitalisée de chacune au terme du placement. Un tel tableau prend la forme suivante.

n° de la période	1	...	k	...	n
Annuité a_i	a_1	...	a_k	...	a_n
Valeur capitalisée grâce à a_i	$a_1 (1 + \tau)^n$...	$a_k (1 + \tau)^{n+1-k}$...	$a_n (1 + \tau)$

Pour obtenir la dernière ligne de ce tableau, on a utilisé que a_1 est placée durant n périodes (à intérêts composés) et donc produira une valeur de $a_1 (1 + \tau)^n$ au terme du placement. De même, a_2 est placé durant $n - 1$ périodes et produira une valeur capitalisée de $a_2 (1 + \tau)^{n-1}$ au terme du placement. Plus généralement, a_k est placé durant $n + 1 - k$ périodes et produira une valeur capitalisée de $a_k (1 + \tau)^{n+1-k}$ au terme du placement.

La *valeur acquise* par la suite d'annuités a_1, \dots, a_k s'obtient simplement en sommant les valeur capitalisées correspondant à chacun des versements. Nous venons donc de voir que :

Proposition 6.4. *La valeur acquise, à l'issue de n périodes, par le versement d'annuités a_1, \dots, a_n en début de période au taux τ par période est donnée par :*

$$V_n = \sum_{k=1}^n a_k (1 + \tau)^{n+1-k} = a_1 (1 + \tau)^n + a_2 (1 + \tau)^{n-1} + \dots + a_n (1 + \tau).$$

Exemple 6.15. Une personne ouvre un PEL rémunéré au taux de 0,5% mensuel et y place 1000€ lors de l'ouverture puis 50€ par mois. À la fin de la première année, il disposera de :

$$\begin{aligned} V_n &= 1000 \times 1,005^{12} + 50 \times 1,005^{11} + 50 \times 1,005^{10} + \dots + 50 \times 1,005 \\ &= 1000 \times 1,005^{12} + 50 \sum_{k=1}^{11} 1,005^k \\ &= 1000 \times 1,005^{12} + 50 \times 1,005 \times \frac{1,005^{11} - 1}{0,005} \\ &= 1628,46\text{€}. \end{aligned}$$

Elle aura donc versé $1000 + 50 \times 11 = 1550\text{€}$ et reçu $1628,46 - 1550 = 78,46\text{€}$ d'intérêts.

Calcul de la valeur actuelle

Pour déterminer la valeur actuelle du versement d'annuités a_1, \dots, a_n en début de période au taux τ par période, on peut dresser un tableau faisant apparaître montant des annuités et valeur actualisée de chacune au terme du placement. Un tel tableau prend la forme suivante.

n° de la période	1	...	k	...	n
Annuité a_i	a_1	...	a_k	...	a_n
Valeur actualisée a_i	a_1	...	$a_k (1 + \tau)^{-k+1}$...	$a_n (1 + \tau)^{-n+1}$

Pour obtenir la dernière ligne de ce tableau, on a utilisé que la valeur actuelle de a_1 placée tout de suite est simplement a_1 . De même, la valeur actuelle de a_2 , placée au début de la deuxième période est, $a_2(1 + \tau)^{-1}$. Plus généralement, la valeur actuelle de a_k , au début de la k^e période, est $a_k(1 + \tau)^{1-k}$.

La valeur actuelle de la suite d'annuités a_1, \dots, a_k s'obtient simplement en sommant les valeurs actuelles correspondant à chacun des versements. Nous venons donc de voir que :

Proposition 6.5. *La valeur actuelle des annuités a_1, \dots, a_n versées en début de période au taux τ par période est donné par :*

$$V_0 = \sum_{k=1}^n a_k (1 + \tau)^{1-k} = a_1 + a_2 (1 + \tau)^{-1} + \dots + a_n (1 + \tau)^{1-n}.$$

Remarque 6.6. En pratique, le calcul de la valeur actuelle d'une suite d'annuités permet de déterminer le montant (actuel) que permet de rembourser cette suite d'annuités, autrement dit le montant que l'on peut emprunter si l'on est prêt à rembourser a_1 durant la première période, a_2 durant la deuxième période, ... Il est rare de devoir rembourser dès le premier jour d'un prêt et on considère plutôt des annuités de fin de période dans ce cadre.

Exemple 6.16. Une personne emprunte au taux de 2% sur 3 mois. Elle est prête à verser 100€ au début du premier mois, 150€ au début du deuxième mois et 120€ au début du troisième mois. La valeur actuelle de cette suite d'annuités est :

$$V_0 = 100 + 150 \times 1,02^{-1} + 120 \times 1,02^{-2} \simeq 362,40€.$$

Elle peut donc emprunter 362,40€ et paiera $100 + 150 + 120 - 362,40 = 7,60€$ d'intérêts.

6.2.2 Annuités de fin de période

On parle d'annuités de *fin de période* lorsque les flux monétaires ont lieu à la fin de chaque période : a_1 est versé à la fin de la première période, a_2 à la fin de la deuxième période, ... Dans ce cas a_k , ne compte pas dans le calcul des intérêts de la k^e période mais dans celui des intérêts de la $(k + 1)^e$ période.

Calcul de la valeur acquise

Pour déterminer la valeur acquise, à l'issue de n périodes, par le versement d'annuités a_1, \dots, a_n en fin de période au taux τ par période, on peut, comme ci-dessus, dresser un tableau faisant apparaître le montant des annuités et la valeur capitalisée de chacune au terme du placement. Un tel tableau prend la forme suivante.

n° de la période	1	...	k	...	n
Annuité a_i	a_1	...	a_k	...	a_n
Valeur capitalisée grâce à a_i	$a_1(1 + \tau)^{n-1}$...	$a_k(1 + \tau)^{n-k}$...	a_n

Pour obtenir la dernière ligne de ce tableau, on a utilisé que a_1 est placée durant $n - 1$ périodes complètes (à intérêts composés) et donc produira une valeur de $a_1(1 + \tau)^{n-1}$ au terme du placement. De même, a_2 est placée durant $n - 2$ périodes et produira une valeur

capitalisée de $a_2(1 + \tau)^{n-2}$ au terme du placement. Plus généralement, a_k est placée durant $n - k$ périodes et produira une valeur capitalisée de $a_k(1 + \tau)^{n-k}$ au terme du placement.

La *valeur acquise* par la suite d'annuités a_1, \dots, a_k s'obtient simplement en sommant les valeurs capitalisées correspondant à chacun des versements. Nous venons donc de voir que :

Proposition 6.6. *La valeur acquise, à l'issue de n périodes, par le versement d'annuités a_1, \dots, a_n en fin de période au taux τ par période est donnée par :*

$$V_n = \sum_{k=1}^n a_k (1 + \tau)^{n-k} = a_1 (1 + \tau)^{n-1} + a_2 (1 + \tau)^{n-2} + \dots + a_n.$$

Exercice 6.3. Reprendre l'Exemple 6.15 en considérant des versements en fin de période et non en début de période.

Calcul de la valeur actuelle

Pour déterminer la valeur actuelle du versement d'annuités a_1, \dots, a_n en fin de période au taux τ par période, on peut dresser, comme précédemment, un tableau faisant apparaître montant des annuités et valeur actualisée de chacune au terme du placement. Un tel tableau prend la forme suivante.

n° de la période	1	...	k	...	n
Annuité a_i	a_1	...	a_k	...	a_n
Valeur actualisée a_i	$a_1(1 + \tau)^{-1}$...	$a_k(1 + \tau)^{-k}$...	$a_n(1 + \tau)^{-n}$

Pour obtenir la dernière ligne de ce tableau, on a utilisé que la valeur actuelle de a_1 placée à la fin de la première période est simplement $a_1(1 + \tau)^{-1}$. De même, la valeur actuelle de a_2 , placée à la fin de la deuxième période est, $a_2(1 + \tau)^{-2}$. Plus généralement, la valeur actuelle de a_k , à la fin de la k^e période, est $a_k(1 + \tau)^{-k}$.

La *valeur actuelle* de la suite d'annuités a_1, \dots, a_k s'obtient simplement en sommant les valeurs actuelles correspondant à chacun des versements. Nous venons donc de voir que :

Proposition 6.7. *La valeur actuelle des annuités a_1, \dots, a_n versées en fin de période au taux τ par période est donnée par :*

$$V_0 = \sum_{k=1}^n a_k (1 + \tau)^{-k} = a_1 (1 + \tau)^{-1} + a_2 (1 + \tau)^{-2} + \dots + a_n (1 + \tau)^{-n}.$$

Exercice 6.4. Reprendre l'Exemple 6.16 en considérant des versements en fin de période et non en début de période.

6.2.3 Cas des annuités constantes

On parle d'*annuités constantes* lorsque $a_1 = a_2 = \dots = a_n = a$ pour une certaine constante a . Les valeurs acquises et actuelles se réécrivent plus simplement dans ce cas.

Proposition 6.8. *1. La valeur acquise par le versement de n annuités constantes égales à a en début de période, au taux τ est donnée par :*

$$V_n = a(1 + \tau) \frac{(1 + \tau)^n - 1}{\tau}.$$

2. La valeur actuelle du versement de n annuités constantes égales à a en début de période, au taux τ est donnée par :

$$V_0 = a(1 + \tau) \frac{1 - (1 + \tau)^{-n}}{\tau}.$$

3. La valeur acquise par le versement de n annuités constantes égales à a en fin de période, au taux τ est donnée par :

$$V_n = a \frac{(1 + \tau)^n - 1}{\tau}.$$

4. La valeur actuelle du versement de n annuités constantes égales à a en fin de période, au taux τ est donnée par :

$$V_0 = a \frac{1 - (1 + \tau)^{-n}}{\tau}.$$

Preuve : On utilise, dans chaque cas, que l'on sait calculer les sommes partielles de suites géométriques (voir Chapitre 5). En particulier :

$$\begin{aligned} \sum_{k=1}^n \left(\frac{1}{1 + \tau} \right)^k &= \frac{1}{1 + \tau} \times \frac{1 - \frac{1}{(1 + \tau)^n}}{1 - \frac{1}{1 + \tau}} = \frac{1}{1 + \tau} \times \frac{\frac{(1 + \tau)^n - 1}{(1 + \tau)^n}}{\frac{1 + \tau - 1}{1 + \tau}} \\ &= (1 + \tau)^{-n} \frac{(1 + \tau)^n - 1}{\tau}. \end{aligned}$$

1. On écrit que :

$$\begin{aligned} V_n &= \sum_{k=1}^n a(1 + \tau)^{n+1-k} = a \sum_{k=1}^n (1 + \tau)^{n+1-k} \\ &= a(1 + \tau)^{n+1} \sum_{k=1}^n (1 + \tau)^{-k} = a(1 + \tau)^{n+1} \sum_{k=1}^n \left(\frac{1}{1 + \tau} \right)^k \\ &= a(1 + \tau)^{n+1} (1 + \tau)^{-n} \frac{(1 + \tau)^n - 1}{\tau} = a(1 + \tau) \frac{(1 + \tau)^n - 1}{\tau}. \end{aligned}$$

2. On écrit que :

$$\begin{aligned} V_0 &= \sum_{k=1}^n a(1 + \tau)^{1-k} = a \sum_{k=1}^n (1 + \tau)^{1-k} \\ &= a(1 + \tau) \sum_{k=1}^n (1 + \tau)^{-k} = a(1 + \tau) \sum_{k=1}^n \left(\frac{1}{1 + \tau} \right)^k \\ &= a(1 + \tau) (1 + \tau)^{-n} \frac{(1 + \tau)^n - 1}{\tau} = a(1 + \tau) \frac{1 - (1 + \tau)^{-n}}{\tau}. \end{aligned}$$

3. On écrit que :

$$\begin{aligned} V_n &= \sum_{k=1}^n a(1 + \tau)^{n-k} = a \sum_{k=1}^n (1 + \tau)^{n-k} \\ &= a(1 + \tau)^n \sum_{k=1}^n (1 + \tau)^{-k} = a(1 + \tau)^n \sum_{k=1}^n \left(\frac{1}{1 + \tau} \right)^k \\ &= a(1 + \tau)^n (1 + \tau)^{-n} \frac{(1 + \tau)^n - 1}{\tau} = a \frac{(1 + \tau)^n - 1}{\tau}. \end{aligned}$$

4. On écrit que :

$$\begin{aligned} V_0 &= \sum_{k=1}^n a(1+\tau)^{-k} = a \sum_{k=1}^n \left(\frac{1}{1+\tau}\right)^k = a(1+\tau)^{-n} \frac{(1+\tau)^n - 1}{\tau} \\ &= a \frac{1 - (1+\tau)^{-n}}{\tau}. \end{aligned}$$

□

6.3 Emprunts indivis

On considère dans cette section des annuités de fin de période et des intérêts composés. Les taux sont effectifs.

Définition 6.10. *Un emprunt indivis est un emprunt que souscrit un emprunteur auprès d'un et un seul prêteur. Il fait l'objet d'un contrat engageant l'emprunteur à verser n annuités a_1, \dots, a_n au prêteur pour rembourser le capital C prêté lors de la souscription ainsi que les intérêts calculés au taux τ fixé dans le contrat.*

Chaque annuité a_k se décompose en :

- un amortissement, noté m_k , servant à rembourser le capital emprunté
- un paiement des intérêts de la période $I_k = C_{k-1}\tau$, où C_{k-1} désigne le capital restant dû au début de la période durant laquelle a lieu le versement de la k^e annuité.

On présente ces données sous la forme d'un *tableau d'amortissement* :

Période n°	Capital restant dû	Intérêts	Amortissement	Annuité
1	C_0	$I_1 = C_0\tau$	m_1	$a_1 = I_1 + m_1$
2	$C_1 = C_0 - m_1$	$I_2 = C_1\tau$	m_2	$a_2 = I_2 + m_2$
...
k	$C_{k-1} = C_{k-2} - m_{k-1}$	$I_k = C_{k-1}\tau$	m_k	$a_k = I_k + m_k$
...
n	$C_{n-1} = C_{n-2} - m_{n-1}$	$I_n = C_{n-1}\tau$	m_n	$a_n = I_n + m_n$

Remarque 6.7. Le remboursement de certains emprunts ne débute qu'après une période fixée à l'avance, c'est par exemple le cas pour la plupart des prêts étudiants. Il convient alors de calculer la valeur capitalisée C_0 de la somme empruntée C avant le début des versements (voir Exercice 6.7). Il faut bien comprendre que C_0 est le montant dû au début de la période durant laquelle a lieu le versement de la première annuité alors que C est le capital emprunté initialement. Ces deux quantités coïncident lorsqu'il n'y a pas de report du paiement de la première annuité mais diffèrent lorsque celui-ci est différé.

Attention : Lorsqu'une annuité a_k est nulle le capital restant dû au début de la $(k+1)^e$ période est donné par :

$$C_k = C_{k-1} + I_k.$$

Proposition 6.9. 1. On a :

$$C_0 = \sum_{k=1}^n m_k = m_1 + m_2 + \dots + m_n.$$

2. Pour tout $k \in \{1, \dots, n-1\}$, on a :

$$C_k = C_0 - \sum_{j=1}^k m_j = C_0 - m_1 - m_2 - \dots - m_k.$$

En particulier,

$$C_{n-1} = m_n.$$

3. Pour tout $k \in \{1, \dots, n-1\}$, on a :

$$a_{k+1} - a_k = m_{k+1} - (1 + \tau)m_k.$$

4. On a :

$$a_n = m_n(1 + \tau).$$

Preuve :

1. Par définition, les amortissement servent à rembourser (exactement) le capital emprunté.
2. On a $C_1 = C_0 - m_1$, $C_2 = C_1 - m_2 = C_0 - m_1 - m_2$, \dots , $C_k = C_{k-1} - m_k = C_0 - \sum_{j=1}^k m_j$, \dots , $C_{n-1} = C_{n-2} - m_{n-1} = C_0 - \sum_{j=1}^{n-1} m_j$. Puisque $C_0 = \sum_{k=1}^n m_k$, on en déduit que :

$$C_{n-1} = C_0 - \sum_{j=1}^{n-1} m_j = \sum_{k=1}^n m_k - \sum_{j=1}^{n-1} m_j = m_1 + \dots + m_{n-1} + m_n - m_1 - \dots - m_{n-1} = m_n.$$

3. On a, pour tout $k \in \{1, \dots, n-1\}$:

$$\begin{aligned} a_{k+1} - a_k &= m_{k+1} + C_k \tau - (m_k + C_{k-1} \tau) \\ &= m_{k+1} + (C_{k-1} - m_k) \tau - m_k - C_{k-1} \tau \\ &= m_{k+1} + \cancel{C_{k-1} \tau} - m_k \tau - m_k - \cancel{C_{k-1} \tau} \\ &= m_{k+1} - (1 + \tau) m_k. \end{aligned}$$

4. On a :

$$a_n = I_n + m_n = C_{n-1} \tau + m_n.$$

Or, $C_{n-1} = m_n$, donc

$$a_n = m_n \tau + m_n = m_n(1 + \tau).$$

□

Exercice 6.5. Le premier amortissement d'un emprunt est $m_1 = 1000\text{€}$ et le deuxième $m_2 = 2000\text{€}$. Les deux premières annuités sont respectivement $a_1 = 2000\text{€}$ et $a_2 = 2950\text{€}$. Déterminer le taux τ de l'emprunt, le montant des intérêts de la première période I_1 puis le capital emprunté C_0 .

Solution : D'après la proposition précédente, on sait que :

$$a_2 - a_1 = m_2 - (1 + \tau)m_1,$$

soit

$$950 = 2950 - 2000 = 2000 - (1 + \tau)1000.$$

On en déduit que :

$$1000(1 + \tau) = 2000 - 950 = 1050,$$

puis que :

$$1 + \tau = \frac{1050}{1000} = 1,05.$$

Ainsi, le taux de l'emprunt est de 5%. On a $I_1 = a_1 - m_1 = 2000 - 1000 = 1000\text{€}$.
Puisque $I_1 = C_0\tau$, on obtient que le capital emprunté est :

$$C_0 = \frac{I_1}{\tau} = \frac{1000}{0,05} = 20000\text{€}.$$

Exercice 6.6. Un emprunt a été souscrit sur 4 mois et on ne dispose que des informations synthétisées dans le tableau suivant.

Moisn° k	Capital restant dû C_{k-1}	Intérêts I_k	Amortissement m_k	Annuité a_k
1			200	240
2			300	
3			300	
4			200	

- Déterminer C_0 et I_1 . En déduire, le taux mensuel τ de l'emprunt.
- Compléter le tableau d'amortissement.

Solution :

- On a :

$$C_0 = m_1 + m_2 + m_3 + m_4 = 200 + 300 + 300 + 200 = 1000\text{€}$$

et

$$I_1 = a_1 - m_1 = 240 - 200 = 40\text{€}.$$

Puisque $I_1 = C_0\tau$, on en déduit que le taux mensuel est

$$\tau = \frac{I_1}{C_0} = \frac{40}{1000} = 4\%.$$

- Le tableau d'amortissement complet est le suivant.

Moisn° k	Capital restant dû C_{k-1}	Intérêts I_k	Amortissement m_k	Annuité a_k
1	1000	40	200	240
2	800	32	300	332
3	500	20	300	320
4	200	8	200	208

Exercice 6.7. Un emprunt de 10000€ est contracté au taux annuel de 1,75%. Aucun versement n'est demandé la première année, puis 5 mensualités sont exigées avec des amortissements constants sur cette période.

1. Quel est le capital dû au début du remboursement ?
2. Déterminer le taux mensuel à appliquer.
3. Déterminer le montant des amortissements.
4. Dresser le tableau d'amortissement.

Solution :

1. Le remboursement du capital et des intérêts débute lors du 13^e mois. La valeur due au début de ce mois est composée du capital emprunté et des intérêts d'une année complète :

$$C_0 = C(1 + \tau) = 10000 \times 1,0175 = 10175\text{€}.$$

2. Le taux mensuel à appliquer est le taux mensuel équivalent au taux annuel $\tau_a = 1,75\%$. Celui-ci est donné par :

$$\tau_m^{\text{equiv}} = (1 + \tau_a)^{\frac{1}{12}} - 1 = 1,0175^{\frac{1}{12}} - 1 \simeq 0,14468\%.$$

3. Puisqu'il n'y a aucun versement durant les 12 premiers mois, on a $m_1 = m_2 = \dots = m_{12} = 0$ (et $a_1 = a_2 = \dots = a_{12} = 0$). Les cinq amortissements suivants sont égaux et permettent de rembourser le capital restant dû au bout d'un an. On a donc :

$$5m_{13} = m_{13} + m_{14} + m_{15} + m_{16} + m_{17} = C_{12}.$$

D'où

$$m_{13} = m_{14} = m_{15} = m_{16} = m_{17} = \frac{C_{12}}{5} = \frac{10175}{5} = 2035\text{€}.$$

4. Le tableau d'amortissement prend la forme suivante.

Période n° k	Capital restant dû C_{k-1}	Intérêts I_k	Amortissement m_k	Annuité a_k
1	10000	14,47	-14,47	0
2	10014,47	14,49	-14,49	0
...
12	10160,30	14,70	-14,70	0
13	10175	14,72	2035	2049,72
14	8140	11,78	2035	2046,78
15	60105	8,83	2035	2043,83
16	4070	5,89	2035	2040,89
17	2035	2,94	2035	2037,94

6.3.1 Emprunts *in fine*

Définition 6.11. *Un emprunt indivis est dit in fine si le remboursement du capital C_0 s'effectue en une seule fois, à la fin du contrat.*

Remarque 6.8. Dans le cas d'un emprunt *in fine* les $n - 1$ premiers amortissements sont nuls ($m_1, \dots, m_{n-1} = 0$) et $m_n = C_0$. Le capital restant à rembourser n'évolue donc pas avant le terme du contrat : $C_0 = C_1 = \dots = C_{n-1}$. Pour $k \in \{1, \dots, n - 1\}$, l'annuité a_k ne sert qu'à rembourser les intérêts de la période et vaut :

$$a_k = I_k = C_{k-1}\tau = C_0\tau.$$

La dernière annuité vaut quant à elle :

$$a_n = I_n + m_n = C_{n-1}\tau + m_n = C_0\tau + C_0 = C_0(1 + \tau).$$

6.3.2 Emprunts à annuités constantes

Définition 6.12. *Un emprunt indivis est dit à annuités constantes si les annuités sont constantes (i.e. $a_1 = a_2 = \dots = a_n = a$ pour un certain réel a).*

Proposition 6.10. *Pour un emprunt à annuités constantes, on a :*

1. pour tout $k \in \{1, \dots, n - 1\}$:

$$m_{k+1} = m_k(1 + \tau);$$

2. $m_1 = \frac{C_0\tau}{(1+\tau)^n - 1}$;

3. $a_k = \frac{C_0\tau}{1 - (1+\tau)^{-n}}$.

Preuve :

1. Puisque les annuités sont constantes, on a $a_{k+1} = a_k$ et on peut écrire que :

$$\begin{aligned} m_{k+1} &= a_{k+1} - I_{k+1} = a_k - C_k\tau \\ &= a_k - (C_{k-1} - m_k)\tau = a_k - C_{k-1}\tau + m_k\tau \\ &= m_k + m_k\tau = m_k(1 + \tau). \end{aligned}$$

2. Nous venons de voir que, pour tout $k \in \{1, \dots, n - 1\}$:

$$m_{k+1} = m_k(1 + \tau).$$

On en déduit que, pour tout $k \in \{1, \dots, n - 1\}$:

$$m_{k+1} = m_1(1 + \tau)^k;$$

Donc,

$$C_0 = \sum_{j=1}^n m_j = \sum_{k=0}^{n-1} m_1(1 + \tau)^k = m_1 \sum_{k=0}^{n-1} (1 + \tau)^k = m_1 \frac{(1 + \tau)^n - 1}{\tau}.$$

D'où le résultat.

3. Il suffit d'écrire que

$$a_k = a_1 = m_1 + I_1 = \frac{C_0\tau}{(1+\tau)^n - 1} + C_0\tau = \frac{C_0\tau(1+\tau)^n}{(1+\tau)^n - 1} = \frac{C_0\tau}{1 - (1+\tau)^{-n}}.$$

□

Remarque 6.9. La proposition précédente montre que m_1, \dots, m_n sont les n premiers termes de la suite géométrique de raison $(1 + \tau)$ et de premier terme $m_1 = \frac{C_0\tau}{(1+\tau)^n - 1}$. En particulier, les amortissements sont croissants.

Exercice 6.8. Dresser le tableau d'amortissement d'un emprunt à annuités constantes de 15000€ sur 61 mois au taux mensuel de 1,5%.

Indication : Commencer par déterminer la suite des amortissements de cet emprunt en utilisant la Proposition 6.10.

6.3.3 Emprunts à amortissements constants

Définition 6.13. Un emprunt indivis est dit à amortissements constants si les amortissements sont constants (i.e. $m_1 = m_2 = \dots = m_n = m$ pour un certain réel m).

Proposition 6.11. Pour un emprunt à amortissements constants, on a :

1. pour tout $k \in \{1, \dots, n\}$:

$$m_k = \frac{C_0}{n};$$

2. pour tout $k \in \{1, \dots, n-1\}$:

$$a_{k+1} = a_k - \frac{C_0\tau}{n}.$$

Preuve :

1. Puisque les amortissements sont constants, on a $m_1 = \dots = m_n = m$. Or, $C_0 = m_1 + \dots + m_n = nm$. D'où, $m = \frac{C_0}{n}$.

2. Puisque les amortissements sont constants, on a $m_{k+1} = m_k = \frac{C_0}{n}$ et on peut écrire que :

$$\begin{aligned} a_{k+1} &= I_{k+1} + m_{k+1} = C_k\tau + m_k \\ &= (C_{k-1} - m_k)\tau + m_k = C_{k-1}\tau + m_k - \frac{C_0\tau}{n} \\ &= a_k - \frac{C_0\tau}{n}. \end{aligned}$$

□

Remarque 6.10. La proposition précédente montre que a_1, \dots, a_n sont les n premiers termes de la suite arithmétique de raison $-\frac{C_0\tau}{n}$ et de premier terme $a_1 = C_0\tau \left(1 + \frac{1}{n}\right)$. En particulier, les annuités sont dégressives.

Exercice 6.9. Dresser le tableau d'amortissement d'un emprunt à amortissement constants de 600€ sur 3 mois au taux mensuel de 3%.

Indication : Commencer par déterminer la suite des amortissements de cet emprunt.

Exercice 6.10. La première annuité d'un emprunt à amortissements constants est de $a_1 = 360$ et le premier amortissement $m_1 = 300$. Le taux mensuel est de 2%. Dresser le tableau d'amortissement de cet emprunt.

Indication : Commencer par déterminer les intérêts de la première période I_1 . En déduire le capital emprunté C_0 puis la durée de l'emprunt (*i.e.* le nombre n de mois). Dresser finalement le tableau.

6.3.4 Taux annuel effectif global (TAEG)

Nous avons jusqu'ici négligé les frais annexes liés à un prêt. En général, le prêteur demande à l'emprunteur de rembourser le capital et les intérêts mais aussi de payer des frais de gestion et d'assurance. Le *taux annuel effectif global (TAEG)* permet d'exprimer le coût global d'un crédit. Le TAEG est défini par la directive 98/7/CE du Parlement européen. En France, il est utilisé comme taux effectif de référence pour les crédits à la consommation depuis le 10 juin 2002. Les directives européennes 2008/48/CE et 2014/17/UE l'ont rendu taux effectif légal pour les crédits à la consommation (à partir du 23 avril 2008) puis pour les crédits immobiliers aux particuliers (à partir du 21 mars 2016).

Le TAEG est défini par la directive 98/7/CE du Parlement européen. Dans le cadre d'un emprunt indivis d'un capital C_0 consenti en une seule fois et remboursable par des annuités a_1, \dots, a_n versées aux temps t_k exprimés en fractions d'années après la souscription du prêt, le TAEG est la solution de l'équation :

$$C_0 = \sum_{k=1}^n \frac{a_k}{(1 + \text{TAEG})^{t_k}}.$$

Remarque 6.11.

- Une année comporte soit 365 jours, soit 365,25 jours, soit 366 si elle est bissextile, 52 semaines ou 12 mois normalisés.
- L'équation définissant le TAEG ne peut pas toujours être résolue de façon exacte et on peut avoir recourt à des approximations numériques.
- Si plusieurs possibilités de prêts sont proposées pour acquérir le même capital, l'emprunteur a tout intérêt à choisir celui présentant le TAEG le plus faible.
- Lorsque l'assurance est facultative, on peut calculer les TAEG avec et hors assurance. Le *Taux annuel effectif d'assurance (TAEA)* fourni un indicateur permettant de comparer des assurances. Il est simplement défini dans le Décret n° 2014-1190 du 15 octobre 2014 comme la différence entre le TAEG avec assurance et le TAEG hors assurance.

Exemple 6.17. Un prêteur consent un emprunt de 3000€. L'emprunteur s'engage en contrepartie à régler 150€ de frais de dossiers et d'assurance puis 1600€ après six mois et un an. Le premier versement de 150€ à lieu au bout de 0 ans, le deuxième (de 1600€) au bout d'une demie année et le troisième (de 1600€ également) au bout d'un an. L'équation définissant le TAEG s'écrit donc :

$$3000 = 150 + \frac{1600}{(1 + \text{TAEG})^{\frac{1}{2}}} + \frac{1600}{1 + \text{TAEG}}.$$

En posant $x = \frac{1}{(1+\text{TAEG})^{\frac{1}{2}}}$, cette équation se réécrit :

$$3000 = 150 + 1600x + 1600x^2,$$

soit

$$1600x^2 + 1600x - 2850 = 0. \tag{6.3.1}$$

Le discriminant du polynôme du second degré $1600x^2 + 1600x - 2850$ est :

$$\Delta = 1600^2 + 4 \times 1600 \times 2850 = 20800000.$$

L'équation précédente admet donc pour solutions :

$$x_1 = \frac{-1600 - \sqrt{20800000}}{2 \times 1600} = \frac{-1600 - 400\sqrt{130}}{2 \times 1600} = \frac{-4 - \sqrt{130}}{8}$$

et

$$x_2 = \frac{-1600 + \sqrt{20800000}}{2 \times 1600} = \frac{-1600 + 400\sqrt{130}}{2 \times 1600} = \frac{-4 + \sqrt{130}}{8}.$$

En utilisant que $x = \frac{1}{(1+\text{TAEG})^{\frac{1}{2}}}$, on déduit que $(1 + \text{TAEG})^{\frac{1}{2}} = \frac{1}{x}$ puis $\text{TAEG} = \frac{1}{x^2} - 1$, où x est une des deux solutions de l'équation 6.3.1. Avec x_1 , on obtient :

$$\frac{1}{x_1^2} - 1 \simeq -0,73,$$

ce qui n'est pas possible, le TAEG étant positif. Avec x_2 , on obtient :

$$\text{TAEG} = \frac{1}{x_2^2} - 1 \simeq 0,1682 = 16,82\%.$$

Le TAEG est donc ici de 16,82%.

Exercice 6.11. Reprendre l'Exemple 6.17 en négligeant les frais de dossier et d'assurance. Comparer les TAEG avec et sans frais de dossier et assurance.

6.4 Rentabilité d'un investissement

Une entreprise peut investir dans de nouveaux équipements afin d'améliorer le volume ou l'efficacité de sa production. Dans cette section, on présente des outils permettant d'évaluer la rentabilité d'un investissement.

6.4.1 Valeur actuelle nette

Définition 6.14. On appelle valeur actuelle nette (VAN) d'un investissement la valeur actuelle des flux de trésorerie provoqués par cet investissement. Si le taux d'actualisation annuel est τ , le montant de l'investissement I et qu'il est prévu que l'investissement entraîne pendant n années des flux de trésorerie r_k en fin de k^e année, la valeur actuelle nette de l'investissement est donnée par :

$$\text{VAN} = -I + \sum_{k=1}^n \frac{r_k}{(1 + \tau)^k}.$$

Remarque 6.12. Dans la pratique, la difficulté est d'estimer le taux d'actualisation et les flux de trésorerie provoqués par l'investissement.

Proposition 6.12. *Si le taux d'actualisation annuel est τ , le montant de l'investissement I et qu'il est prévu que l'investissement entraîne pendant n années des flux de trésorerie constant égaux à r à la fin de chaque année, la valeur actuelle nette de l'investissement est donnée par :*

$$\text{VAN} = -I + r \frac{1 - (1 + \tau)^{-n}}{\tau}.$$

Preuve : Exercice. □

6.4.2 Taux de rentabilité interne

Définition 6.15. *On appelle taux de rentabilité interne (TRI) d'un investissement le taux d'actualisation annulant la valeur actuelle nette de celui-ci. Si le montant de l'investissement est I et qu'il est prévu que l'investissement entraîne pendant n années des flux de trésorerie r_k en fin de k^e année, le taux de rentabilité interne TRI de l'investissement est solution de l'équation :*

$$-I + \sum_{k=1}^n \frac{r_k}{(1 + \text{TRI})^k} = 0.$$

Remarque 6.13. On choisit généralement de réaliser un investissement lorsque son TRI est suffisamment supérieur au taux d'intérêt bancaire (il faut tenir compte des risques). Le TRI donne un critère pour décider si un investissement est susceptible d'être rentable mais n'est en aucun cas un critère objectif pour choisir entre deux investissements ; il faut avoir recourt à la VAN.

Proposition 6.13. *Si le montant de l'investissement est I et qu'il est prévu que l'investissement entraîne pendant n années des flux de trésorerie r_k en fin de k^e année, le taux de rentabilité interne TRI de l'investissement est solution de l'équation :*

$$-I + r \frac{1 - (1 + \text{TRI})^{-n}}{\text{TRI}} = 0.$$

Preuve : Exercice. □

Chapitre 7

Séries chronologiques

7.1 Premières définitions, exemples et motivations

Définition 7.1. On appelle série chronologique ou série temporelle toute suite (finie) y d'observations numériques d'une grandeur effectuées au cours de plusieurs années (mois, semaines, ...) à intervalle de temps réguliers, appelés saisons.

Notations 7.1. On note :

- y_t la valeur de la t^e observation,
- n le nombre d'années (mois, semaines, ...),
- p le nombre de saisons (périodes) dans une année.

Exemple 7.1. On a relevé les chiffres d'affaires trimestriels, exprimés en K€, d'une entreprise au cours des années 2012 à 2015. Ceux-ci sont présentés dans le tableau suivant.

Année \ Trimestre	1	2	3	4
	2012	20	25	50
2013	35	30	65	105
2014	40	34	75	135
2015	50	37	80	170

Ici, le nombre d'années est $n = 4$, le nombre de périodes par année est $p = 4$ et on a $y_1 = 20, y_2 = 25, y_3 = 50, y_4 = 70, y_5 = 35, \dots, y_{16} = 170$.

Exemple 7.2. On a relevé la consommation mensuelle en eau, exprimée en Hm^3 , d'un producteur de melons au cours des années 2013 à 2015. Les relevés sont présentés dans le tableau suivant.

	Janv.	Fév.	Mars	Avr.	Mai	Juin	Juil.	Août	Sept.	Oct.	Nov.	Déc.
2013	1	1,5	3	5	10	20	45	50	30	2	1	0,5
2014	3,5	3	5,5	9	11	24	49	50	31	4	4	3,5
2015	7	6	8	9	15	25	52	55	37	7	5	6

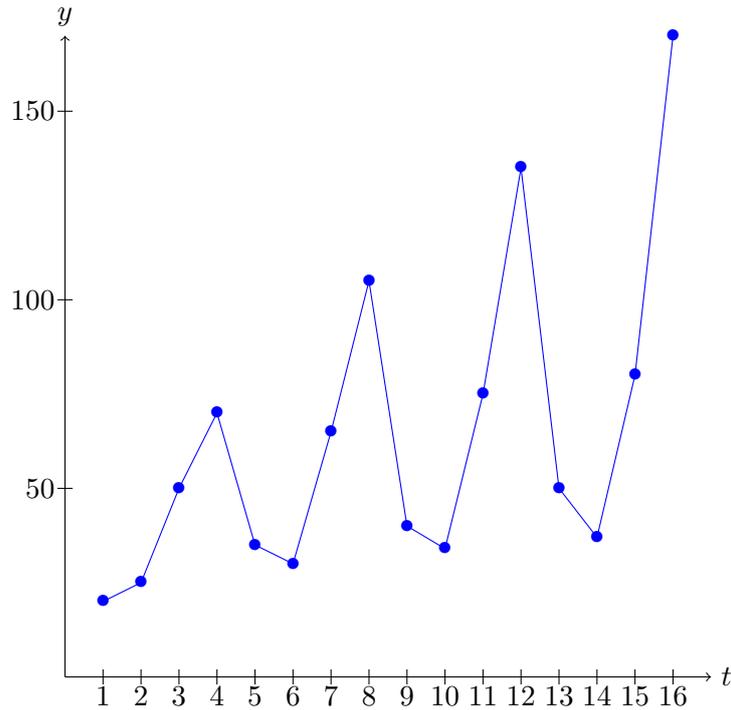


FIGURE 7.1 – Chiffres d'affaires y_t (en K€) de la t^e période (Exemple 7.1).

Ici, le nombre d'années est $n = 3$, le nombre de périodes par année est $p = 12$ et on a $y_1 = 1, y_2 = 1,5, \dots, y_{12} = 0,5, y_{13} = 3,5, \dots, y_{36} = 6$.

7.1.1 Motivations et objectifs

Pour toute série chronologique, trois questions essentielles se posent :

- Une *tendance générale* (« *trend* ») se dégage-t-elle de la série ? Quelle est l'évolution de la série en temps long ?
- Peut-on détecter des phénomènes saisonniers ? L'évolution de la série présente-t-elle des caractéristique se répétant d'année en année ?
- Peut-on prévoir des valeurs futures de la série ?

Dans les deux exemples considérés ci-dessus, on observe que les séries ont tendance à augmenter d'année en année (voir Figures 7.1 et 7.2). Une tendance générale (un « *trend* ») se dégage donc dans les deux cas. On peut également observer que l'allure générale des variation au sein d'une année est semblable d'une année à l'autre. Des effets saisonniers se dégagent. Par exemple, on peut penser que la consommation d'eau du producteur de melon de l'Exemple 7.2 est liée au variations climatiques et au périodes de production qui se répètent assez régulièrement d'année en année. Bien entendu, une troisième composante doit-être prise en compte : la composante *aléatoire*, que l'on ne peut pas prévoir. Elle peut, par exemple, prendre en compte les aléas du climat d'une année.

Pour répondre aux questions posées ci-dessus, nous fournirons des méthodes permettant :

- de décomposer une série chronologique en composantes *générale*, *saisonnière* et *aléatoire* afin de l'analyser (voir Section 7.2) ;

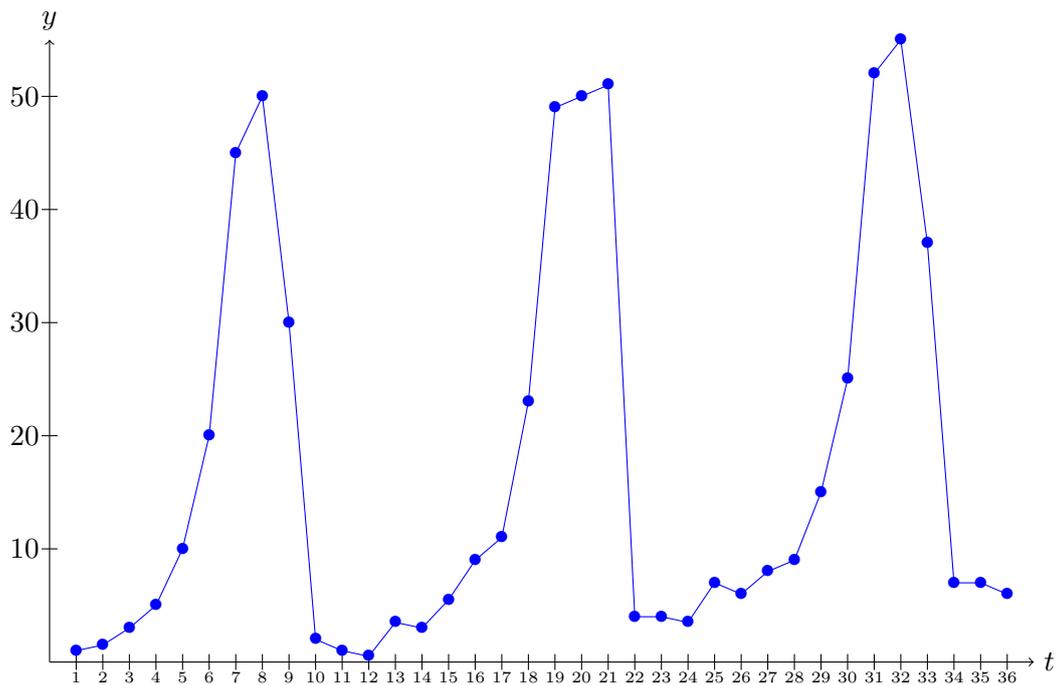


FIGURE 7.2 – Consommation d’eau y_t (en Hm^3) au cours du t^{e} mois (Exemple 7.2).

- d’estimer les composantes de la série chronologique (voir Section 7.3);
- prédire les valeurs futures de la série (voir Section 7.4).

7.2 Analyse d’une série chronologique

7.2.1 Décomposition d’une série chronologique

On veut écrire toute valeur y_t d’une série chronologique sous la forme :

$$y_t = f(g_t, s_t, a_t),$$

où f est une certaine fonction et :

- g_t désigne la composante « tendance générale » de la série au temps t ,
- s_t désigne la composante « saisonnière » de la série au temps t ,
- a_t désigne la composante « aléatoire » de la série au temps t .

Différentes fonctions f conduisent à des modèles différents. Dans la suite, on s’intéressera aux modèles *additif* (voir Sous-section 7.2.2) et *multiplicatif* (voir Sous-section 7.2.3).

Exemple 7.1 (suite). Reprenons l’exemple des chiffres d’affaires (Exemple 7.1).

Graphiquement, on constate que les chiffres d’affaires ont tendance à augmenter d’année en année. Rigoureusement, on peut justifier ce fait en faisant une régression linéaire. On constate que la tendance générale à une faible courbure, ce qui rend légitimes les calculs que nous ferons dans la suite.

Graphiquement, on voit que le chiffre d’affaire présente des « pics positifs » au 4^e trimestre de chaque année et des « pics négatifs » lors des deux premiers trimestres de chaque

année. On cherchera, dans la suite, à décrire l'impact d'une saison (ici un trimestre) sur la série chronologique. On supposera que la composante saisonnière est périodique de période p ($s_{k+p} = s_k$, pour tout k) et a une influence nulle sur une année.

La composante aléatoire est plus délicate à visualiser. Elle correspond aux irrégularités des cycles de la série. Nous supposons par la suite que celle-ci est négligeable.

7.2.2 Modèle additif

Définition 7.2. On parle de modèle additif lorsque la série chronologique $\mathbf{y} = y_t$ se décompose sous la forme :

$$y_t = g_t + s_t + a_t,$$

où g_t désigne la composante « tendance générale », s_t désigne la composante saisonnière et a_t désigne la composante aléatoire de la série au temps t .

Remarque 7.1. Rappelons que la composante saisonnière est supposée p -périodique (*i.e.* $s_{k+p} = s_k$, pour tout k) et d'influence nulle sur une année (*i.e.* $s_1 + s_2 + \dots + s_p = 0$ pour le modèle additif). La composante aléatoire est supposée négligeable (*i.e.* $a_t \simeq 0$, pour tout t , pour le modèle additif).

Critère 7.1. Pour savoir si le modèle additif est adapté, on trace les lignes polygonales passant par les pics « positifs » d'une part et « négatifs » d'autre part. Si celles-ci sont proches de droites et si la largeur de la bande délimitée par celles-ci est essentiellement constante, on choisit d'utiliser le modèle additif.

Exemple 7.2 (suite). Le modèle additif est adapté pour l'Exemple 7.2 (voir Figure 7.3).

7.2.3 Modèle multiplicatif

Définition 7.3. On parle de modèle multiplicatif lorsque la série chronologique $\mathbf{y} = y_t$ se décompose sous la forme :

$$y_t = g_t \times s_t \times a_t,$$

où g_t désigne la composante « tendance générale », s_t désigne la composante saisonnière et a_t désigne la composante aléatoire de la série au temps t .

Remarque 7.2. Rappelons que la composante saisonnière est supposée p -périodique (*i.e.* $s_{k+p} = s_k$, pour tout k) et d'influence nulle sur une année (*i.e.* $s_1 \times s_2 \times \dots \times s_p = 1$ pour le modèle multiplicatif). La composante aléatoire est supposée négligeable (*i.e.* $a_t \simeq 1$, pour tout t , pour le modèle multiplicatif).

Critère 7.2. Pour savoir si le modèle multiplicatif est adapté, on trace les lignes polygonales passant par les pics « positifs » d'une part et « négatifs » d'autre part. Si celles-ci sont proches de droites et si la largeur de la bande délimitée par celles-ci est clairement croissante ou décroissante, on choisit d'utiliser le modèle multiplicatif.

Exemple 7.1 (suite). Le modèle multiplicatif est adapté pour l'Exemple 7.1 (voir Figure 7.4).

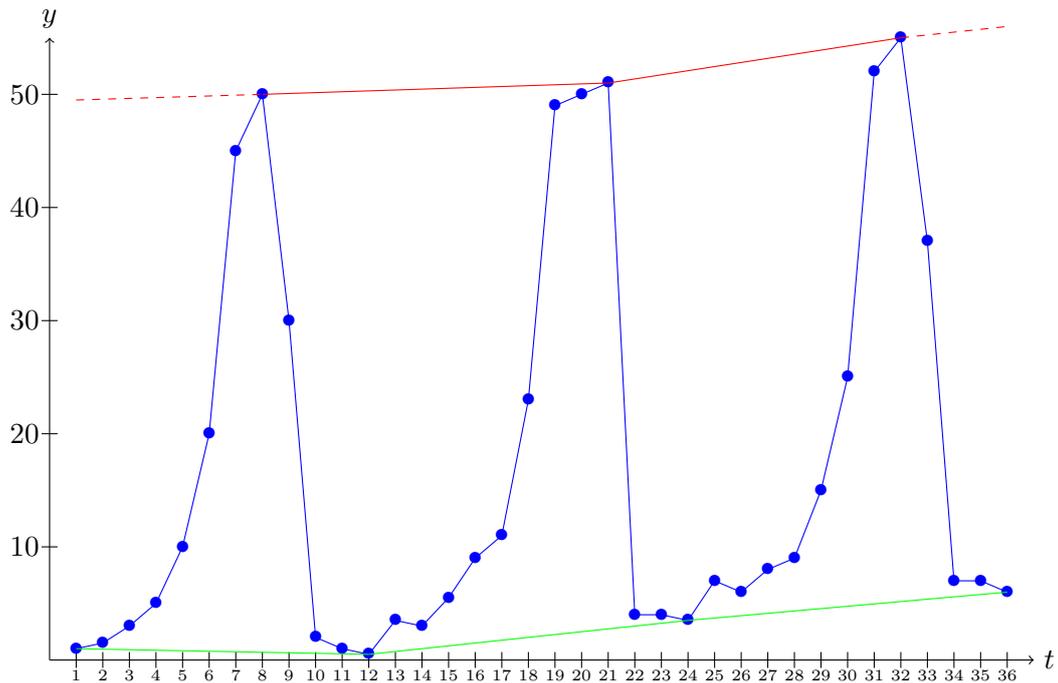


FIGURE 7.3 – Consommation d’eau y_t (en Hm^3) au cours du t^{e} mois de l’Exemple 7.2 (bleu) et les lignes polygonales passant par les pics « positifs » (rouge) et « négatifs » (vert).

7.2.4 Quelques manipulations avec un tableur

Supposons que les données constituant une série chronologique soient enregistrées dans un fichier dans un format lisible par un tableur tel que Excel, LibreOffice ou Numbers. Si ce n’est pas déjà fait, il est pratique de numéroter les différentes périodes. On peut, ensuite, repérer facilement celles correspondant aux pics « positifs » et « négatifs » et éventuellement reporter leurs numéros et les valeurs correspondantes. Pour obtenir un graphique permettant de décider si un modèle additif ou multiplicatif est adapté à la situation, on peut :

1. sélectionner les plages de données contenant les numéros de périodes t et les valeurs de la série chronologique y_t ;
2. insérer un graphique de type « X/Y dispersion » ;
3. ajouter sur ce graphique des plages/séries de données correspondants aux pics « positifs » et « négatifs » (en spécifiant les plages de données pour X et Y, nom de la série, ... ; si la fenêtre de création de graphique a été fermée, il est possible de la réouvrir grâce à un « clic droit » puis la sélection de « Plages de données... ») ;
4. ajouter les titres, sous-titres, légendes, ... nécessaires à une bonne compréhension du graphique et pour le choix du modèle (pensez au clic droit pour accéder aux menus d’édition).

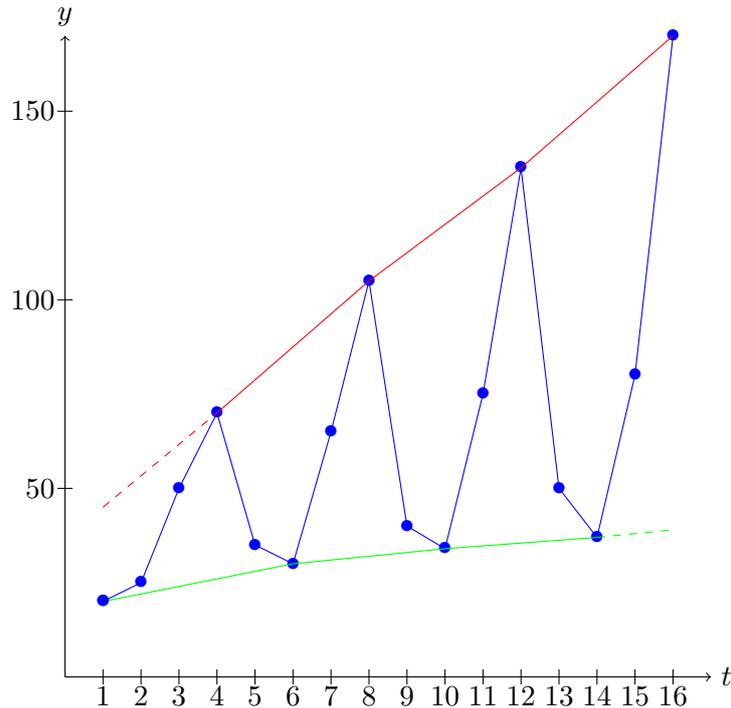


FIGURE 7.4 – Chiffres d'affaires y_t (en K€) du t^e trimestre de l'Exemple 7.1 (bleu) et les lignes polygonales passant par les pics « positifs » (rouge) et « négatifs » (vert).

7.3 Estimation des composantes d'une série chronologique

7.3.1 Estimation de la tendance générale

Une première méthode possible pour estimer la tendance générale g_t est d'effectuer une régression de y_t en fonction de t (voir Chapitre 3). Celle-ci conduit à des calculs assez simples mais n'est pas toujours adaptée à cause des effets saisonniers (en particulier quand ceux-ci sont importants). Nous détaillons ici une autre méthode basée sur l'utilisation de *moyennes mobiles* pour estimer la tendance générale.

Définition 7.4. Soit \mathbf{y} une série chronologique portant sur n années décomposées en p saisons et soit $k \in \{2, \dots, np\}$. On appelle moyenne mobile (ou glissante) d'ordre k au temps $t \in \{1 + \lfloor \frac{k}{2} \rfloor, \dots, np - \lfloor \frac{k}{2} \rfloor\}$ la quantité :

$$M_t^{(k)} = \begin{cases} \frac{1}{k} \left(y_{t-\frac{k-1}{2}} + \dots + y_t + \dots + y_{t+\frac{k-1}{2}} \right) & \text{si } k \text{ est impair} \\ \frac{1}{k} \left(\frac{y_{t-\frac{k}{2}}}{2} + y_{t-\frac{k}{2}+1} + \dots + y_t + \dots + y_{t+\frac{k}{2}-1} + \frac{y_{t+\frac{k}{2}}}{2} \right) & \text{si } k \text{ est pair} \end{cases} .$$

Exemple 7.3. Les moyennes mobiles d'ordre 2, 3, 4 et 5 sont données par :

$$M_t^{(2)} = \frac{1}{2} \left(\frac{y_{t-1}}{2} + y_t + \frac{y_{t+1}}{2} \right),$$

$$M_t^{(3)} = \frac{1}{3} (y_{t-1} + y_t + y_{t+1}),$$

$$M_t^{(4)} = \frac{1}{4} \left(\frac{y_{t-2}}{2} + y_{t-1} + y_t + y_{t+1} + \frac{y_{t+2}}{2} \right)$$

et

$$M_t^{(5)} = \frac{1}{5} (y_{t-2} + y_{t-1} + y_t + y_{t+1} + y_{t+2}).$$

Remarque 7.3.

1. Les moyennes mobiles (ou glissantes) doivent leur nom au fait que le sous-ensemble des données utilisées pour les calculer se déplace (« glisse ») dans le jeu complet de données lorsque t varie.
2. L'utilisation des moyennes mobile permet de lisser la courbe et « supprimer » la composante aléatoire par moyennisation. Ainsi, elles permettent une meilleure visualisation de la tendance générale. Leur défaut est la perte d'information sur les premières et dernières valeurs du temps.
3. Lorsque k est un multiple de p , la composante saisonnière n'influe pas sur la moyenne mobile d'ordre k .
4. Il est nécessaire de choisir convenablement l'ordre k des moyennes mobiles utilisées. En général, on choisit $k = p$ (le nombre de période dans l'année) de sorte à ce que la composante saisonnière n'influe pas sur cette estimation tout en gardant le plus d'information possible sur l'évolution de la tendance. C'est ce que nous ferons pour l'Exemple 7.1 ($k = p = 4$). Dans l'Exemple 7.2, on choisira un ordre inférieur $k = 7 < p = 12$ pour illustrer les conséquences d'un mauvais choix de l'ordre des moyennes mobiles. À titre d'exercice, on pourra reprendre cet exemple en utilisant un « bon » choix de l'ordre des moyennes mobiles.

Proposition 7.1. *Soit y une série chronologique dont la tendance générale est notée g_t .*

Alors, un estimateur de g_t est donné par :

$$\hat{g}_t = M_t^{(k)}.$$

Autrement dit, on peut considérer que :

$$g_t \simeq \hat{g}_t = M_t^{(k)}.$$

Remarque 7.4. L'estimation de la tendance générale ne dépend pas du fait que l'on utilise un modèle additif ou multiplicatif.

Exemple 7.1 (suite). L'estimation de la tendance des chiffres d'affaires trimestriels de l'Exemple 7.1 par des moyennes mobiles d'ordre 4 est donnée dans le tableau suivant et représentée dans la Figure 7.5.

Trimestre Année	1	2	3	4
2012			43,125	45,625
2013	48,125	54,375	59,375	60,5
2014	62,25	67,25	72,25	73,875
2015	74,875	79,875		

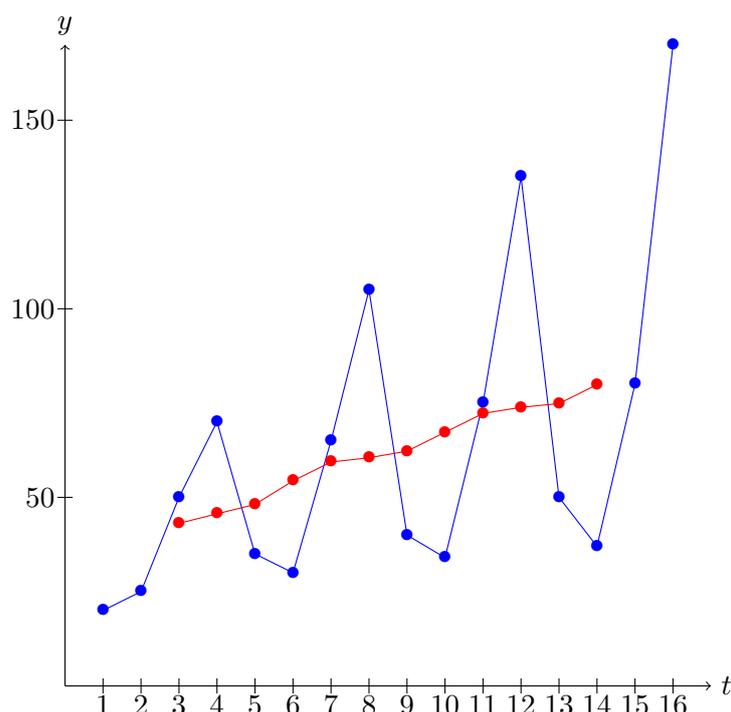


FIGURE 7.5 – Chiffres d'affaires y_t (en K€) du t^e trimestre (bleu, Exemple 7.1) et estimation de la tendance g_t par $\hat{g}_t = M_t^{(4)}$ (rouge).

Exemple 7.2 (suite). Les valeurs approchées à une décimale de l'estimation de la tendance la consommation mensuelle d'eau de l'Exemple 7.2 par des moyennes mobiles d'ordre 7 sont données dans le tableau suivant et représentées dans la Figure 7.6.

	Janv.	Fév.	Mars	Avr.	Mai	Juin	Juil.	Août	Sept.	Oct.	Nov.	Déc.
2013				12,2	19,2	23,3	23,1	22,6	21,2	18,9	12,9	6,5
2014	3,5	4,8	8,1	15	21,6	25,6	25,4	24,7	23,6	21,2	15,1	9,1
2015	5,9	7,5	10,5	17,4	24,3	28,7	28,6	28	26,7			

7.3.2 Estimation de la composante saisonnière

L'estimation de la composante saisonnière s_t dépend du fait que l'on utilise un modèle additif ou multiplicatif.

Méthode pour l'estimation de la composante saisonnière dans le modèle additif

Rappelons que l'on veut écrire les éléments de la série sous la forme :

$$y_t = g_t + s_t + a_t$$

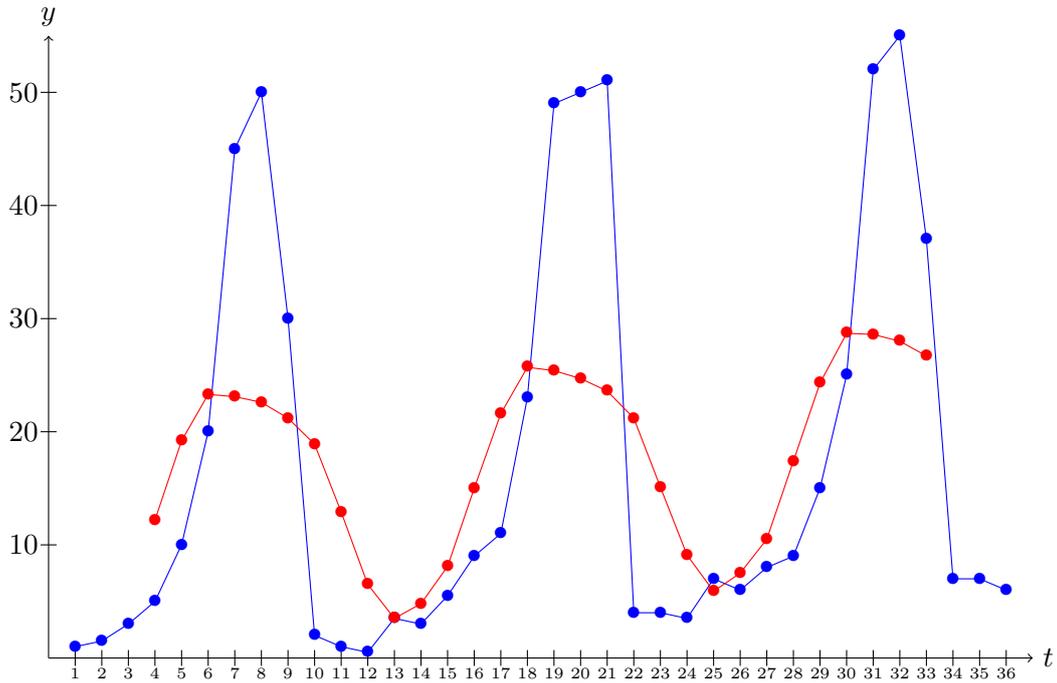


FIGURE 7.6 – Consommation d'eau y_t (en Hm^3) au cours du t^{e} mois (bleu, Exemple 7.1) et estimation de la tendance g_t par $\hat{g}_t = M_t^{(7)}$ (rouge).

où g_t est la tendance générale, s_t la composante saisonnière et a_t la composante aléatoire de la série au temps t . Rappelons aussi que l'on a supposé que la composante aléatoire est négligeable ($a_t \simeq 0$) et que l'on sait déjà estimer la tendance générale à l'aide de moyennes mobiles d'ordre k par $\hat{g}_t = M_t^{(k)}$. On cherche donc un estimateur \hat{s}_t de s_t de façon à pouvoir écrire :

$$y_t \simeq \hat{g}_t + \hat{s}_t.$$

Pour cela, on effectue les étapes suivantes :

1. pour $t \in \{1 + \lfloor \frac{k}{2} \rfloor, \dots, np - \lfloor \frac{k}{2} \rfloor\}$, on calcule les écarts entre la valeur de la série et l'estimation de la tendance au temps t : $e_t = y_t - \hat{g}_t$;
2. pour $t \in \{1, \dots, p\}$, on calcule la moyenne **arithmétique** \bar{e}_t des écarts correspondant à la même période de l'année, *i.e.* des e_{t+kp} ;
3. on calcule la moyenne **arithmétique** \bar{e} des \bar{e}_t : $\bar{e} = \frac{\bar{e}_1 + \dots + \bar{e}_p}{p}$;
4. pour $t \in \{1, \dots, p\}$, on calcule les *coefficients saisonniers normalisés* $\hat{s}_t = \bar{e}_t - \bar{e}$.

Remarque 7.5.

1. Il suffit de connaître les \hat{s}_t pour $t \in \{1, \dots, p\}$ puisque l'on a supposé la composante saisonnière p -périodique.
2. Dans le point 2., le nombre de termes dans les calculs des moyennes peut être différent d'une période à l'autre. C'est le cas lorsque $k < p$: on aura alors $n - 1$ termes dans les moyennes correspondant aux $\lfloor \frac{k}{2} \rfloor$ premières et $\lfloor \frac{k}{2} \rfloor$ dernières périodes et n termes dans les autres.

3. Les points 3. et 4. permettent d'assurer que la participation de la composante saisonnière est d'influence nulle sur une année.

Proposition 7.2. Dans le cadre d'un modèle additif, une estimation de la composante saisonnière est donnée par \hat{s}_t défini ci-dessus.

Exemple 7.2 (suite). Reprenons l'Exemple 7.2. Le tableau suivant donne les valeurs approchées à une décimale près des écarts $e_t = y_t - \hat{g}_t$ obtenus pour chaque période $t \in \{4, \dots, 33\}$, celles des moyennes \bar{e}_t des écarts correspondant à chacun des mois d'une année :

$$\bar{e}_1 = \frac{e_{13} + e_{25}}{2}, \dots, \bar{e}_3 = \frac{e_{15} + e_{27}}{2}, \bar{e}_4 = \frac{e_4 + e_{16} + e_{28}}{3}, \dots, \bar{e}_9 = \frac{e_9 + e_{21} + e_{33}}{3},$$

$$\bar{e}_{10} = \frac{e_{10} + e_{22}}{2}, \dots, \bar{e}_{12} = \frac{e_{12} + e_{24}}{2},$$

ainsi que l'estimation de la composante saisonnière $\hat{s}_t = \bar{e}_t - \bar{e}$ où $\bar{e} = \frac{\bar{e}_1 + \bar{e}_2 + \dots + \bar{e}_{12}}{12} \simeq 0,05$.

	Janv.	Fév.	Mars	Avr.	Mai	Juin	Juil.	Août	Sept.	Oct.	Nov.	Déc.
2013				-7,2	-9,2	-3,3	21,9	27,4	8,8	-16,9	-11,9	-6
2014	0	-1,8	-2,6	-6	-10,6	-1,6	23,6	25,3	7,4	-17,2	-11,1	-5,6
2015	1,1	-1,5	-2,5	-8,4	-9,3	-3,7	23,4	27	10,3			
\bar{e}_t	0,5	-1,6	-2,5	-7,2	-9,7	-2,9	23	26,6	8,8	-17	-11,5	-5,8
\hat{s}_t	0,5	-1,7	-2,6	-7,2	-9,8	-2,9	22,9	26,5	8,8	-17,1	-11,5	-5,8

Méthode pour l'estimation de la composante saisonnière dans le modèle multiplicatif

Rappelons que l'on veut écrire les éléments de la série sous la forme :

$$y_t = g_t \times s_t \times a_t$$

où g_t est la tendance générale, s_t la composante saisonnière et a_t la composante aléatoire de la série au temps t . Rappelons aussi que l'on a supposé que la composante aléatoire est négligeable ($a_t \simeq 1$) et que l'on sait déjà estimer la tendance générale à l'aide de moyennes mobiles d'ordre k par $\hat{g}_t = M_t^{(k)}$. On cherche donc un estimateur \hat{s}_t de s_t de façon à pouvoir écrire :

$$y_t \simeq \hat{g}_t \times \hat{s}_t.$$

Pour cela, on effectue les étapes suivantes :

1. pour $t \in \{1 + \lfloor \frac{k}{2} \rfloor, \dots, np - \lfloor \frac{k}{2} \rfloor\}$, on calcule les ratios entre la valeur de la série et l'estimation de la tendance au temps t : $r_t = \frac{y_t}{\hat{g}_t}$;
2. pour $t \in \{1, \dots, p\}$, on calcule la moyenne **géométrique** \bar{r}_t des ratios correspondant à la même période de l'année, *i.e.* des r_{t+kp} ;
3. on calcule la moyenne **géométrique** \bar{r} des \bar{r}_t : $\bar{r} = \sqrt[p]{\bar{r}_1 \times \dots \times \bar{r}_p}$;

4. pour $t \in \{1, \dots, p\}$, on calcule les *coefficients saisonniers normalisés* $\hat{s}_t = \frac{\bar{r}_t}{\bar{r}}$.

Remarque 7.6.

1. Il suffit de connaître les \hat{s}_t pour $t \in \{1, \dots, p\}$ puisque l'on a supposé la composante saisonnière p -périodique.
2. Dans le point 2., le nombre de termes dans les calculs des moyennes peut être différent d'une période à l'autre. C'est le cas lorsque $k < p$: on aura alors $n - 1$ termes dans les moyennes correspondant aux $\lfloor \frac{k}{2} \rfloor$ premières et $\lfloor \frac{k}{2} \rfloor$ dernières périodes et n termes dans les autres.
3. Les points 3. et 4. permettent d'assurer que la participation de la composante saisonnière est d'influence nulle sur une année.

Proposition 7.3. *Dans le cadre d'un modèle multiplicatif, une estimation de la composante saisonnière est donnée par \hat{s}_t défini ci-dessus.*

Exemple 7.1 (suite). Reprenons l'Exemple 7.1. Le tableau suivant donne les valeurs approchées à deux décimales près des ratios $r_t = \frac{y_t}{\hat{g}_t}$ obtenus pour chaque période $t \in \{3, 4, \dots, 14\}$, celles des moyennes \bar{r}_t des ratios correspondant à chacun des trimestres d'une année :

$$\bar{r}_1 = \sqrt[3]{r_5 r_9 r_{13}}, \quad \bar{r}_2 = \sqrt[3]{r_6 r_{10} r_{14}}, \quad \bar{r}_3 = \sqrt[3]{r_3 r_7 r_{11}}, \quad \text{et} \quad \bar{r}_4 = \sqrt[3]{r_4 r_8 r_{12}},$$

ainsi que l'estimation de la composante saisonnière $\hat{s}_t = \frac{\bar{r}_t}{\bar{r}}$ où $\bar{r} = \sqrt[4]{\bar{r}_1 \bar{r}_2 \bar{r}_3 \bar{r}_4} \simeq 0,89$.

Année \ Trimestre	1	2	3	4
2012			1,16	1,53
2013	0,73	0,55	1,09	1,74
2014	0,64	0,51	1,04	1,83
2015	0,67	0,46		
\bar{r}_t	0,68	0,51	1,10	1,69
\hat{s}_t	0,76	0,57	1,23	1,90

7.4 Prédiction

Étant donnée une série chronologique \mathbf{y} portant sur n années découpées en p saisons, on souhaite prédire les valeurs futures de la série. Pour cela, on effectue la démarche suivante.

1. On estime la tendance à l'aide de moyennes mobiles $\hat{g}_t = M_t^{(k)}$, $t \in \{1, \dots, np\}$ (voir Sous-section 7.3.1).
2. On estime la composante saisonnière par \hat{s}_t comme détaillé dans la Sous-section 7.3.2.
3. On effectue une régression linéaire sur les \hat{g}_t pour prédire leurs valeurs futures (voir Chapitre 3).
4. On en déduit des prédictions des valeurs futures de la série en utilisant que :

- $y_t \simeq \hat{g}_t + \hat{s}_t$ si l'on utilise un modèle additif,
- $y_t \simeq \hat{g}_t \times \hat{s}_t$ si l'on utilise un modèle multiplicatif.

Exemple 7.1 (suite). Reprenons l'Exemple 7.1. En utilisant les résultats déjà obtenus dans les sections précédentes, on obtient que la droite de régression de \hat{g}_t en t a pour équation :

$$y = \hat{g}_t = 3,37t + 33,15.$$

On en déduit les prédictions suivantes pour l'année 2016 :

$$y_{17} \simeq (3,37 \times 17 + 33,15)\hat{s}_{17} \simeq (3,37 \times 17 + 33,15) \times 0,76 \simeq 68,73,$$

$$y_{18} \simeq (3,37 \times 18 + 33,15)\hat{s}_{18} \simeq (3,37 \times 17 + 33,15) \times 0,57 \simeq 53,47,$$

$$y_{19} \simeq (3,37 \times 19 + 33,15)\hat{s}_{19} \simeq (3,37 \times 17 + 33,15) \times 1,22 \simeq 118,56,$$

$$y_{20} \simeq (3,37 \times 20 + 33,15)\hat{s}_{20} \simeq (3,37 \times 17 + 33,15) \times 1,89 \simeq 190,04.$$

Celles-ci sont représentées dans la Figure 7.7.

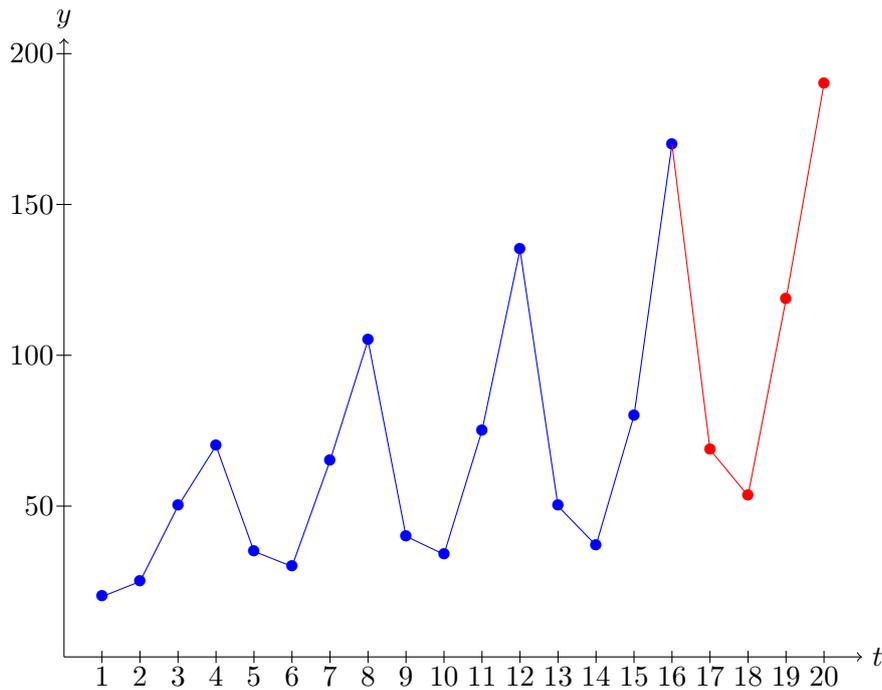


FIGURE 7.7 – Chiffres d'affaires y_t (en K€) de la t^e période de l'Exemple 7.1 (bleu) et prédictions pour l'année 2016 (rouge).

Exemple 7.2 (suite). Reprenons l'Exemple 7.2. En utilisant les résultats déjà obtenus dans les sections précédentes, on obtient que la droite de régression de \hat{g}_t en t a pour équation :

$$y = \hat{g}_t = 0,04t + 17,1.$$

On en déduit les prédictions pour l'année 2016, en calculant, pour $t \in \{37, \dots, 48\}$,

$$y_t \simeq 0,04t + 17,1 + \hat{\varepsilon}_t.$$

Les prédictions sont les suivantes et sont représentées dans la Figure 7.7 :

$$\begin{aligned} y_{37} &\simeq 19,08, & y_{38} &\simeq 16,92, & y_{39} &\simeq 16,06, & y_{40} &\simeq 11,4, & y_{41} &\simeq 8,94, & y_{42} &\simeq 15,88, \\ y_{43} &\simeq 41,72, & y_{44} &\simeq 45,36, & y_{45} &\simeq 27,65, & y_{46} &\simeq 1,84, & y_{47} &\simeq 7,48, & y_{48} &\simeq 13,19. \end{aligned}$$

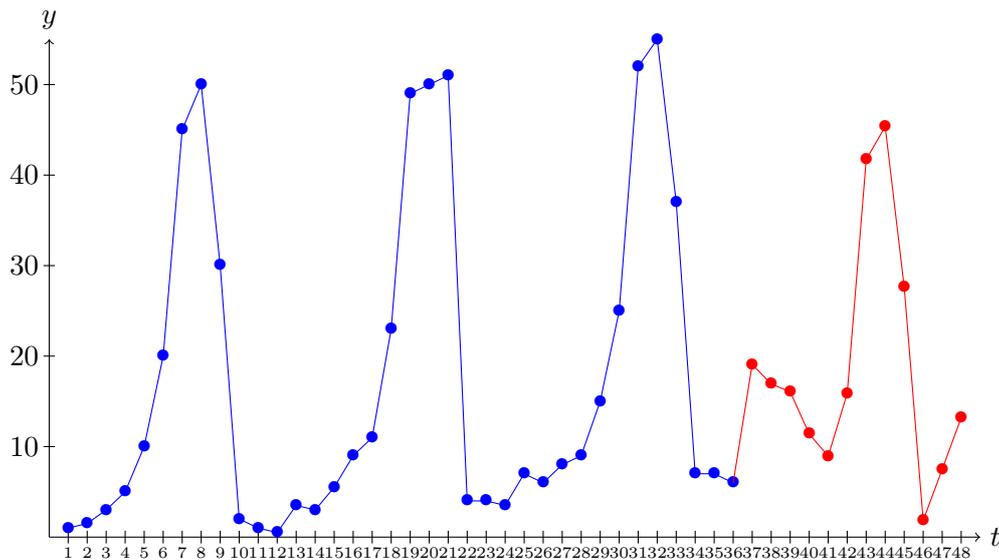


FIGURE 7.8 – Consommation d'eau y_t (en Hm^3) au cours du t^{e} mois de l'Exemple 7.2 (bleu) et prévisions pour l'année 2016 (rouge).

Remarque 7.7. Les prévisions obtenues dans l'Exemple 7.2 sont moins satisfaisantes que celles obtenues dans l'Exemple 7.2. On peut observer sur la Figure 7.6 que l'estimation de la tendance par les moyennes mobiles d'ordre $k = 7$ continue à osciller et n'a donc pas une faible courbure. Pour pallier à ce problème et obtenir de meilleures prévisions, on peut augmenter l'ordre des moyennes mobiles considérées. Il y a, en fait, un compromis à faire pour obtenir une tendance proche d'une droite sans perdre trop d'information en lissant la courbe. Si l'ordre des moyennes mobiles est trop faible, la tendance aura une allure chaotique mais, s'il est trop élevé, on perd trop d'information pour obtenir des résultats convenables. En général, le choix de l'ordre des moyennes mobiles comme le nombre de saisons (*i.e.* $k = p$) est approprié.

7.5 Quelques manipulations avec un tableur

Avec Excel ou Libre Office, pour illustrer les résultats, on peut :

1. sélectionner les plages de données contenant les numéros de périodes t et les valeurs de la série chronologique y_t ;

2. insérer un graphique de type « X/Y dispersion » ;
3. ajouter sur ce graphique des plages/séries de données correspondants à l'estimation de la tendance générale par les moyennes mobiles et aux prédictions (en spécifiant les plages de données pour X et Y - attention au recollement entre les données connues et les prévisions -, nom de la série, ... ; si la fenêtre de création de graphique a été fermée, « clic droit>Plages de données... ») ;
4. utiliser un clic droit sur les points correspondants aux moyennes mobiles pour ajouter la droite de tendance et demander d'extrapoler sur autant de périodes que nécessaire pour les prévisions ;
5. ajouter les titres, sous-titres, légendes, ... nécessaires à une bonne compréhension du graphique et pour le choix du modèle (pensez au clic droit pour accéder aux menus d'édition).

Chapitre 8

Ensembles et dénombrement

Dans ce chapitre, on présente des éléments de la théorie des ensembles et des méthodes de dénombrement nécessaires au calcul des probabilités.

8.1 Éléments de la théorie des ensembles

8.1.1 Ensembles, sous-ensembles

Définition 8.1. 1. On appelle ensemble toute collection E d'objets satisfaisant certaines propriétés. Chacun de ces objets est appelé un élément de l'ensemble E .

2. Si x est un élément de l'ensemble E , on dit que x appartient à E et on note $x \in E$; sinon, on note $x \notin E$.

3. L'ensemble ne contenant aucun élément est appelé ensemble vide et est noté \emptyset .

Exemple 8.1.

1. Un ensemble peut être décrit soit par énumération de ses éléments, soit par compréhension des propriétés vérifiées par ses éléments. Par exemple, l'ensemble des nombres premiers inférieurs à 20 est :

$$\begin{aligned} E &= \{n \in \mathbf{N} : n \text{ est premier et } n \leq 20\} \\ &= \{2; 3; 5; 7; 11; 13; 17; 19\}. \end{aligned}$$

On peut observer que $2 \in E$ et $4 \notin E$.

2. Parmi les ensembles classiques, on peut citer :

- \mathbf{N} : l'ensemble des entiers naturels,
- \mathbf{Z} : l'ensemble des entiers relatifs,
- \mathbf{Q} : l'ensemble des nombres rationnels,
- \mathbf{R} : l'ensemble des nombres réels,
- \mathbf{C} : l'ensemble des nombres complexes.

Remarque 8.1. Il n'existe pas d'ensemble de tous les ensembles.

Définition 8.2. Soit E un ensemble ayant un nombre fini d'éléments.

On appelle cardinal de E , le nombre d'éléments de E . Celui-ci est noté $|E|$, $\#(E)$ ou $\text{Card}(E)$.

Exemple 8.2. Si $E = \{n \in \mathbf{N} : n \text{ est pair et } n \leq 10\}$, $|E| = 6$.

Remarque 8.2. On a $|\emptyset| = 0$.

Définition 8.3. Soient A et B deux ensembles.

On dit que A est inclus dans B si pour tout $x \in A$, on a $x \in B$. On note alors $A \subset B$ et A est un sous-ensemble de B .

Exemple 8.3. On a :

$$\mathbf{N} \subset \mathbf{Z} \subset \mathbf{Q} \subset \mathbf{R},$$

$$\{0; 1; \sqrt{2}\} \subset \mathbf{R},$$

mais

$$\{0; 1; \sqrt{2}\} \not\subset \mathbf{Q}.$$

Proposition 8.1. Soient A, B et C trois ensembles.

On a :

1. [réflexivité] $A \subset A$;
2. [transitivité] si $A \subset B$ et $B \subset C$, alors $A \subset C$;
3. [antisymétrie] si $A \subset B$ et $B \subset A$, alors $A = B$.

Preuve :

1. Évident !
2. Soit $x \in A$. Puisque $A \subset B$, on a $x \in B$ et puisque $B \subset C$, on a $x \in C$. Ainsi, $A \subset C$.
3. Raisonnons par contraposition. Si $A \neq B$, on peut trouver un élément $x \in A$ qui n'appartient pas à B ce qui contredit $A \subset B$ ou un élément $x' \in B$ qui n'appartient pas à A ce qui contredit $B \subset A$. D'où le résultat.

□

8.1.2 Opérations sur les ensembles

Réunion et intersection

Définition 8.4. Soient A et B deux sous-ensembles d'un ensemble E .

1. On appelle union de A et de B l'ensemble :

$$A \cup B = \{x \in E : x \in A \text{ ou } x \in B\}.$$

2. On appelle intersection de A et de B l'ensemble :

$$A \cap B = \{x \in E : x \in A \text{ et } x \in B\}.$$

Remarque 8.3.

1. Dans la définition de la réunion de deux ensembles, le « ou » utilisé est un ou *inclusif*, c'est-à-dire que $A \cup B$ contient tous les éléments contenus dans A , dans B et dans les deux.
2. Si $A_1, A_2, \dots, A_n \subset E$, on note :

$$\bigcup_{k=1}^n A_k = A_1 \cup A_2 \cup \dots \cup A_n$$

$$\bigcap_{k=1}^n A_k = A_1 \cap A_2 \cap \dots \cap A_n.$$

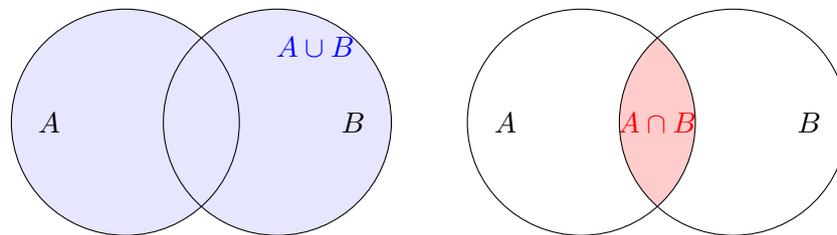


FIGURE 8.1 – Représentation de l'union (bleu) et l'intersection (rouge) de A et de B par des diagrammes de Venn (en « patatoïdes »).

Exemple 8.4. Soient les deux sous-ensembles de \mathbf{N} :

$$A = \{1; 3; 5; 7; 9\} \quad \text{et} \quad B = \{4; 5; 6\}.$$

On a :

$$A \cup B = \{1; 3; 4; 5; 6; 7; 9\} \quad \text{et} \quad A \cap B = \{5\}.$$

Définition 8.5. Soient $A, B, A_1, \dots, A_n \subset E$.

1. On dit que A et B sont *disjoints* si $A \cap B = \emptyset$.
2. On dit que A_1, \dots, A_n forment une *partition* de E s'ils sont deux à deux *disjoints* et

$$\bigcup_{k=1}^n A_k = E.$$

Proposition 8.2. Soient $A, B, C \subset E$.

On a :

1.

$$A \cap B \subset A, \quad A \cap B \subset B, \quad A \subset A \cup B \quad \text{et} \quad B \subset A \cup B;$$

2. [Commutativité]

$$A \cup B = B \cup A \quad \text{et} \quad A \cap B = B \cap A;$$

3. [Associativité]

$$A \cup (B \cup C) = (A \cup B) \cup C = A \cup B \cup C$$

et

$$A \cap (B \cap C) = (A \cap B) \cap C = A \cap B \cap C;$$

4. [Distributivité]

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

et

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

Proposition 8.3 (Inclusion-exclusion). *Soient E un ensemble fini et $A, B, A_1, \dots, A_n \subset E$.*

1. On a :

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

Plus généralement, on a :

$$\left| \bigcup_{k=1}^n A_k \right| = \sum_{k=1}^n |A_k| - \sum_{k < l} |A_k \cap A_l| + \dots + (-1)^{n+1} \left| \bigcap_{k=1}^n A_k \right|.$$

2. Si A_1, \dots, A_n forment une partition de E alors

$$\sum_{k=1}^n |A_k| = |E|.$$

Différence et complémentaire

Définition 8.6. *Soient A et B deux sous-ensembles d'un ensemble E .*

1. On appelle différence de A et de B l'ensemble :

$$A \setminus B = \{x \in E : x \in A \text{ et } x \notin B\}.$$

On lit $A \setminus B$ « A privé de B ».

2. On appelle complémentaire de A dans E l'ensemble $E \setminus A$. Cet ensemble peut également être noté $\complement_E A$ ou, lorsqu'il n'y a pas de confusion possible, $\complement A$, A^c ou \bar{A} .

Remarque 8.4. On a $A \setminus A = \emptyset$, $A \setminus \emptyset = A$ et, en général, $A \setminus B \neq B \setminus A$.

Exemple 8.4 (suite). Reprenons les ensembles $A = \{1; 3; 5; 7; 9\}$, $B = \{4; 5; 6\}$ et $E = \mathbf{N}$ de l'Exemple 8.4. On a :

$$A \setminus B = \{1; 3; 7; 9\}, \quad B \setminus A = \{4; 6\} \quad \text{et} \quad A^c = \{0; 2; 4; 6; 8\} \cup (\mathbf{N} \cap [10; +\infty[).$$

Proposition 8.4. *Soient $A, B \subset E$.*

1. Si $B \subset A$, alors $|A \setminus B| = |A| - |B|$.

2. En toute généralité, $|A \setminus B| = |A| - |A \cap B|$.

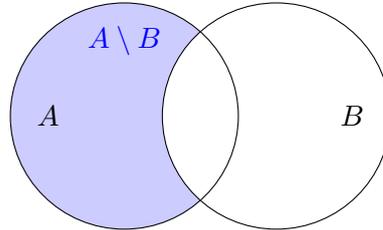


FIGURE 8.2 – Représentation de $A \setminus B$ (bleu) par un diagramme de Venn.

Preuve : Il suffit de vérifier 2.. Écrivons que :

$$A = (A \setminus B) \cup (A \cap B).$$

D'après la Proposition 8.3, on a :

$$|A| = |A \setminus B| + |A \cap B| - |(A \setminus B) \cap (A \cap B)| = |A \setminus B| + |A \cap B|.$$

On en déduit que :

$$|A \setminus B| = |A| - |A \cap B|.$$

□

Proposition 8.5 (Règles de de Morgan). *Soient E un ensemble et $A_1, \dots, A_n \subset E$.*

On a :

$$\left(\bigcup_{k=1}^n A_k \right)^c = \bigcap_{k=1}^n A_k^c$$

et

$$\left(\bigcap_{k=1}^n A_k \right)^c = \bigcup_{k=1}^n A_k^c$$

Différence symétrique

Définition 8.7. *Soient A et B deux sous-ensembles d'un ensemble E .*

On appelle différence symétrique de A et de B l'ensemble :

$$A \Delta B = \{x \in E : x \in A \text{ ou } x \in B \text{ mais } x \notin A \cap B\}.$$

On lit $A \Delta B$ « A delta B ».

Remarque 8.5. On a $A \Delta B = B \Delta A$.

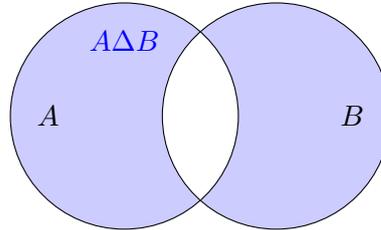


FIGURE 8.3 – Représentation de $A\Delta B$ (bleu) par un diagramme de Venn.

Exemple 8.4 (suite). Reprenons les ensembles $A = \{1; 3; 5; 7; 9\}$, $B = \{4; 5; 6\}$ et $E = \mathbf{N}$ de l'Exemple 8.4. On a :

$$A\Delta B = B\Delta A = \{1; 3; 4; 6; 7; 9\}.$$

Proposition 8.6. Soient E un ensemble et $A, B \subset E$.

On a :

$$\begin{aligned} A\Delta B &= (A \cup B) \setminus (A \cap B) \\ &= (A \setminus B) \cup (B \setminus A). \end{aligned}$$

Preuve : Pour montrer l'égalité entre deux ensembles, une présentation pratique est de réaliser une *table de vérité* (ou de Boole), indiquant dans chaque case si certaines propriétés d'appartenance sont vraies ou fausses selon qu'un élément soit, ou non, dans les ensembles de base intervenant. Dans le cas présent, celles-ci prennent les formes suivantes.

$x \in A$	$x \in B$	$x \in A\Delta B$	$x \in A \cup B$	$x \in A \cap B$	$x \in (A \cup B) \setminus (A \cap B)$
vrai	vrai	faux	vrai	vrai	faux
vrai	faux	vrai	vrai	faux	vrai
faux	vrai	vrai	vrai	faux	vrai
faux	faux	faux	faux	faux	faux

On observe que les valeurs « vrai » et « faux » sont identiques dans les colonnes de $A\Delta B$ et $(A \cup B) \setminus (A \cap B)$, ce qui prouve l'égalité de ces deux ensembles.

$x \in A$	$x \in B$	$x \in A\Delta B$	$x \in A \setminus B$	$x \in B \setminus A$	$x \in (A \setminus B) \cup (B \setminus A)$
vrai	vrai	faux	faux	faux	faux
vrai	faux	vrai	vrai	faux	vrai
faux	vrai	vrai	faux	vrai	vrai
faux	faux	faux	faux	faux	faux

On observe que les valeurs « vrai » et « faux » sont identiques dans les colonnes de $A\Delta B$ et $(A \setminus B) \cup (B \setminus A)$, ce qui prouve l'égalité de ces deux ensembles. \square

Produit cartésien

Définition 8.8. Soient A et B deux ensembles.

On appelle ensemble produit de A par B l'ensemble :

$$A \times B = \{(x; y) : x \in A \text{ et } y \in B\}.$$

Remarque 8.6. En général, $A \times B \neq B \times A$.

Proposition 8.7. Soient A et B deux ensembles finis.

On a :

$$|A \times B| = |A| \times |B|.$$

Remarque 8.7. On note :

$$A^n = \underbrace{A \times A \times \cdots \times A}_{n \text{ fois}}$$

Corollaire 8.1 (Principe fondamental du dénombrement). Si deux expériences, pouvant produire respectivement n et m résultats différents, sont réalisées simultanément, le nombre total de résultats possibles pour ce couple d'expériences est nm .

Exemple 8.5. On lance un dé à 6 faces, donnant un des nombres $1, 2, \dots, 6$ et une pièce tombant soit sur pile soit sur face. Les résultats possibles de cette expérience sont au nombre de $6 \times 2 = 12$:

(1; pile), (1; face),
 (2; pile), (2; face),
 (3; pile), (3; face),
 (4; pile), (4; face),
 (5; pile), (5; face),
 (6; pile), (6; face).

8.2 Combinatoire

8.2.1 p -listes

Définition 8.9. Soit A un ensemble et $p \in \mathbf{N}$.

On appelle p -liste d'éléments de A tout élément de A^p .

Remarque 8.8. Il s'agit d'une liste **ordonnée** de p éléments de A **avec répétitions** possibles. On appelle parfois A un *alphabet* et une p -liste un *mot de longueur p* . Notons que les *mots* ne sont pas nécessairement dans le dictionnaire!

Il découle immédiatement de la Proposition 8.7 que :

Proposition 8.8. *Soit A un ensemble à n éléments et $p \in \mathbf{N}$.*

Le nombre de p -listes d'éléments de A est n^p .

Exercice 8.1. Déterminer le nombre de mots de 5 lettres avec l'alphabet fondamental français.

Solution : Cet alphabet comporte 26 lettres donc le nombre de mots de 5 lettres avec cet alphabet est :

$$26^5 = 11881376.$$

Exercice 8.2. Le Système d'immatriculation des véhicules (SIV) français est basé sur une séquence de deux lettres - trois chiffres - deux lettres. Les lettres I, O et U sont exclues et les associations de lettres SS et WW sont interdites à gauche et l'association SS est interdite à droite. Déterminer le nombre de plaques minéralogiques possibles.

8.2.2 Permutations

Définition 8.10. *Soit A un ensemble fini.*

On appelle permutation de A toute façon d'ordonner les éléments de A .

Remarque 8.9. Il s'agit d'une liste **ordonnée de tous les éléments de A (sans répétition possible)**.

Proposition 8.9. *Soit A un ensemble à n éléments.*

Le nombre de permutations de A est $n! = 1 \times 2 \times \dots \times n$.

Remarque 8.10. L'expression $n!$ se lit « factorielle n ». Par convention, $0! = 1$.

Preuve : Raisonnons par récurrence sur $n \in \mathbf{N}^*$.

Initialisation : Si A ne contient qu'un élément x_1 , il n'y a clairement qu'une permutation des éléments de A et $1! = 1$.

Hérédité : Supposons que pour un certain $n \in \mathbf{N}^*$, le nombre de permutations des éléments d'un ensemble à n éléments soit $n!$ et montrons que le nombre de permutations des éléments d'un ensemble à $n + 1$ éléments est $(n + 1)!$.

Soit $A = \{x_1, x_2, \dots, x_{n+1}\}$ un ensemble à $n + 1$ éléments. Toute permutation de A s'obtient en insérant à l'une des $n + 1$ places possible l'élément x_{n+1} dans une des $n!$ permutations de $\{x_1, x_2, \dots, x_n\}$. Par le Corolaire 8.1, il y a $(n + 1) \times n! = (n + 1)!$ solutions possibles et la propriété considérée est héréditaire à partir du rang $n = 1$.

Conclusion : La propriété considérée étant héréditaire à partir du rang $n = 1$ et vérifiée à ce rang, le principe du raisonnement par récurrence permet de conclure que le nombre de permutations d'un ensemble à n éléments est $n!$. □

Exemple 8.6. Si 10 concurrents se présentent pour une épreuve sportive individuelle, le nombre de façons d'ordonner leurs passages est :

$$10! = 1 \times 2 \times \dots \times 10 = 3628800.$$

8.2.3 Arrangements

Définition 8.11. Soit A un ensemble à n éléments et $0 \leq k \leq n$.

On appelle arrangement de k éléments de A toute façon d'ordonner k éléments distincts de A .

Remarque 8.11. Il s'agit d'une liste **ordonnée** de k éléments de A **sans répétition** possible.

Proposition 8.10. Soit A un ensemble à n éléments et $0 \leq k \leq n$.

Le nombre d'arrangements de k éléments de A est :

$$A_n^k = \frac{n!}{(n-k)!}.$$

Preuve : Pour la première place dans la liste ordonnée, il y a n possibilités, pour la deuxième place, il y a $(n-1)$ possibilités, ..., pour la k^{e} place, il y a $(n-k+1)$ possibilités. Par le principe du dénombrement (Corolaire 8.1), il y a :

$$n \times (n-1) \times \cdots \times (n-k+1) = \frac{n!}{(n-k)!}$$

possibilité. □

Exercice 8.3. Déterminer le nombre de mots de 5 lettres avec l'alphabet fondamental français ne contenant que des lettres distinctes.

Solution : Cet alphabet comporte 26 lettres donc le nombre de mots de 5 lettres ne contenant que des lettres distinctes avec cet alphabet est :

$$A_{26}^5 = 26 \times 25 \times 24 \times 23 \times 22 = 7893600,$$

ce qui est substantiellement moins que les 11881376 mots de 5 lettres avec répétitions possibles (voir Exercice 8.1).

8.2.4 Combinaisons

Définition 8.12. Soit A un ensemble à n éléments et $0 \leq k \leq n$.

On appelle combinaison de k éléments de A tout sous-ensemble de k éléments distincts de A .

Remarque 8.12. Il s'agit d'une liste **non ordonnée** de k éléments de A **sans répétition** possible.

Proposition 8.11. Soit A un ensemble à n éléments et $0 \leq k \leq n$.

Le nombre de combinaisons de k éléments de A est le coefficient binomial :

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

$n \backslash k$	0	1	2	3	4	5
0	1					
1	1	1				
2	1	2	1			
3	1	3	3	1		
4	1	4	6	4	1	
5	1	5	10	10	5	1

FIGURE 8.4 – Les six premières lignes du triangle de Pascal donnant les $\binom{n}{k}$.

Remarque 8.13. Ce nombre est parfois noté C_n^k . Il est utile de simplifier l'expression avant de la taper sur une calculatrice pour éviter de dépasser sa mémoire.

Preuve : À toute combinaison de k éléments parmi n correspondent $k!$ arrangements de k éléments parmi n puisqu'il y a $k!$ façons de les ordonner. Ainsi,

$$A_n^k = k! \binom{n}{k}$$

et donc

$$\binom{n}{k} = \frac{A_n^k}{k!} = \frac{n!}{k!(n-k)!}.$$

□

Exemple 8.7. Considérons un jeu de loto à 49 numéros et 6 numéros par grille. À chaque tirage 6 numéros sont sélectionnés sans remise. Il y a donc :

$$\binom{49}{6} = 13983816$$

tirages possibles et une chance de gain au premier rang d'environ 0,0000072%.

La proposition suivante s'établit par des calculs directs.

Proposition 8.12. On a

1.

$$\binom{n}{k} = \binom{n}{n-k}, \quad \binom{n}{0} = \binom{n}{n} = 1, \quad \binom{n}{1} = \binom{n}{n-1} = n;$$

2.

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}.$$

Le point 2 de la proposition est parfois appelé *règle de Pascal*. Il en découle une présentation des coefficients binomiaux dans un triangle dit *triangle* de Pascal.

Les coefficients binomiaux permettent d'établir une formule pour développer $(a+b)^n$, pour tout $n \in \mathbf{N}^*$.

Proposition 8.13 (Formule du binôme de Newton). *Soient $a, b \in \mathbf{R}$ et $n \in \mathbf{N}^*$.*

On a :

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

Preuve : Par récurrence sur $n \in \mathbf{N}^*$ en exercice. □

Exercice 8.4. Calculer pour tout $n \in \mathbf{N}$:

$$\sum_{k=0}^n (-1)^k \binom{n}{k}.$$

Solution : Pour $n = 0$, on a :

$$\binom{0}{0} = 1$$

et pour $n \in \mathbf{N}^*$, il suffit de remarquer que :

$$\sum_{k=0}^n (-1)^k \binom{n}{k} = \sum_{k=0}^n \binom{n}{k} (-1)^k 1^{n-k} = (1 - 1)^n = 0^n = 0.$$

Chapitre 9

Calcul élémentaire des probabilités

9.1 Vocabulaire probabiliste

Définition 9.1. 1. On appelle expérience aléatoire une procédure donnant un résultat imprévisible mais dont on connaît l'ensemble des résultats possibles.

2. L'ensemble des résultats possibles d'une expérience aléatoire est appelé univers, ensemble fondamental ou ensemble des possibles et est noté Ω .

3. Un élément ω de Ω est appelé une issue ou un événement élémentaire de l'expérience aléatoire.

4. Un sous-ensemble A de Ω est appelé un événement.

Exemple 9.1.

1. Si on lance un dé à 6 faces, on aura $\Omega = \{1; 2; 3; 4; 5; 6\}$.
2. Si on tire 10 pièces à pile ou face, les issues sont les éléments de $\{\text{pile ; face}\}^{10}$.
3. Si on s'intéresse à la durée de vie d'une ampoule, on obtiendra un temps aléatoire dans $\Omega = \mathbf{R}_+$.
4. Si on regarde le déplacement d'une particule dans un verre d'eau, on observera une trajectoire continue.

Remarque 9.1. Pour une expérience donnée, il existe plusieurs façons de modéliser l'univers. Le choix du modèle est primordial et des calculs plus ou moins compliqués peuvent en dépendre.

Exemple 9.2. Le grand duc de Toscane, féru de jeux, avait remarqué que, bien qu'il y ait autant de façons d'écrire 9 et 10 comme somme de 3 nombres compris entre 1 et 6, on obtient plus souvent un total de 10 lorsqu'on lance 3 dés. Il demanda à Galilée de lui donner une explication de ce paradoxe.

La clef est de choisir le bon modèle. On a, ici, au moins deux choix possibles :

- distinguer les dés et considérer que les événements élémentaires sont les triplets de $\Omega_1 = \{1; 2; \dots; 6\}^3$
- ne pas distinguer les dés et considérer que les éléments de Ω_2 sont de la forme $\{d_1; d_2; d_3\}$, avec $d_1; d_2; d_3 \in \{1; 2; \dots; 6\}$, sans que l'ordre soit important.

Nous reprendrons cet exemple par la suite et verrons lequel de ces deux modèles permet d'expliquer le paradoxe.

Les probabilités sont intimement liées aux ensembles. Toutefois, les terminologies diffèrent quelque peu entre théorie des ensembles et théorie des probabilités. Le tableau suivant indique les différences des terminologies.

Vocabulaire ensembliste	Notation	Vocabulaire probabiliste
ensemble	A	événement
ω appartient à A	$\omega \in A$	ω réalise A
A est inclus dans B	$A \subset B$	A implique B
A union B	$A \cup B$	A ou B
A inter B	$A \cap B$	A et B
A et B sont disjoints	$A \cap B = \emptyset$	A est incompatible avec B
le complémentaire de A	A^c	le contraire de A

La probabilité d'un événement s'interprète comme la proportion de d'issues permettant de le réaliser ou la fréquence à laquelle il est réalisé. Ainsi, si un événement a une probabilité proche de 0, il n'est que très peu probable de le voir réalisé alors que s'il a une probabilité proche de 1, on a de grandes chances de le voir ce réaliser. Un événement de probabilité 1 est dit *certain* alors qu'un événement de probabilité 0 est dit *impossible*.

Donnons maintenant une définition formelle de ce que l'on entend par une probabilité.

Définition 9.2. *Soit une expérience aléatoire et Ω l'univers associé.*

Une probabilité \mathbf{P} sur Ω est une application définie sur les événements et vérifiant :

1. pour tout événement A :

$$0 \leq \mathbf{P}[A] \leq 1;$$

2. $\mathbf{P}[\Omega] = 1$;

3. [**σ -additivité**] : pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'événements telle que $A_i \cap A_j = \emptyset$ si $i \neq j$, on a :

$$\mathbf{P} \left[\bigcup_{n=0}^{+\infty} A_n \right] = \sum_{n=0}^{\infty} \mathbf{P} [A_n].$$

On appelle alors (Ω, \mathbf{P}) un espace probabilisé.

Proposition 9.1. *Soient (Ω, \mathbf{P}) un espace probabilisé et A et B deux événements.*

On a :

1. $\mathbf{P} [A^c] = 1 - \mathbf{P} [A]$;
2. $\mathbf{P} [\emptyset] = 0$;
3. $\mathbf{P} [A \cup B] = \mathbf{P} [A] + \mathbf{P} [B] - \mathbf{P} [A \cap B]$;
4. $\mathbf{P} [A \cup B] \leq \mathbf{P} [A] + \mathbf{P} [B]$;
5. si $A \subset B$, alors $\mathbf{P} [A] \leq \mathbf{P} [B]$.

Preuve :

1. On a $\Omega = A \cup A^c$ donc $\mathbf{P} [A] + \mathbf{P} [A^c] = \mathbf{P} [\Omega]$ d'où $\mathbf{P} [A^c] = 1 - \mathbf{P} [A]$.
2. C'est une conséquence du point précédent avec $A = \Omega$.

3. On a $A = A \setminus (A \cap B) \cup (A \cap B)$ et cette union est disjointe donc

$$\mathbf{P}[A] = \mathbf{P}[A \setminus (A \cap B)] + \mathbf{P}[A \cap B].$$

De même,

$$\mathbf{P}[B] = \mathbf{P}[B \setminus (A \cap B)] + \mathbf{P}[A \cap B].$$

En utilisant que les événements $A \setminus (A \cap B)$, $B \setminus (A \cap B)$ et $A \cap B$ sont deux à deux disjoints et que leur union est $A \cup B$, on obtient que :

$$\begin{aligned} \mathbf{P}[A \cup B] &= \mathbf{P}[A \setminus (A \cap B)] + \mathbf{P}[B \setminus (A \cap B)] + \mathbf{P}[A \cap B] \\ &= \mathbf{P}[A \setminus (A \cap B)] + \mathbf{P}[A \cap B] + \mathbf{P}[B \setminus (A \cap B)] + \mathbf{P}[A \cap B] - \mathbf{P}[A \cap B] \\ &= \mathbf{P}[A] + \mathbf{P}[B] - \mathbf{P}[A \cap B]. \end{aligned}$$

4. C'est une conséquence du point précédent et du fait que $\mathbf{P}[A \cap B] \geq 0$.

5. Puisque $A \subset B$, $B = A \cup (B \setminus A)$ et cette union est disjointe. Ainsi,

$$\mathbf{P}[B] = \mathbf{P}[A] + \underbrace{\mathbf{P}[B \setminus A]}_{\geq 0} \geq \mathbf{P}[A].$$

□

Exercice 9.1. On donne $\mathbf{P}[A] = 0,3$, $\mathbf{P}[A \cap B] = 0,1$, $\mathbf{P}[A \cup B] = 0,8$, $\mathbf{P}[C] = 0,2$ et B et C sont incompatibles. Déterminer $\mathbf{P}[B]$, $\mathbf{P}[B \cap C]$ et $\mathbf{P}[B \cup C]$.

Proposition 9.2 (Formule des probabilités totales - version 1). Soient (Ω, \mathbf{P}) un espace probabilisé et $(B_i)_{i \in I}$ une partition finie ou dénombrable de Ω .

Pour tout événement A , on a :

$$\mathbf{P}[A] = \sum_{i \in I} \mathbf{P}[A \cap B_i].$$

Preuve : Puisque $(B_i)_{i \in I}$ une partition de Ω , on a :

$$\mathbf{P}[A] = \mathbf{P}\left[A \cap \bigcup_{i \in I} B_i\right] = \mathbf{P}\left[\bigcup_{i \in I} (A \cap B_i)\right] = \sum_{i \in I} \mathbf{P}[A \cap B_i].$$

□

Corollaire 9.1. Si Ω est fini ou dénombrable (i.e. ses éléments peuvent être indicés par les entiers naturels) alors pour tout événement A , on a :

$$\mathbf{P}[A] = \sum_{\omega \in A} \mathbf{P}[\omega].$$

9.2 Équiprobabilité

Définition[-Théorème] 9.3. *Soit une expérience aléatoire dont l'univers Ω est fini.*

On parle d'équiprobabilité lorsque les événements élémentaires ont la même probabilité $p = \frac{1}{|\Omega|}$.

Preuve : Notons p la valeur commune des probabilités des événements élémentaires. En utilisant la Proposition 9.1, on obtient que :

$$1 = \mathbf{P}[\Omega] = \sum_{\omega \in \Omega} \mathbf{P}[\omega] = |\Omega|p.$$

□

Proposition 9.3. *Soit une expérience aléatoire dont l'univers Ω est fini et telle que les événements élémentaires soient équiprobables.*

Pour tout événement $A \subset \Omega$, on a :

$$\mathbf{P}[A] = \frac{|A|}{|\Omega|}.$$

Preuve : En utilisant la Proposition 9.1 et la définition précédente, on obtient :

$$\mathbf{P}[A] = \sum_{\omega \in A} \mathbf{P}[\omega] = \sum_{\omega \in A} \frac{1}{|\Omega|} = \frac{|A|}{|\Omega|}.$$

□

Remarque 9.2. La formule précédente n'est valable qu'en cas d'équiprobabilité des événements élémentaires.

Exemple 9.2 (suite). Reprenons l'Exemple 9.2. Si l'on ne distingue pas les dés, les événements élémentaires ne sont pas équiprobables alors qu'ils le sont si l'on distingue les dés. Il convient donc de choisir l'univers $\Omega_1 = \{1; 2; \dots; 6\}^3$ pour lever le paradoxe apparent.

Les façons d'écrire 9 comme somme de 3 entiers compris entre 1 et 6 sont :

$$9 = 6 + 2 + 1, 9 = 5 + 3 + 1, 9 = 5 + 2 + 2, 9 = 4 + 4 + 1, 9 = 4 + 3 + 2, \text{ et } 9 = 3 + 3 + 3.$$

Les façons d'écrire 10 comme somme de 3 entiers compris entre 1 et 6 sont :

$$10 = 6 + 3 + 1, 10 = 6 + 2 + 2, 10 = 5 + 3 + 2, 10 = 5 + 4 + 1, 10 = 4 + 4 + 2, \text{ et } 10 = 4 + 3 + 3.$$

En distinguant les dés, on voit que les triplets permettant d'obtenir pour somme 9 sont :

$$\begin{aligned} &(6; 2; 1), (6; 1; 2), (2; 6; 1), (2; 1; 6) (1; 6; 2), (1; 2; 6), \\ &(5; 3; 1), (5; 1; 3), (3; 5; 1), (3; 1; 5) (1; 5; 3), (1; 3; 5), \\ &(5; 2; 2), (2; 5; 2), (2; 2; 5), \end{aligned}$$

$$(4; 4; 1), (4; 1; 4), (1; 4; 4),$$

$$(4; 3; 2), (4; 2; 3), (3; 4; 2), (3; 2; 4) (2; 4; 3), (2; 3; 4),$$

et

$$(3; 3; 3).$$

Les triplets permettant d'obtenir pour somme 10 sont quant à eux :

$$(6; 3; 1), (6; 1; 3), (3; 6; 1), (3; 1; 6) (1; 6; 3), (1; 3; 6),$$

$$(6; 2; 2), (2; 6; 2), (2; 2; 6),$$

$$(5; 3; 2), (5; 2; 3), (3; 5; 2), (3; 2; 5) (2; 5; 3), (2; 3; 5),$$

$$(5; 4; 1), (5; 1; 4), (4; 5; 1), (4; 1; 5) (1; 5; 4), (1; 4; 5),$$

$$(4; 4; 2), (4; 2; 4), (2; 4; 4),$$

$$(4; 4; 2), (4; 2; 4), (2; 4; 4),$$

$$(4; 3; 3), (3; 4; 3), (3; 3; 4),$$

Il y a donc 25 triplets équiprobables permettant d'obtenir pour somme 9 et 27 triplets équiprobables permettant d'obtenir pour somme 10. Puisque $|\Omega_1| = 6^3 = 216$, la probabilité d'obtenir pour somme 9 est $\frac{25}{216} \simeq 0,116$ alors que la probabilité d'obtenir pour somme 10 est $\frac{27}{216} = \frac{1}{8} = 0,125$.

9.3 Indépendance

Définition 9.4. Soient (Ω, \mathbf{P}) un espace probabilisé et A et B deux événements.

On dit que A et B sont indépendants si :

$$\mathbf{P}[A \cap B] = \mathbf{P}[A]\mathbf{P}[B].$$

Remarque 9.3.

1. Si A et B sont indépendants, la connaissance de la réalisation de l'un n'influence pas la probabilité de l'autre.
2. Si $\mathbf{P}[A] > 0$, $\mathbf{P}[B] > 0$ et A et B sont incompatibles, alors A et B ne sont pas indépendants. En effet,

$$0 = \mathbf{P}[A \cap B] \neq \mathbf{P}[A]\mathbf{P}[B] > 0.$$

3. Si $\mathbf{P}[A] = 0$ ou 1 et B est un événement quelconque, A et B sont indépendants (le vérifier en exercice).

Exemple 9.3. On place un jeu de 9 boules rouges, numérotées de 1 à 9, et un jeu de 9 boules blanches, numérotées de 1 à 9, dans une urne. On tire ensuite une boule au hasard dans l'urne. Les événements A « la boule tirée est rouge » et B « la boule tirée porte le numéro 1 » sont indépendants. En effet, on a :

$$\mathbf{P}[A] = \frac{1}{2}, \quad \mathbf{P}[B] = \frac{1}{9} \quad \text{et} \quad \mathbf{P}[A \cap B] = \frac{1}{18} = \frac{1}{2} \times \frac{1}{9}.$$

Exemple 9.4. On place un jeu de 9 boules rouges, numérotées de 1 à 9, et un jeu de 9 boules blanches, numérotées de 11 à 19, dans une urne. On tire ensuite une boule au hasard dans l'urne. Les événements A « la boule tirée est rouge » et B « la boule tirée porte le numéro 13 » ne sont pas indépendants. En effet, on a :

$$\mathbf{P}[A] = \frac{1}{2}, \quad \mathbf{P}[B] = \frac{1}{18} \quad \text{et} \quad \mathbf{P}[A \cap B] = 0 \neq \frac{1}{2} \times \frac{1}{18}.$$

Proposition 9.4. Soient (Ω, \mathbf{P}) un espace probabilisé et A et B deux événements indépendants.

Alors,

1. A et B^c sont indépendants ;
2. A^c et B sont indépendants ;
3. A^c et B^c sont indépendants.

Preuve : Il suffit de prouver le premier point. On a :

$$\begin{aligned} \mathbf{P}[A \cap B^c] &= \mathbf{P}[A] - \mathbf{P}[A \cap B] \\ &= \mathbf{P}[A] - \mathbf{P}[A] \mathbf{P}[B] && \text{par indépendance de } A \text{ et } B \\ &= \mathbf{P}[A] (1 - \mathbf{P}[B]) = \mathbf{P}[A] \mathbf{P}[B^c] \end{aligned}$$

ce qui prouve que A et B^c sont indépendants. □

Remarque 9.4. Attention : si A et B sont indépendants et B et C sont indépendants, on ne peut pas affirmer en général que A et C sont indépendants. Pour le voir, il suffit de considérer deux événements indépendants A et B tels que $0 < \mathbf{P}[A], \mathbf{P}[B] < 1$ et l'événement $C = A^c$.

Définition 9.5. Soient (Ω, \mathbf{P}) un espace probabilisé et $(A_i)_{i \in I}$ une famille finie ou dénombrable d'événements.

On dit que les $(A_i)_{i \in I}$ sont mutuellement indépendants si pour toute sous-famille finie A_{i_1}, \dots, A_{i_k} des $(A_i)_{i \in I}$, on a :

$$\mathbf{P}[A_{i_1} \cap \dots \cap A_{i_k}] = \mathbf{P}[A_{i_1}] \times \dots \times \mathbf{P}[A_{i_k}].$$

9.4 Probabilités conditionnelles

Définition 9.6. Soient (Ω, \mathbf{P}) un espace probabilisé et A et B deux événements avec $\mathbf{P}[B] > 0$.

On appelle probabilité conditionnelle de A sachant B la quantité :

$$\mathbf{P}[A|B] = \frac{\mathbf{P}[A \cap B]}{\mathbf{P}[B]}.$$

Remarque 9.5.

1. La probabilité conditionnelle de A sachant B représente la probabilité d'observer A si l'on sait que B est réalisé.

2. On a ainsi $\mathbf{P}[A \cap B] = \mathbf{P}[A|B]\mathbf{P}[B]$.
3. La probabilité conditionnelle de A sachant B est parfois notée $\mathbf{P}_B[A]$.
4. Si $\mathbf{P}[A], \mathbf{P}[B] > 0$, les assertions suivantes sont équivalentes :
 - (a) A et B sont indépendants ;
 - (b) $\mathbf{P}[A|B] = \mathbf{P}[A]$;
 - (c) $\mathbf{P}[B|A] = \mathbf{P}[B]$.

Exemple 9.5. Une urne contient une boule rouge portant le numéro 1, deux boules rouges portant le numéro 2, quatre boules blanches portant le numéro 1 et trois boules blanches portant le numéro 2. On tire au hasard une boule dans l'urne. Notons A l'événement « la boule tirée est blanche » et B l'événement « la boule tirée porte le numéro 2 ».

On a :

$$\mathbf{P}[A] = \frac{7}{10}, \quad \mathbf{P}[B] = \frac{1}{2}, \quad \mathbf{P}[A \cap B] = \frac{3}{10}, \quad \mathbf{P}[A|B] = \frac{3}{5} \quad \text{et} \quad \mathbf{P}[B|A] = \frac{3}{7}.$$

Proposition 9.5. Soient (Ω, \mathbf{P}) un espace probabilisé et B un événement avec $\mathbf{P}[B] > 0$. Alors, $\mathbf{P}[\cdot|B]$ est une probabilité sur B .

Preuve : Puisque, pour tout $A \subset B$,

$$0 \leq \mathbf{P}[A \cap B] \leq \mathbf{P}[B],$$

on a :

$$0 \leq \mathbf{P}[A|B] = \frac{\mathbf{P}[A \cap B]}{\mathbf{P}[B]} \leq 1.$$

D'autre part, $\mathbf{P}[B|B] = 1$. Finalement, si $(A_n)_{n \in \mathbf{N}} \subset B$ est une suite d'événements deux à deux incompatibles, $(A_n \cap B)_{n \in \mathbf{N}} \subset B$ est une suite d'événements deux à deux incompatibles et on a :

$$\begin{aligned} \mathbf{P}\left[\bigcup_{n=1}^{+\infty} A_n | B\right] &= \frac{\mathbf{P}\left[\left(\bigcup_{n=1}^{+\infty} A_n\right) \cap B\right]}{\mathbf{P}[B]} \\ &= \frac{\mathbf{P}\left[\bigcup_{n=1}^{+\infty} (A_n \cap B)\right]}{\mathbf{P}[B]} \\ &= \frac{\sum_{n=0}^{+\infty} \mathbf{P}[A_n \cap B]}{\mathbf{P}[B]} \\ &= \sum_{n=0}^{+\infty} \mathbf{P}[A_n | B]. \end{aligned}$$

□

On en déduit immédiatement le corollaire suivant.

Corollaire 9.2. 1. $\mathbf{P}[A^c|B] = 1 - \mathbf{P}[A|B]$;

2. $\mathbf{P}[\emptyset|B] = 0$;

3. $\mathbf{P}[A \cup C|B] = \mathbf{P}[A|B] + \mathbf{P}[C|B] - \mathbf{P}[A \cap C|B]$.

Proposition 9.6 (Conditionnements successifs). *Soient (Ω, \mathbf{P}) un espace probabilisé et A_1, \dots, A_n des événements tels que $\mathbf{P}[A_1], \mathbf{P}[A_1 \cap A_2], \dots, \mathbf{P}[A_1 \cap A_2 \cap \dots \cap A_{n-1}] > 0$.*

On a :

$$\mathbf{P}[A_1 \cap A_2 \cap \dots \cap A_n] = \mathbf{P}[A_n|A_1 \cap A_2 \cap \dots \cap A_{n-1}] \mathbf{P}[A_{n-1}|A_1 \cap A_2 \cap \dots \cap A_{n-2}] \dots \mathbf{P}[A_2|A_1] \mathbf{P}[A_1].$$

Preuve : Par récurrence sur n en exercice. □

Exemple 9.6. Un jeu de 52 cartes contient 13 cartes de chaque couleur (pique, carreau, cœur, trèfle). On tire successivement 5 cartes sans remise. On s'intéresse à la probabilité de tirer successivement deux carreaux, deux cœurs puis un pique. En définissant les événements :

- E_1 : « la 1^{re} carte tirée est un carreau »,
- E_2 : « la 2^e carte tirée est un carreau »,
- E_3 : « la 3^e carte tirée est un cœur »,
- E_4 : « la 4^e carte tirée est un cœur »,
- E_5 : « la 5^e carte tirée est un pique »,

on doit calculer $\mathbf{P}[E_1 \cap E_2 \cap E_3 \cap E_4 \cap E_5]$. Il est clair que :

$$\mathbf{P}[E_1] = \frac{13}{52} = \frac{1}{4}, \quad \mathbf{P}[E_2|E_1] = \frac{12}{51} = \frac{4}{17},$$

$$\mathbf{P}[E_3|E_1 \cap E_2] = \frac{13}{50}, \quad \mathbf{P}[E_4|E_1 \cap E_2 \cap E_3] = \frac{12}{49},$$

et

$$\mathbf{P}[E_5|E_1 \cap E_2 \cap E_3 \cap E_4] = \frac{13}{48}.$$

En utilisant la Proposition 9.6, on obtient que :

$$\begin{aligned} \mathbf{P}[E_1 \cap E_2 \cap E_3 \cap E_4 \cap E_5] &= \mathbf{P}[E_1] \times \mathbf{P}[E_2|E_1] \times \mathbf{P}[E_3|E_1 \cap E_2] \times \mathbf{P}[E_4|E_1 \cap E_2 \cap E_3] \\ &\quad \times \mathbf{P}[E_5|E_1 \cap E_2 \cap E_3 \cap E_4] \\ &= \frac{1}{4} \times \frac{4}{17} \times \frac{13}{50} \times \frac{12}{49} \times \frac{13}{48} \\ &= \frac{169}{166600} \\ &\simeq 0,00101. \end{aligned}$$

Proposition 9.7 (Formule des probabilités totales - version 2). *Soient (Ω, \mathbf{P}) un espace probabilisé et $(B_i)_{i \in I}$ une partition finie ou dénombrable de Ω telle que, pour tout $i \in I$, $\mathbf{P}[B_i] > 0$.*

Pour tout événement A , on a :

$$\mathbf{P}[A] = \sum_{i \in I} \mathbf{P}[A|B_i] \mathbf{P}[B_i].$$

Preuve : Puisque $(B_i)_{i \in I}$ une partition de Ω , on a :

$$\mathbf{P}[A] = \mathbf{P}\left[A \cap \bigcup_{i \in I} B_i\right] = \mathbf{P}\left[\bigcup_{i \in I} (A \cap B_i)\right] = \sum_{i \in I} \mathbf{P}[A \cap B_i] = \sum_{i \in I} \mathbf{P}[A|B_i]\mathbf{P}[B_i]$$

la dernière égalité étant valide puisque, pour tout $i \in I$, $\mathbf{P}[B_i] > 0$. □

Exemple 9.7. Dans une usine, deux chaînes de production A et B fabriquent le même produit. On sait que 10% des produits fabriqués par A sont défectueux alors que seulement 1% des produits fabriqués par B le sont. De plus, 20% des produits sont fabriqués par la chaîne A.

On prélève uniformément au hasard un produit. Notons A l'événement « le produit prélevé provient de la chaîne A » et D l'événement « le produit prélevé est défectueux ». On a, bien sûr, que A^c est l'événement « le produit prélevé provient de la chaîne B » et, d'après les données, que :

$$\mathbf{P}[A] = 0,2, \quad \mathbf{P}[A^c] = 1 - 0,2 = 0,8, \quad \mathbf{P}[D|A] = 0,1, \quad \text{et} \quad \mathbf{P}[D|A^c] = 0,01.$$

En utilisant que A et A^c forment une partition de l'univers et la formule des probabilités totales (Proposition 9.7), on obtient que :

$$\mathbf{P}[D] = \mathbf{P}[D|A]\mathbf{P}[A] + \mathbf{P}[D|A^c]\mathbf{P}[A^c] = 0,1 \times 0,2 + 0,01 \times 0,8 = 0,028.$$

On en déduit que 2,8% des produits fabriqués par l'usine sont défectueux.

Proposition 9.8 (Formule de Bayes). *Soient (Ω, \mathbf{P}) un espace probabilisé, $(B_i)_{i \in I}$ une partition finie ou dénombrable de Ω telle que, pour tout $i \in I$, $\mathbf{P}[B_i] > 0$ et A un événement tel que $\mathbf{P}[B_i] > 0$.*

Pour tout $i \in I$, on a :

$$\mathbf{P}[B_i|A] = \frac{\mathbf{P}[A|B_i]\mathbf{P}[B_i]}{\sum_{j \in I} \mathbf{P}[A|B_j]\mathbf{P}[B_j]}.$$

Preuve : Il suffit d'écrire que :

$$\mathbf{P}[B_i|A] = \frac{\mathbf{P}[B_i \cap A]}{\mathbf{P}[A]} = \frac{\mathbf{P}[A|B_i]\mathbf{P}[B_i]}{\mathbf{P}[A]} = \frac{\mathbf{P}[A|B_i]\mathbf{P}[B_i]}{\sum_{j \in I} \mathbf{P}[A|B_j]\mathbf{P}[B_j]}.$$

où l'on a utilisé la formule des probabilités (Proposition 9.7) totales dans la dernière égalité. □

Exemple 9.7 (suite). Reprenons l'Exemple 9.7 et déterminons la probabilité pour qu'un produit sélectionné au hasard provienne de la chaîne A sachant qu'il est défectueux. Par la formule de Bayes (Proposition 9.8), cette probabilité est :

$$\mathbf{P}[A|D] = \frac{\mathbf{P}[D|A]\mathbf{P}[A]}{\mathbf{P}[D|A]\mathbf{P}[A] + \mathbf{P}[D|A^c]\mathbf{P}[A^c]} = \frac{0,1 \times 0,2}{0,1 \times 0,2 + 0,01 \times 0,8} \simeq 0,7143.$$

Ainsi, 71,43% des produits défectueux de l'usine ont été fabriqués par la chaîne A.

Chapitre 10

Variables aléatoires

10.1 Premières définitions et exemples

Définition 10.1. Soit (Ω, \mathbf{P}) un espace probabilisé.

On appelle variable aléatoire toute application $X : \Omega \rightarrow \mathbf{R}$.

Si l'ensemble $X(\Omega)$ des valeurs numériques prises par X est fini ou dénombrable, la variable aléatoire X est dite discrète. Si $X(\Omega)$ est une réunion d'intervalles de \mathbf{R} , la variable aléatoire X est dite continue.

Exemple 10.1.

1. Si l'on lance un dé classique, la variable aléatoire X donnant le résultat du lancé est une variable aléatoire discrète prenant ses valeurs dans $X(\Omega) = \{1; 2; 3; 4; 5; 6\}$.
2. Si l'on lance simultanément n pièces équilibrées, la variable aléatoire comptant le nombre de faces obtenues est discrète et prend ses valeurs dans $X(\omega) = \{1; 2; \dots; n\}$.
3. La durée de vie X d'une ampoule est une variable aléatoire continue. Une ampoule donnée fonctionnera un temps $t > 0$ avant de griller et il n'y a *a priori* pas de durée de vie maximale. Ainsi, $X(\Omega) = \mathbf{R}_+^*$.

Les ensembles $\{X = x\}, \{X \leq x\}, \{X \geq x\}, \{X < x\}, \{X > x\}$ sont des événements. On note généralement $\mathbf{P}[X = x], \mathbf{P}[X \leq x], \mathbf{P}[X \geq x], \mathbf{P}[X < x], \mathbf{P}[X > x]$ au lieu de $\mathbf{P}[\{X = x\}], \mathbf{P}[\{X \leq x\}], \mathbf{P}[\{X \geq x\}], \mathbf{P}[\{X < x\}], \mathbf{P}[\{X > x\}]$.

Définition 10.2. Soit (Ω, \mathbf{P}) un espace probabilisé et X une variable aléatoire.

On appelle fonction de répartition de X la fonction :

$$F_X : X(\Omega) \rightarrow [0; 1].$$
$$x \mapsto F_X(x) = \mathbf{P}[X \leq x]$$

Remarque 10.1.

1. S'il n'y a pas de confusion possible, on note simplement F la fonction de répartition d'une variable aléatoire X .
2. La fonction de répartition F d'une variable aléatoire X est croissante, sa limite en $-\infty$ est 0 et sa limite en $+\infty$ est 1. Si X est discrète F est une fonction en escalier alors que si X est continue F est continue.

Définition 10.3. Soit (Ω, \mathbf{P}) un espace probabilisé et X et Y deux variables aléatoires sur cet espace.

On dit que X et Y sont indépendantes si pour tous $A, B \subset \mathbf{R}$, on a :

$$\mathbf{P}[X \in A, Y \in B] = \mathbf{P}[X \in A] \mathbf{P}[Y \in B].$$

Remarque 10.2. Il suffit de le vérifier pour A de la forme $] - \infty; a]$ et B de la forme $] - \infty; b]$.

10.2 Variables aléatoires discrètes

10.2.1 Loi d'une variable aléatoire discrète

Définition[-Théorème] 10.4. Soit (Ω, \mathbf{P}) un espace probabilisé et X une variable aléatoire discrète sur cet espace.

On appelle loi de la variable aléatoire X la probabilité sur $X(\Omega)$ définie par :

$$\begin{aligned} \mathbf{P}_X : X(\Omega) &\longrightarrow [0; 1]. \\ x &\longmapsto \mathbf{P}_X(x) = \mathbf{P}[X = x] \end{aligned}$$

Remarque 10.3. On note indifféremment $\mathbf{P}_X(x)$, $\mathbf{P}[X = x]$ ou $\mathbf{P}[\{X = x\}]$. L'ensemble $X(\Omega)$ est appelé *support* de X .

Exemple 10.2. Si l'on lance un dé équilibré à 6 faces, le résultat est la variable aléatoire X dont la loi est caractérisée par :

$$\mathbf{P}[X = 1] = \mathbf{P}[X = 2] = \mathbf{P}[X = 3] = \mathbf{P}[X = 4] = \mathbf{P}[X = 5] = \mathbf{P}[X = 6] = \frac{1}{6}.$$

Il s'agit de la loi uniforme sur $\{1; 2; \dots; 6\}$ (voir Section 10.2.3).

Proposition 10.1. Si X est une variable aléatoire discrète prenant les valeurs $x_1 < x_2 < \dots$, la fonction de répartition F de X est donnée par :

$$F(x) = \mathbf{P}[X \leq x] = \mathbf{P}[X = x_1] + \mathbf{P}[X = x_2] + \dots + \mathbf{P}[X = x_k], \quad \text{pour tout } x \in [x_k; x_{k+1}[.$$

10.2.2 Moments d'une variable aléatoire discrète

Définition 10.5. Soit X une variable aléatoire discrète prenant pour valeurs x_1, x_2, \dots .

Si $\sum_i |x_i| \mathbf{P}[X = x_i] < +\infty$, on appelle espérance de X la quantité :

$$\mathbf{E}[X] = \sum_{i=1}^N x_i \mathbf{P}[X = x_i].$$

Remarque 10.4.

1. Pour reprendre le vocabulaire statistique, l'espérance mathématique correspond à une moyenne espérée si l'on considère un grand nombre de copies indépendantes de même loi que X .
2. Si $N = +\infty$, l'espérance peut être infinie ou ne pas avoir de sens alors que si $N < +\infty$ l'espérance est toujours bien définie.

3. L'espérance est un paramètre de position et ne suffit pas à caractériser la loi d'une variable aléatoire. On peut compléter l'information sur la loi d'une variable aléatoire en considérant les *moments* d'ordres supérieurs et la variance de la variable aléatoire.

Définition 10.6. Soit X une variable aléatoire discrète prenant pour valeurs x_1, x_2, \dots .

1. On appelle moment d'ordre k de X , si elle existe, la quantité $\mathbf{E}[X^k]$.
2. On appelle variance de X , si elle existe la quantité, $\mathbf{V}[X] = \mathbf{E}[(X - \mathbf{E}[X])^2]$.
3. On appelle écart-type de X , si elle existe la quantité, $\sigma = \sqrt{\mathbf{V}[X]}$.

Proposition 10.2. Soient X et Y deux variables aléatoires admettant un moment d'ordre 2 fini et $a, b \in \mathbf{R}$.

On a :

1. $\mathbf{E}[aX + bY] = a\mathbf{E}[X] + b\mathbf{E}[Y]$;
2. $\mathbf{V}[X] = \mathbf{E}[X^2] - \mathbf{E}[X]^2$;
3. $\mathbf{V}[aX + b] = a^2\mathbf{V}[X]$;
4. Si X et Y sont **indépendantes**,

$$\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$$

et

$$\mathbf{V}[X + Y] = \mathbf{V}[X - Y] = \mathbf{V}[X] + \mathbf{V}[Y].$$

10.2.3 Lois discrètes usuelles

Loi uniforme discrète

Définition 10.7. On appelle loi uniforme sur $\{1, 2, \dots, n\}$ la loi dont le support est $\{1, 2, \dots, n\}$ et dont les probabilités des événements élémentaires sont identiques, c'est-à-dire telle qu'une variable aléatoire (v.a.) suivant cette loi vérifie :

$$\mathbf{P}[X = k] = \frac{1}{n}, \quad \text{pour tout } k \in \{1, 2, \dots, n\}.$$

Cette loi est notée $\mathcal{U}(\{1, 2, \dots, n\})$.

Remarque 10.5.

1. Cette loi modélise l'équiprobabilité.
2. On note $X \sim \mathcal{U}(\{1, 2, \dots, n\})$ pour dire que la v.a. X suit la loi $\mathcal{U}(\{1, 2, \dots, n\})$.

Exemple 10.3. Considérons une urne contenant 50 boules indiscernables numérotées de 1 à 50. Si l'on choisit une boule au hasard dans l'urne, la variable aléatoire indiquant le numéro de la boule suit une loi uniforme sur $\{1, 2, \dots, 50\}$.

Proposition 10.3. Soit $X \sim \mathcal{U}(\{1, 2, \dots, n\})$.

On a :

$$\mathbf{E}[X] = \frac{n+1}{2} \quad \text{et} \quad \mathbf{V}[X] = \frac{n^2-1}{12}.$$

Preuve : On a :

$$\begin{aligned} \mathbf{E}[X] &= \sum_{k=1}^n \frac{k}{n} = \frac{1}{n} \sum_{k=1}^n k \\ &= \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}. \end{aligned}$$

On a utilisé, dans ce calcul, que la suite $(u_n)_{n \in \mathbf{N}}$ de terme général $u_n = n$ est arithmétique de raison 1 et que l'on sait en calculer les sommes partielles.

On montre, par récurrence, que pour tout $n \in \mathbf{N}$:

$$\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}.$$

On en déduit que :

$$\begin{aligned} \mathbf{V}[X] &= \mathbf{E}[X^2] - \mathbf{E}[X]^2 = \frac{1}{n} \sum_{k=1}^n k^2 - \left(\frac{n+1}{2}\right)^2 \\ &= \frac{1}{n} \frac{n(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} = \frac{2(n+1)(2n+1)}{12} - \frac{3(n+1)^2}{12} \\ &= \frac{(n+1)(4n+2) - (n+1)(3n+3)}{12} = \frac{(n+1)(n-1)}{12} = \frac{n^2-1}{12}. \end{aligned}$$

□

Loi de Bernoulli

Considérons une expérience aléatoire possédant (exactement) deux issues alternatives du type « succès » ou « échec », « vrai » ou « faux », « pile » ou « face », ... Une telle expérience est appelée *épreuve de Bernoulli*. On modélise une telle expérience par une v.a. X telle que l'événement $\{X = 1\}$ représente un succès et l'événement $\{X = 0\}$ un échec. On a alors $X(\Omega) = \{0; 1\}$ et si $\mathbf{P}[X = 1] = p \in [0; 1]$, $\mathbf{P}[X = 0] = 1 - \mathbf{P}[X = 1] = 1 - p$.

Définition 10.8. On dit qu'une v.a. X telle que $X(\Omega) = \{0; 1\}$, $\mathbf{P}[X = 1] = p \in [0; 1]$, $\mathbf{P}[X = 0] = 1 - p$ suit une loi de Bernoulli de paramètre p .

On note alors $X \sim \text{Ber}(p)$.

Exemple 10.4. Considérons un tirage à pile ou face d'une pièce bien équilibrée. La variable aléatoire définie par :

$$X = \begin{cases} 1 & \text{si la pièce tombe sur face} \\ 0 & \text{si la pièce tombe sur pile} \end{cases}$$

suit la loi de Bernoulli de paramètre $\frac{1}{2}$.

Proposition 10.4. Si $X \sim \text{Ber}(p)$, on a :

$$\mathbf{E}[X] = p \quad \text{et} \quad \mathbf{V}[X] = p(1-p).$$

Preuve : Exercice. □

Loi binomiale

Lorsqu'une épreuve de Bernoulli est répétée plusieurs fois, disons n fois, **indépendamment**, on peut s'intéresser au nombre de succès obtenus lors de ces n expériences ; ce nombre est simplement la somme des variables de Bernoulli servant à modéliser ces expériences répétées.

Définition 10.9. On dit qu'une v.a. X suit une loi binomiale de paramètres $n \in \mathbf{N}$ et $p \in [0; 1]$ si X s'écrit sous la forme :

$$X = \sum_{k=1}^n X_k,$$

où X_1, \dots, X_n sont **indépendantes et identiquement distribuées (i.i.d.)** de loi de Bernoulli de paramètre p .

On note alors $X \sim \mathcal{Bin}(n, p)$.

Remarque 10.6. Le support de la loi $\mathcal{Bin}(n, p)$ est $\{0, \dots, n\}$.

Exemple 10.5. Considérons 10 tirages successifs à pile ou face d'une pièce tombant sur face avec probabilité p . Pour $k = 1, \dots, n$, la variable aléatoire définie par :

$$X_k = \begin{cases} 1 & \text{si la } k^{\text{e}} \text{ pièce tombe sur face} \\ 0 & \text{si la } k^{\text{e}} \text{ pièce tombe sur pile} \end{cases}$$

suit la loi de Bernoulli de paramètre p . De plus, X_1, \dots, X_n sont indépendantes. Ainsi, le nombre de faces obtenu lors de ces tirages est :

$$X = \sum_{k=1}^n X_k$$

et suit la loi $\mathcal{Bin}(10, p)$.

Proposition 10.5. Soit $X \sim \mathcal{Bin}(n, p)$.

1. Pour tout $k \in \{1, \dots, n\}$, on a :

$$\mathbf{P}[X = k] = \binom{n}{k} p^k (1 - p)^{n-k}.$$

2. On a :

$$\mathbf{E}[X] = np \quad \text{et} \quad \mathbf{V}[X] = np(1 - p).$$

Preuve :

1. Pour déterminer les probabilités des événements élémentaires d'une variable aléatoire suivant la loi $\mathcal{Bin}(n, p)$, il faut, dans un premier temps, déterminer le nombre de possibilités d'obtenir k succès au cours de n expériences successives. Il s'agit du nombre de combinaisons (non ordonnées) de k objets pris parmi n , c'est-à-dire $\binom{n}{k}$. Il suffit ensuite de multiplier par les probabilités des k succès et $n - k$ échecs pour obtenir le résultat.

2. Puisque $X_1, \dots, X_n \sim \text{Ber}(p)$ et $X = \sum_{k=1}^n X_k$, on a :

$$\mathbf{E}[X] = \mathbf{E}\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n \mathbf{E}[X_i] = \sum_{i=1}^n p = np,$$

et, X_1, \dots, X_n étant de plus indépendantes :

$$\mathbf{V}[X] = \mathbf{V}\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n \mathbf{V}[X_i] = \sum_{i=1}^n p(1-p) = np(1-p).$$

□

Exemple 10.5 (suite). Reprenons l'Exemple 10.5 et supposons que $p = 0,4$. On a alors :

$$\begin{aligned} \mathbf{P}[3 \leq X \leq 5] &= \mathbf{P}[X = 3] + \mathbf{P}[X = 4] + \mathbf{P}[X = 5] \\ &= \binom{10}{3} 0,4^3 (1-0,4)^7 + \binom{10}{4} 0,4^4 (1-0,4)^6 + \binom{10}{5} 0,4^5 (1-0,4)^5 \\ &\simeq 0,6665. \end{aligned}$$

On a 66,65% de chances d'observer entre 3 et 5 faces lors d'une série de 10 lancers. L'espérance de la variable X est $\mathbf{E}[X] = 10 \times 0,4 = 4$: si on répète une grand nombre de fois l'expérience, on s'attend à observer en moyenne 4 faces par série de 10 lancers.

Loi de Poisson

Lorsque le nombre d'épreuves n devient très important, la manipulation de la loi binomiale devient fastidieuse voire impossible. On peut alors remplacer son utilisation par celle de la *loi de Poisson* sous certaines conditions (Voir Théorème 10.1). Celle-ci évalue le nombre aléatoire d'événements (rares) de même probabilité pendant une durée donnée comme, par exemple, le nombre d'appels reçus par un standard téléphonique en une heure, le nombre de voitures se présentant à un péage dans une journée, ...

Définition 10.10. On dit qu'une v.a. X suit une loi de Poisson de paramètre $\lambda > 0$, si son support est $X(\Omega) = \mathbf{N}$ et, pour tout $k \in \mathbf{N}$:

$$\mathbf{P}[X = k] = e^{-\lambda} \frac{\lambda^k}{k!}.$$

On note alors $X \sim \mathcal{P}(\lambda)$.

Remarque 10.7. On a clairement que pour tout $k \in \mathbf{N}$:

$$P[X = k + 1] = \frac{\lambda}{k + 1} \mathbf{P}[X = k].$$

Proposition 10.6 (Admise). Si $X \sim \mathcal{P}(\lambda)$, on a :

$$\mathbf{E}[X] = \mathbf{V}[X] = \lambda.$$

Exemple 10.6. Si, en moyenne, 10 voitures se présentent à un péage donné en une heure, on modélise le nombre de voiture se présentant au péage en une heure par une v.a. de loi $\mathcal{P}(10)$. Cette modélisation sera justifiée dans la suite.

Proposition 10.7. Soient X_1 et X_2 deux v.a. indépendantes de lois de Poisson de paramètres respectifs λ_1 et λ_2 .

On a :

$$X_1 + X_2 \sim \mathcal{P}(\lambda_1 + \lambda_2).$$

Preuve : Pour tout $k \in \mathbf{N}$, on a :

$$\begin{aligned} \mathbf{P}[X_1 + X_2 = k] &= \sum_{i=0}^k \mathbf{P}[X_1 = i, X_2 = k - i] \\ &= \sum_{i=0}^k \mathbf{P}[X_1 = i] P[X_2 = k - i] && (X_1 \text{ et } X_2 \text{ indépendantes}) \\ &= \sum_{i=0}^k e^{-\lambda_1} \frac{\lambda_1^i}{i!} e^{-\lambda_2} \frac{\lambda_2^{k-i}}{(k-i)!} \\ &= e^{-(\lambda_1 + \lambda_2)} \sum_{i=0}^k \frac{\lambda_1^i}{i!} \frac{\lambda_2^{k-i}}{(k-i)!} \\ &= e^{-(\lambda_1 + \lambda_2)} \frac{1}{k!} \sum_{i=0}^k \frac{k!}{i!(k-i)!} \lambda_1^i \lambda_2^{k-i} \\ &= e^{-(\lambda_1 + \lambda_2)} \frac{1}{k!} \sum_{i=0}^k \binom{k}{i} \lambda_1^i \lambda_2^{k-i} \\ &= e^{-(\lambda_1 + \lambda_2)} \frac{(\lambda_1 + \lambda_2)^k}{k!} \end{aligned}$$

où l'on a utilisé la formule du binôme de Newton dans la dernière égalité. □

Approximation binomiale/Poisson Dans cette sous-section nous allons justifier le fait, évoqué ci-dessus, que l'on peut approcher certaines lois binomiales par une loi de Poisson. Pour cela nous commençons par définir une notion de convergence pour les suites de variables aléatoires.

Définition 10.11. Soit $(X_n)_{n \in \mathbf{N}}$ une suite de variables aléatoires réelles dont les fonctions de répartition sont notées F_n et Y une variable aléatoire de fonction de répartition F .

On dit que $(X_n)_{n \in \mathbf{N}}$ converge en loi vers Y , si, en tout point de continuité x de F , on a :

$$\lim_{n \rightarrow +\infty} F_n(x) = F(x).$$

On note alors :

$$X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} Y.$$

Proposition 10.8 (Admise). Soit $(X_n)_{n \in \mathbf{N}}$ une suite de variables aléatoires discrètes et Y une variable aléatoire discrète.

Alors, $(X_n)_{n \in \mathbf{N}}$ converge en loi vers Y , si, et seulement si :

$$\lim_{n \rightarrow +\infty} \mathbf{P}[X_n = x] = \mathbf{P}[Y = x], \quad \text{pour tout } x \in Y(\Omega).$$

Théorème 10.1 (Admis). Soit $\lambda > 0$, $X_n \sim \mathcal{Bin}(n, \frac{\lambda}{n})$ et $Y \sim \mathcal{P}(\lambda)$.

On a :

$$X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} Y.$$

Remarque 10.8. On peut donc approcher la loi $\mathcal{Bin}(n, p)$ par la loi $\mathcal{P}(n \times p)$. Dans la pratique, on ne fait une telle approximation que si p est proche de 0, $n \geq 30$, $p \leq 0,1$, $np \leq 10$, sans quoi l'approximation est mauvaise.

Exemple 10.6 (suite). Justifions l'utilisation de la loi de Poisson $\mathcal{P}(10)$ dans l'Exemple 10.6. Supposons que l'on observe s'il y a eu une arrivée de voiture seulement à n instants fixés dans l'heure (n est voué à être grand). Alors, la variable X_n comptant le nombre d'instants où l'on a observé une arrivée de voiture suit une loi binomiale $\mathcal{Bin}(n, p)$. Puisqu'en moyenne on observe 10 arrivées de voitures en une heure, on a $np = 10$, soit $p = \frac{10}{n}$. Plus n est grand, plus la discrétisation du temps est fine et s'approche de la réalité. L'idée est donc de faire tendre n vers l'infini. Le Théorème 10.1 affirme que la loi limite est la loi $\mathcal{P}(10)$.

Loi géométrique

Considérons une expérience de Bernoulli dont la probabilité de succès est $p \in]0; 1[$. Si l'on répète plusieurs fois indépendamment cette expérience, on peut s'intéresser au nombre aléatoire de répétition de l'expérience nécessaires pour obtenir un premier succès.

Définition 10.12. Considérons une expérience de Bernoulli dont la probabilité de succès est $p \in]0; 1[$ que l'on répète jusqu'au premier succès. On appelle loi géométrique de paramètre $p \in]0; 1[$ la loi du rang du premier succès. Cette loi est notée $\mathcal{G}(p)$.

Proposition 10.9. Soit $X \sim \mathcal{G}(p)$.

1. Le support de X est $X(\Omega) = \mathbf{N}^*$ et pour tout $k \in \mathbf{N}^*$, on a :

$$\mathbf{P}[X = k] = (1 - p)^{k-1}p.$$

2. On a :

$$\mathbf{E}[X] = \frac{1}{p} \quad \text{et} \quad \mathbf{V}[X] = \frac{1-p}{p^2}.$$

Preuve : Exercice! □

Exemple 10.7. Une personne rentre ivre chez elle et prélève au hasard une clef dans son trousseau, en contenant 5, pour tenter d'ouvrir la porte. Si elle échoue, elle remet la clef dans son trousseau et recommence. Le nombre de tentatives X jusqu'à l'ouverture de la porte suit alors une loi géométrique de paramètre $\frac{1}{5} = 0,2$. En moyenne, la porte sera ouverte après 5 tentatives et la probabilité qu'elle soit ouverte après k tentatives est :

$$\mathbf{P}[X = k] = 0,2 \times 0,8^{k-1}.$$

10.3 Variables aléatoires continues

10.3.1 Loi d'une variable aléatoire continue et densité

Le point délicat pour la définition de la loi d'une v.a. continue X est que la probabilité $\mathbf{P}[X = x]$ pour que X prenne la valeur x est identiquement nulle. Ainsi, il est impossible de définir la loi d'une v.a. continue par la donnée de ses probabilités élémentaires. On définit donc la loi d'une variable aléatoire par sa fonction de répartition :

$$F_X(x) = \mathbf{P}[X \leq x], \quad x \in \mathbf{R},$$

qui la caractérise.

On a les propriétés suivantes.

Proposition 10.10. *Soit X une v.a. continue et F_X sa fonction de répartition.*

1. *La fonction F_X est croissante et continue et vérifie :*

$$\lim_{x \rightarrow -\infty} F_X(x) = 0 \quad \text{et} \quad \lim_{x \rightarrow +\infty} F_X(x) = 1.$$

2. *Pour tous $a, b \in \mathbf{R}$, on a :*

$$\mathbf{P}[a < X < b] = \mathbf{P}[a < X \leq b] = \mathbf{P}[a \leq X < b] = \mathbf{P}[a \leq X \leq b] = F_X(b) - F_X(a).$$

Définition 10.13. *On dit qu'une v.a. continue admet pour densité f_X si F_X est dérivable (sauf éventuellement en un nombre fini de points) et sa dérivée est f_X .*

Remarque 10.9. La plupart des lois continues usuelles sont, en fait, définies par leurs densités.

Quelques mots sur l'intégration

Si f est une fonction continue sur un intervalle $[a; b]$, on appelle « intégrale de f entre a et b » l'aire (algébrique) sous la courbe représentative de f entre a et b (voir Figure 10.1). Cette quantité est notée :

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

Remarque 10.10. Ainsi, si X est une v.a. réelle admettant f_X pour densité, sa fonction de répartition est donnée par :

$$F(x) = \mathbf{P}[X \leq x] = \int_{-\infty}^x f_X(t) \, dt.$$

Une construction rigoureuse de l'intégrale est hors de la portée de ce cours. On peut toutefois énoncer un résultat permettant de calculer des intégrales de fonctions continues.

Définition 10.14. *Soit f une fonction continue sur un intervalle I . On dit que f admet pour primitive F si, F est dérivable sur I et pour tout $x \in I$, $F'(x) = f(x)$.*

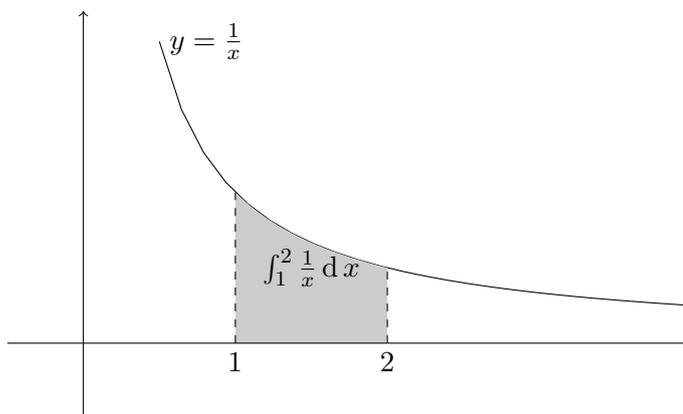


FIGURE 10.1 – Illustration de la notion d'intégrale.

Théorème 10.2 (Théorème fondamentale de l'analyse). *Soit f une fonction admettant F pour primitive sur I . Pour tous $a, b \in I$, on a :*

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

Exemple 10.8. Soient $f : x \mapsto \frac{1}{x}$ et $F : x \mapsto \ln(x)$. On a pour tout $x \in \mathbf{R}_+^*$, $F'(x) = f(x)$ et F est une primitive de f sur \mathbf{R}_+^* . Ainsi, l'aire grisée sur la Figure 10.1 vaut :

$$\int_1^2 \frac{1}{x} \, dx = \int_1^2 f(x) \, dx = F(2) - F(1) = \ln(2) - \ln(1) = \ln(2).$$

Remarque 10.11. Dans la pratique, pour les lois usuelles des tables existent et permettent de se passer des calculs d'intégrales. On observe que, si X est une v.a. de densité f_X , on a :

$$\mathbf{P}[a < X < b] = \mathbf{P}[a < X \leq b] = \mathbf{P}[a \leq X < b] = \mathbf{P}[a \leq X \leq b] = F_X(b) - F_X(a) = \int_a^b f_X(x) \, dx$$

et

$$\int_{-\infty}^{+\infty} f_X(x) \, dx = 1.$$

Proposition 10.11 (Linéarité de l'intégrale). *Soient f et g deux fonctions intégrables sur $[a; b]$ et $\alpha, \beta \in \mathbf{R}$.*

Alors,

$$\int_a^b (\alpha f(x) + \beta g(x)) \, dx = \alpha \int_a^b f(x) \, dx + \beta \int_a^b g(x) \, dx.$$

10.3.2 Moments d'une variable aléatoire continue

Définition 10.15. *Soit X une variable aléatoire réelle admettant f_X pour densité.*

1. Si $\int_{-\infty}^{+\infty} |x| f_X(x) \, dx < \infty$, l'espérance de X est la quantité :

$$\mathbf{E}[X] = \int_{-\infty}^{+\infty} x f_X(x) \, dx.$$

2. Si $\int_{-\infty}^{+\infty} |x|^k f_X(x) dx < \infty$, on appelle moment d'ordre k de X la quantité :

$$\mathbf{E}[X^k].$$

3. Si $\int_{-\infty}^{+\infty} |x|^2 f_X(x) dx < \infty$, on appelle variance de X , la quantité

$$\mathbf{V}[X] = \mathbf{E}[(X - \mathbf{E}[X])^2] = \mathbf{E}[X^2] - \mathbf{E}[X]^2.$$

L'écart-type de X , est alors $\sigma = \sqrt{\mathbf{V}[X]}$.

10.3.3 Lois à densités usuelles

Loi uniforme continue

Définition 10.16. Soient $a < b$ deux réels.

On dit qu'une v.a. X suit la loi uniforme sur $[a; b]$ si sa densité est donnée par :

$$f_X(x) = \frac{1}{b-a} \mathbf{1}_{[a;b]}(x), \quad x \in \mathbf{R}.$$

On note alors $X \sim \mathcal{U}([a; b])$.

Remarque 10.12. Il s'agit de l'analogue continue de la loi uniforme discrète.



FIGURE 10.2 – Densité f_X de la loi uniforme sur $[1; 4]$.

Proposition 10.12. Soit $X \sim \mathcal{U}([a; b])$.

1. La fonction de répartition F de X est donnée par :

$$F_X(x) = \begin{cases} 0 & \text{si } x \leq a \\ \frac{x-a}{b-a} & \text{si } a \leq x \leq b \\ 1 & \text{si } x \geq b \end{cases} .$$

En particulier, pour tout intervalle I , on a :

$$\mathbf{P}[X \in I] = \frac{|I \cap [a; b]|}{|[a; b]|}$$

où $|J|$ désigne la longueur d'un intervalle J .

2. On a :

$$\mathbf{E}[X] = \frac{a+b}{2} \quad \text{et} \quad \mathbf{V}[X] = \frac{(b-a)^2}{12}.$$

Preuve :

1. Il est clair que si $x \leq a$, $F_X(x) = 0$ et si $x \geq b$, $F_X(x) = 1$. Soit $x \in [a; b]$. On a :

$$\begin{aligned} F_X(x) &= \mathbf{P}[X \leq x] = \int_{-\infty}^x f_X(t) dt = \int_{-\infty}^x \frac{1}{b-a} \mathbf{1}_{[a;b]}(t) dt \\ &= \frac{1}{b-a} \int_a^x dt = \frac{1}{b-a} [t]_a^x = \frac{x-a}{b-a}. \end{aligned}$$

2. On a :

$$\begin{aligned} \mathbf{E}[X] &= \int_{-\infty}^{+\infty} x f_X(x) dx = \int_{-\infty}^{+\infty} x \frac{1}{b-a} \mathbf{1}_{[a;b]}(x) dx \\ &= \frac{1}{b-a} \int_a^b x dx = \frac{1}{b-a} \left[\frac{x^2}{2} \right]_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2} \end{aligned}$$

et

$$\begin{aligned} \mathbf{E}[X^2] &= \int_{-\infty}^{+\infty} x^2 f_X(x) dx = \int_{-\infty}^{+\infty} x^2 \frac{1}{b-a} \mathbf{1}_{[a;b]}(x) dx \\ &= \frac{1}{b-a} \int_a^b x^2 dx = \frac{1}{b-a} \left[\frac{x^3}{3} \right]_a^b = \frac{b^3 - a^3}{3(b-a)} = \frac{a^2 + ab + b^2}{3}. \end{aligned}$$

On en déduit que :

$$\begin{aligned} \mathbf{V}[X] &= \mathbf{E}[X^2] - \mathbf{E}[X]^2 = \frac{a^2 + ab + b^2}{3} - \left(\frac{a+b}{2} \right)^2 \\ &= \frac{a^2 + ab + b^2}{3} - \frac{a^2 + 2ab + b^2}{4} = \frac{4a^2 + 4ab + 4b^2 - 3a^2 - 6ab - 3b^2}{12} \\ &= \frac{(b-a)^2}{12}. \end{aligned}$$

□

Lois exponentielles

Définition 10.17. Soit $\lambda > 0$.

On dit qu'une v.a. X suit la loi exponentielle de paramètre λ si sa densité est donnée par :

$$f_X(x) = \lambda e^{-\lambda x} \mathbf{1}_{[0;+\infty[}(x), \quad x \in \mathbf{R}.$$

On note alors $X \sim \mathcal{E}(\lambda)$.

Remarque 10.13. Elle est utilisée pour modéliser des phénomènes *sans mémoire* ou *sans vieillissement* tels que le temps d'attente avant le prochain tremblement de terre ou la prochaine désintégration dans un réacteur nucléaire ou encore la durée de vie de certains appareils comme des ampoules. Ceci est justifié par le deuxième point de la Proposition 10.16.

Proposition 10.13. Soit $X \sim \mathcal{E}(\lambda)$, $\lambda > 0$.

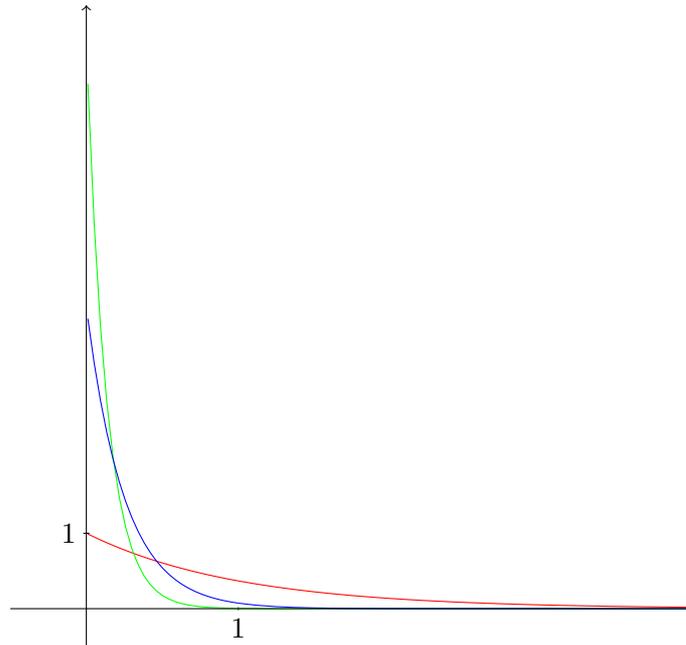


FIGURE 10.3 – Densités des lois exponentielles de paramètre 1 (rouge), 4 (bleu) et 7,5 (vert).

1. La fonction de répartition F_X de X est donnée par :

$$F_X(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ 1 - e^{-\lambda x} & \text{si } x \geq 0 \end{cases} .$$

2. [Perte de mémoire] Pour tous $s, t \geq 0$:

$$\mathbf{P}[X > s + t | X > t] = \mathbf{P}[X > s].$$

3. On a :

$$\mathbf{E}[X] = \frac{1}{\lambda} \quad \text{et} \quad \mathbf{V}[X] = \frac{1}{\lambda^2}.$$

Preuve :

1. Il est clair que si $x \leq 0$, $F_X(x) = 0$. Soit $x \geq 0$. On a :

$$\begin{aligned} F_X(x) &= \mathbf{P}[X \leq x] = \int_{-\infty}^x f_X(t) \, dt = \int_{-\infty}^x \lambda e^{-\lambda t} \mathbf{1}_{[0; +\infty[}(t) \, dt \\ &= \int_0^x \lambda e^{-\lambda t} \, dt = \left[-\lambda e^{-\lambda t} \right]_0^x = 1 - \lambda e^{-\lambda x}. \end{aligned}$$

2. On a :

$$\begin{aligned} \mathbf{P}[X > s + t | X > t] &= \frac{\mathbf{P}[X > s + t, X > t]}{\mathbf{P}[X > t]} = \frac{\mathbf{P}[X > s + t]}{\mathbf{P}[X > t]} \\ &= \frac{1 - \mathbf{P}[X \leq s + t]}{1 - \mathbf{P}[X \leq t]} = \frac{1 - F_X(s + t)}{1 - F_X(t)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} \\ &= e^{-\lambda s} = 1 - F_X(s) \mathbf{P}[X > s]. \end{aligned}$$

3. Admis.

□

Exemple 10.9. On a observé que la durée de vie d'une ampoule d'un modèle donné est d'en moyenne 1000 heures. Considérons une ampoule de ce modèle et intéressons nous à sa durée de vie X (exprimée en heures). La v.a. X est continue et sans mémoire. On considère donc que X suit une loi exponentielle. Puisque l'on s'attend à avoir une durée de vie moyenne de 1000 heures, le paramètre de cette loi exponentielle est $\lambda = \frac{1}{1000}$ de sorte que $\mathbf{E}[X] = \frac{1}{\lambda} = 1000$.

Ainsi, la probabilité pour que l'ampoule fonctionne au plus 100h est :

$$\mathbf{P}[X \leq 100] = 1 - e^{-\frac{1}{1000} \times 100} = 1 - e^{-\frac{1}{10}} \simeq 0,01.$$

De même, la probabilité pour que l'ampoule fonctionne plus de 4500 heures est :

$$\mathbf{P}[X > 4500] = 1 - \mathbf{P}[X \leq 4500] = e^{-\frac{1}{1000} \times 4500} = 1 - e^{-4,5} \simeq 0,01.$$

Loi normale centrée réduite $\mathcal{N}(0; 1)$

Définition 10.18. On dit qu'une v.a. X suit la loi normale (ou gaussienne) centrée réduite si sa densité est donnée par :

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in \mathbf{R}^d.$$

On note alors $X \sim \mathcal{N}(0; 1)$.

Remarque 10.14. La loi normale centrée réduite et, plus généralement les lois normales, sont parmi les lois les mieux adaptées pour décrire des phénomènes provenant de plusieurs événements aléatoires. Elles sont apparues dans les écrits de Moivre (1733) comme limites de lois binomiales. Il s'agit de l'origine du Théorème Central Limite (TCL), un théorème limite des probabilités absolument fondamental (voir Section 10.4). En 1777, Laplace a obtenu une bonne approximation de l'erreur entre cette loi normale et la loi binomiale. Quelques années plus tard, en 1809, Gauss assimile des erreurs d'observation en astronomie à la courbe, dite des erreurs, de la densité de la loi normale. Ceci explique pourquoi elle est également nommée loi de Gauss ou de Laplace-Gauss.

Proposition 10.14. Soit $X \sim \mathcal{N}(0; 1)$.

1. La fonction densité f_X de X est paire i.e. $f_X(-x) = f_X(x)$ pour tout réel x . En particulier, on a :

$$F_X(-x) = 1 - F_X(x).$$

2.

$$\mathbf{E}[X] = 0 \quad \text{et} \quad \mathbf{V}[X] = 1.$$

Preuve :

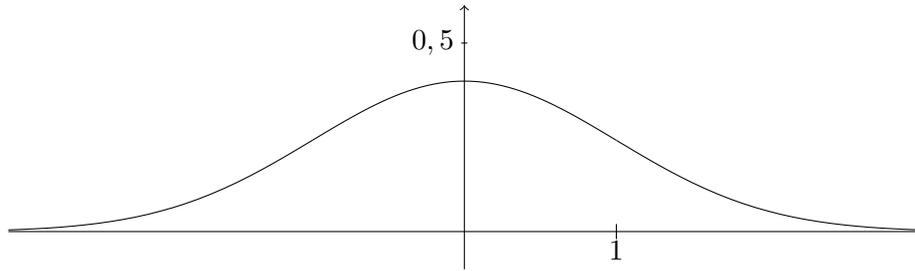


FIGURE 10.4 – Le graphe de la densité de la loi $\mathcal{N}(0;1)$ est la célèbre *courbe en cloche* de Gauss. On peut observer sur ce graphique sa symétrie par rapport à l’axe des ordonnées.

1. La parité de la densité est claire. On en déduit que :

$$\begin{aligned}
 F_X(-x) &= \int_{-\infty}^{-x} f_X(t) dt \\
 &= \int_{-\infty}^{+\infty} f_X(t) dt - \int_{-x}^{+\infty} f_X(t) dt = 1 - \int_{-x}^{+\infty} f_X(t) dt \\
 &= \int_{-\infty}^{+\infty} f_X(t) dt && \text{par parité de } f_X \\
 &= 1 - F_X(x).
 \end{aligned}$$

2. Admis. □

Remarque 10.15. Il n’existe pas d’expression de la fonction de répartition de la loi normale centrée réduite sous forme close autre que :

$$F_X(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt, \quad x \in \mathbf{R}.$$

On a donc recours à des tables telles que celle présentée en Annexe A. L’utilisation de cette table est détaillée dans cette même annexe.

Lois normales générales $\mathcal{N}(m; \sigma^2)$

Définition 10.19. On dit qu’une v.a. X suit la loi normale (ou gaussienne) de moyenne m et de variance σ^2 si sa densité est donnée par :

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}, \quad x \in \mathbf{R}.$$

On note alors $X \sim \mathcal{N}(m; \sigma^2)$.

Le changement de variable $y = \frac{x-m}{\sigma}$ permet de voir que l’on peut ramener les calculs de probabilités pour la loi $\mathcal{N}(m; \sigma^2)$ à des calculs de probabilités pour la loi $\mathcal{N}(0; 1)$. On taira, ici, les détails techniques et on utilisera le résultat suivant.

Proposition 10.15. Soit $X \sim \mathcal{N}(m; \sigma^2)$.

1.

$$\mathbf{E}[X] = m \quad \text{et} \quad \mathbf{V}[X] = \sigma^2.$$

2.

$$Z := \frac{X - m}{\sigma} \sim \mathcal{N}(0; 1).$$

Remarque 10.16. Le deuxième point de cette proposition est fondamental puisqu'il explicite la manière de ramener les calculs de probabilités pour la loi $\mathcal{N}(m; \sigma^2)$ à des calculs de probabilités pour la loi $\mathcal{N}(0; 1)$.

Exemple 10.10. Soit $X \sim \mathcal{N}(60; 36)$.

Déterminons $\mathbf{P}[X < 45]$ puis $\mathbf{P}[45 < X < 50]$.

On utilise que, d'après la proposition précédente, la v.a. $Z := \frac{X-60}{6}$ suit une loi normale centrée réduite.

On a :

$$\begin{aligned} X < 45 &\iff X - 60 < -15 \\ &\iff Z = \frac{X - 60}{6} < -2,5 \end{aligned}$$

donc

$$\begin{aligned} \mathbf{P}[X < 45] &= \mathbf{P}[Z < -2,5] = 1 - \mathbf{P}[Z \leq 2,5] \\ &\simeq 1 - 0,9938 = 0,0062. \end{aligned}$$

On a :

$$\begin{aligned} 45 < X < 50 &\iff -15 < X - 60 < -10 \\ &\iff -2,5 < Z = \frac{X - 60}{6} < -\frac{5}{3}. \end{aligned}$$

donc

$$\begin{aligned} \mathbf{P}[45 < X < 50] &= \mathbf{P}\left[-2,5 < Z < -\frac{5}{3}\right] = \mathbf{P}\left[Z < -\frac{5}{3}\right] - \mathbf{P}[Z \leq -2,5] \\ &\simeq \mathbf{P}\left[Z \leq -\frac{5}{3}\right] - 0,0062 = 1 - \mathbf{P}\left[Z \leq \frac{5}{3}\right] - 0,0062 \\ &= 0,9938 - \mathbf{P}\left[Z \leq \frac{5}{3}\right]. \end{aligned}$$

La valeur de $\mathbf{P}\left[Z \leq \frac{5}{3}\right]$ n'est pas donnée dans la table de la fonction de répartition de la loi normale centrée réduite et nous devons utiliser la méthode de l'interpolation linéaire pour en trouver une valeur approchée α précise.

On a :

$$1,66 < \frac{5}{3} \simeq 1,666 < 1,67$$

donc

$$\mathbf{P}[Z \leq 1,66] < \mathbf{P}\left[Z \leq \frac{5}{3}\right] < \mathbf{P}[Z \leq 1,67]$$

donc

$$0,9515 < \alpha < 0,9525.$$

On en déduit que :

$$\frac{\alpha - 0,9515}{0,9525 - 0,9515} = \frac{\frac{5}{3} - 1,66}{1,67 - 1,66}$$

d'où

$$\alpha = \frac{\frac{5}{3} - 1,66}{0,01} \times 0,001 + 0,9515 \simeq 0,9522$$

et

$$\mathbf{P}[45 < X < 50] \simeq 0,9938 - 0,9522 = 0,0416.$$

Proposition 10.16. Soient $\alpha, \beta \in \mathbf{R}$, $X_1 \sim \mathcal{N}(m_1; \sigma_1^2)$ et $X_2 \sim \mathcal{N}(m_2; \sigma_2^2)$.

Si X_1 et X_2 sont **indépendantes**, alors :

$$\alpha X_1 + \beta X_2 \sim \mathcal{N}(\alpha m_1 + \beta m_2; \alpha^2 \sigma_1^2 + \beta^2 \sigma_2^2).$$

On peut également s'intéresser à la concentration de la masse autour de la moyenne pour des v.a. normales. Plus précisément on a :

Proposition 10.17. Soit $X \sim \mathcal{N}(m; \sigma^2)$.

On a :

$$\mathbf{P}[m - \sigma \leq X \leq m + \sigma] \simeq 0,6827,$$

$$\mathbf{P}[m - 2\sigma \leq X \leq m + 2\sigma] \simeq 0,9545$$

et

$$\mathbf{P}[m - 3\sigma \leq X \leq m + 3\sigma] \simeq 0,9973.$$

10.4 Complément : quelques mots sur les théorèmes limite en probabilité

10.4.1 Loi forte des grands nombres

Théorème 10.3 (Loi forte des grands nombres). Soit $(X_n)_{n \in \mathbf{N}^*}$ une suite de v.a. indépendantes et identiquement distribuées (i.i.d.) définies sur le même espace probabilisé $(\Omega; \mathbf{P})$ et telles que $\mathbf{E}[|X_1|] < +\infty$.

Alors, il existe $\Omega_1 \subset \Omega$, avec $\mathbf{P}[\Omega_1] = 1$, tel que sur Ω_1 :

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{i=1}^n X_i = \mathbf{E}[X_1].$$

Remarque 10.17. Ce résultat justifie que si l'on considère des v.a. $(X_i)_{1 \leq i \leq n}$ représentant les résultats de n expériences identiques mais indépendantes, on s'attend à ce que la moyenne empirique observée soit proche de l'espérance $\mathbf{E}[X_1]$ d'une seule de ces variables.

10.4.2 Théorème Central Limite

Théorème 10.4 (Théorème Central Limite (TCL)). *Soit $(X_n)_{n \in \mathbf{N}^*}$ une suite de v.a. i.i.d. définies sur le même espace probabilisé. Supposons $m = \mathbf{E}[X_1]$ existe et soit fini et que $\sigma^2 = \mathbf{V}[X_1]$ existe et soit fini et non nulle.*

Alors,

$$\frac{\sum_{i=1}^n X_i - nm}{\sqrt{n\sigma^2}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} Z$$

où $Z \sim \mathcal{N}(0; 1)$.

Remarque 10.18.

1. Ce théorème est fondamental pour déterminer des intervalles de confiance et justifie le fait que les lois normales sont adaptées pour modéliser des phénomènes aléatoires issus d'un grand nombre de répétitions indépendantes d'événements aléatoires.
2. On a :

$$\mathbf{E} \left[\sum_{i=1}^n X_i \right] = nm \quad \text{et} \quad \mathbf{V} \left[\sum_{i=1}^n X_i \right] = n\sigma^2.$$

Théorème 10.5 (Théorème de Moivre-Laplace). *Soit $(B_n)_{n \in \mathbf{N}^*}$ une suite de v.a. indépendantes avec $B_n \sim \mathcal{Bin}(n; p)$, $p \in]0; 1[$.*

Alors,

$$\frac{B_n - np}{\sqrt{np(1-p)}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} Z$$

où $Z \sim \mathcal{N}(0; 1)$.

Preuve : Il suffit d'écrire B_n comme somme de v.a. i.i.d. de loi de Bernoulli de paramètre p et d'appliquer le TCL. □

Remarque 10.19. Dans la pratique, on considère que l'approximation fournie par ce théorème est bonne si $n \geq 30$, $p \geq 0, 1$ et $np > 15$.

10.4.3 Applications en statistique inférentielle

Afin de réduire les tailles des populations étudiées en statistique, les statistiques *inférentielles* s'appuient sur les probabilités (et en particulier sur les théorèmes limite) pour estimer les paramètres de séries statistiques dont il serait trop coûteux d'explorer toutes les valeurs. C'est, par exemple, le principe des sondages : au lieu d'interroger tous les membres d'une population, on n'interroge qu'un *échantillon* de la population que l'on espère représentatif. On parle alors d'échantillonnage.

Plus formellement, on suppose qu'un caractère que l'on souhaite étudié dans une population de taille N (très grand) est tel que la valeur observée pour chaque individu suit une certaine loi de probabilité \mathcal{L} , commune à tous les individus, mais inconnue. On cherche alors à en estimer les paramètres. On peut, bien entendu avoir un *a priori* sur la loi en question et, par exemple, s'attendre à ce qu'il s'agisse d'une loi normale dont les paramètres ne sont pas connus. Pour estimer ces paramètres, on dispose de $n \ll N$ observations X_1, \dots, X_n que l'on suppose toutes de loi \mathcal{L} et indépendantes. On parle d'échantillon de taille n .

Estimations ponctuelles

Supposons que l'on souhaite estimer la « vraie » moyenne m et la « vraie » variance σ^2 de la loi de la v.a. X étudiée. Si l'on avait accès aux observations X_1, \dots, X_N pour toute la population, on écrirait simplement que :

$$m = \frac{1}{N} \sum_{i=1}^N X_i \quad \text{et} \quad \sigma^2 = \frac{1}{N} \sum_{i=1}^N (X_i - m)^2.$$

Puisque l'on n'a accès qu'à un sous-échantillon X_1, \dots, X_n de la population (de taille très inférieure à N) ceci est impossible. On estime donc la « vraie » moyenne m par la *moyenne empirique* :

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$$

et la « vraie » variance σ^2 par la *variance empirique corrigée* :

$$\bar{S}_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2.$$

Ces quantités sont des variables aléatoires que l'on appelle des *estimateurs*. On note respectivement \bar{X} et \bar{S}^2 des réalisations de celles-ci que l'on appelle des *estimations ponctuelles*. Ces estimateurs sont ceux utilisés car ils ont de bonnes qualités. En particulier, ils sont *sans biais*, c'est-à-dire que :

$$\mathbf{E} [\bar{X}_n] = m \quad \text{et} \quad \mathbf{E} [\bar{S}_n^2] = \sigma^2.$$

Estimation par intervalles de confiances

Le principe Supposons que l'on dispose d'une estimation ponctuelle $\bar{\theta}$ d'un paramètre θ . On souhaite déterminer un intervalle centré en $\bar{\theta}$ c'est-à-dire de la forme :

$$I_{1-\alpha} = [\bar{\theta} - l_\alpha; \bar{\theta} + l_\alpha]$$

tel que :

$$\mathbf{P} [\theta \in I_{1-\alpha}] = 1 - \alpha, \quad \alpha \in]0; 1[.$$

Le réel $1 - \alpha$ est la *confiance* accordée à cette estimation et α est appelé le *risque*. Ceci s'interprète de la façon suivante : si la vraie valeur de θ ne se trouve pas dans $I_{1-\alpha}$, la probabilité d'avoir observé $\bar{\theta}$ n'est que de α . Typiquement, on prend $\alpha = 0,1$, $\alpha = 0,05$ ou $\alpha = 0,01$. Lorsque le risque diminue, la longueur de l'intervalle de confiance augmente. Il n'y a donc pas un intervalle de confiance meilleur que les autres. Il faut trouver un compromis.

Intervalle de confiance pour la moyenne avec variance connue Supposons que l'on dispose d'un échantillon X_1, \dots, X_n , avec $n \geq 30$, d'observations indépendantes suivant une loi commune de moyenne m inconnue et de variance σ^2 connue. Le TCL (Théorème 10.4) assure que l'on peut utiliser l'approximation en loi suivante :

$$Z := \frac{\sum_{i=1}^n X_i - nm}{\sqrt{n\sigma^2}} = \frac{\bar{X}_n - m}{\sigma/\sqrt{n}} \sim \mathcal{N}(0; 1).$$

On cherche un intervalle de la forme $I_{1-\alpha} = [\bar{X} - l_\alpha; \bar{X} + l_\alpha]$ tel que :

$$\mathbf{P}[m \in I_{1-\alpha}] = 1 - \alpha.$$

On a :

$$\begin{aligned} \mathbf{P}[m \in I_{1-\alpha}] = 1 - \alpha &\iff \mathbf{P}\left[\bar{X}_n - l_\alpha \leq m \leq \bar{X}_n + l_\alpha\right] = 1 - \alpha \\ &\iff \mathbf{P}\left[\bar{X}_n - l_\alpha \geq m \text{ ou } m \geq \bar{X}_n + l_\alpha\right] = \alpha \\ &\iff \mathbf{P}\left[\bar{X}_n - m \geq l_\alpha \text{ ou } \bar{X}_n - m \leq -l_\alpha\right] = \alpha \\ &\iff \mathbf{P}\left[\frac{\bar{X}_n - m}{\sigma/\sqrt{n}} \geq \frac{\sqrt{n}}{\sigma}l_\alpha \text{ ou } \frac{\bar{X}_n - m}{\sigma/\sqrt{n}} \leq -\frac{\sqrt{n}}{\sigma}l_\alpha\right] = \alpha \\ &\iff \mathbf{P}\left[Z \geq \frac{\sqrt{n}}{\sigma}l_\alpha \text{ ou } Z \leq -\frac{\sqrt{n}}{\sigma}l_\alpha\right] = \alpha. \end{aligned}$$

Puisque Z suit la loi $\mathcal{N}(0; 1)$, on peut trouver un réel t_α tel que :

$$\mathbf{P}[Z \geq t_\alpha \text{ ou } Z \leq -t_\alpha] = \alpha.$$

Remarque 10.20. Dans la pratique, il existe une table (dite de d'écart-réduit) pour déterminer t_α . Les valeurs les plus fréquemment utilisées sont :

$$t_{0,1} \simeq 1,598, \quad t_{0,05} \simeq 1,960 \quad \text{et} \quad t_{0,01} \simeq 2,576.$$

Par identification, on obtient que :

$$\frac{\sqrt{n}}{\sigma}l_\alpha = t_\alpha,$$

soit encore

$$l_\alpha = t_\alpha \frac{\sigma}{\sqrt{n}}.$$

On en déduit que l'intervalle de confiance $I_{1-\alpha}$ cherché s'écrit sous la forme :

$$I_{1-\alpha} = \left[\bar{X} - t_\alpha \frac{\sigma}{\sqrt{n}}; \bar{X} + t_\alpha \frac{\sigma}{\sqrt{n}} \right].$$

On vient donc de montrer que :

Proposition 10.18 (Intervalle de confiance pour la moyenne avec variance connue). *Si l'on dispose d'un échantillon X_1, \dots, X_n , avec $n \geq 30$, d'observations indépendantes suivant une loi commune de moyenne m inconnue et de variance σ^2 connue. L'estimation par intervalle de confiance de m au risque α est donnée par :*

$$I_{1-\alpha} = \left[\bar{X} - t_\alpha \frac{\sigma}{\sqrt{n}}; \bar{X} + t_\alpha \frac{\sigma}{\sqrt{n}} \right],$$

où t_α est tel que pour une v.a. normale centrée réduite on ait :

$$\mathbf{P}[Z \geq t_\alpha \text{ ou } Z \leq -t_\alpha] = \alpha.$$

Remarque 10.21. On peut également construire des intervalles de confiance pour la moyenne avec variance inconnue, pour la variance avec moyenne connue ou non, ... Ceci sera développé dans le cours de deuxième année. Lorsque les deux paramètres sont inconnus, la loi normale centrée réduite est remplacée par certaines de lois dérivées de celle-ci (les *lois de Student*). Le raisonnement est pour autant tout à fait analogue.

Intervalle de confiance pour une proportion Supposons que l'on souhaite estimer, dans une population de taille N , la proportion p d'individus vérifiant une certaine condition. On peut, par exemple s'intéresser à la proportion de personne favorable à une mesure, souhaitant voter pour un candidat lors d'une élection, connaissant l'existence d'un produit, ... On suppose que l'on a, pour cela, interrogé $n \geq 30$ individus, que les réponses sont indépendantes et sont codées par 1 si l'individu remplit la condition et 0 sinon. Ainsi, on dispose de n v.a. de Bernoulli de paramètre p , indépendantes X_1, \dots, X_n . La quantité

$$Y_n = \sum_{i=1}^n X_i$$

est le nombre de personne dans l'échantillon satisfaisant la condition et suit la loi $\mathcal{B}in(n; p)$. Il est naturel d'estimer la « vraie » proportion p d'individus satisfaisant la condition par :

$$\bar{p}_n = \frac{Y_n}{n}.$$

Le Théorème de Moivre-Laplace (Théorème 10.5) assure que l'on peut utiliser l'approximation en loi suivante :

$$\frac{Y_n - np}{\sqrt{np(1-p)}} \sim \mathcal{N}(0; 1).$$

On en déduit que l'on peut utiliser l'approximation suivante :

$$Z := \frac{\bar{p}_n - p}{\sqrt{\bar{p}_n(1-\bar{p}_n)/n}} \sim \mathcal{N}(0; 1).$$

On cherche un intervalle de la forme $I_{1-\alpha} = [\bar{p}_n - l_\alpha; \bar{p}_n + l_\alpha]$ tel que :

$$\mathbf{P}[p \in I_{1-\alpha}] = 1 - \alpha.$$

On a :

$$\begin{aligned} \mathbf{P}[p \in I_{1-\alpha}] = 1 - \alpha &\iff \mathbf{P}[\bar{p}_n - l_\alpha \leq p \leq \bar{p}_n + l_\alpha] = 1 - \alpha \\ &\iff \mathbf{P}[\bar{p}_n - l_\alpha \geq p \text{ ou } p \geq \bar{p}_n + l_\alpha] = \alpha \\ &\iff \mathbf{P}[\bar{p}_n - p \geq l_\alpha \text{ ou } \bar{p}_n - p \leq -l_\alpha] = \alpha \\ &\iff \mathbf{P}\left[\frac{\bar{p}_n - p}{\sqrt{\bar{p}_n(1-\bar{p}_n)/n}} \geq \sqrt{\frac{n}{\bar{p}_n(1-\bar{p}_n)}} l_\alpha \text{ ou } \frac{\bar{p}_n - p}{\sqrt{\bar{p}_n(1-\bar{p}_n)/n}} \leq -\sqrt{\frac{n}{\bar{p}_n(1-\bar{p}_n)}} l_\alpha\right] = \alpha \\ &\iff \mathbf{P}\left[Z \geq \sqrt{\frac{n}{\bar{p}_n(1-\bar{p}_n)}} l_\alpha \text{ ou } Z \leq -\sqrt{\frac{n}{\bar{p}_n(1-\bar{p}_n)}} l_\alpha\right] = \alpha. \end{aligned}$$

Puisque Z suit la loi $\mathcal{N}(0; 1)$, on peut trouver un réel t_α tel que :

$$\mathbf{P}[Z \geq t_\alpha \text{ ou } Z \leq -t_\alpha] = \alpha.$$

Par identification, on obtient que :

$$\sqrt{\frac{n}{\bar{p}_n(1-\bar{p}_n)}} l_\alpha = t_\alpha,$$

soit encore

$$l_\alpha = \sqrt{\frac{\bar{p}_n(1 - \bar{p}_n)}{n}} t_\alpha.$$

On en déduit que l'intervalle de confiance $I_{1-\alpha}$ cherché s'écrit sous la forme :

$$I_{1-\alpha} = \left[\bar{p} - \sqrt{\frac{\bar{p}(1 - \bar{p})}{n}} t_\alpha; \bar{p} + \sqrt{\frac{\bar{p}(1 - \bar{p})}{n}} t_\alpha \right].$$

On vient donc de montrer que :

Proposition 10.19 (Intervalle de confiance pour une proportion). *L'estimation par intervalle de confiance pour une proportion p dans un échantillon de taille $n \geq 30$ au risque α est donnée par :*

$$I_{1-\alpha} = \left[\bar{p} - \sqrt{\frac{\bar{p}(1 - \bar{p})}{n}} t_\alpha; \bar{p} + \sqrt{\frac{\bar{p}(1 - \bar{p})}{n}} t_\alpha \right],$$

où t_α est tel que pour une v.a. normale centrée réduite on ait :

$$\mathbf{P}[Z \geq t_\alpha \text{ ou } Z \leq -t_\alpha] = \alpha.$$

Exercice 10.1. À l'approche d'un référendum, on a effectué un sondage pour connaître la proportion p de personnes projetant de voter « oui ». On a dans ce cadre obtenu 110 réponses « oui » et 90 réponses « non ».

1. Déterminer une estimation ponctuelle de p .
2. Déterminer une estimation de p par intervalle de confiance au risque de 10%, 5% puis 1%.
3. Commenter.

Annexe A

Table de la fonction de répartition de la loi normale centrée réduite

La table donnée deux pages plus loin contient les valeurs arrondies à 4 décimales près de la fonction de répartition de la loi $\mathcal{N}(0; 1)$

$$F(x) = \mathbf{P}[X \leq x] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$$

pour $x \in \{0; 0,01; \dots; 3,99\}$.

A.1 Lecture dans la table

A.1.1 Lecture directe

Pour toute valeur de $x \in \{0; 0,01; \dots; 3,99\}$, la valeur de $F(x)$ s'obtient par simple lecture dans la table. Plus précisément, si $x = l + c$ où l est l'entête d'une ligne et c l'entête d'une colonne, la valeur de $F(x)$ se trouve dans la case d'intersection de la ligne l et la colonne c .

Par exemple, $F(0,45) = \mathbf{P}[X \leq 0,45]$ se trouve dans la case d'intersection de la ligne repérée par 0,4 et la colonne repérée par 0,05. On lit donc dans la table que :

$$F(0,45) \simeq 0,6736.$$

A.1.2 Cas des valeurs non indiquées dans la table

Cas des x négatifs

Si $x \in \{-3,99; -3,98; \dots; -0,01\}$, on utilise que par symétrie de la loi $\mathcal{N}(0; 1)$:

$$F(x) = 1 - F(-x).$$

Par exemple, pour déterminer $F(-0,52)$, on écrit que :

$$F(-0,52) = 1 - F(0,52).$$

Or, par lecture directe dans la table, on a que $F(0,52) \simeq 0,6985$. On en déduit que :

$$F(-0,52) = 1 - F(0,52) \simeq 1 - 0,6985 = 0,3015.$$

Cas général

Si $x \leq -3.99$, on conclue que $F(x) \simeq 0$ et si $x \geq 3.99$, on conclue que $F(x) \simeq 1$. Finalement, pour $x \in [-3,99; 3,99] \setminus \{-3,99; -3,98; \dots; 3,99\}$, on commence par encadrer le plus précisément possible x par deux nombres dans $\{-3,99; -3,98; \dots; 3,99\}$ puis on réalise une interpolation linéaire.

Par exemple, pour déterminer $F(2/3)$, on écrit que :

$$0,66 \leq \frac{2}{3} \leq 0,67$$

et donc

$$F(0,66) \leq F\left(\frac{2}{3}\right) \leq F(0,67).$$

Or, par lecture directe dans la table, on a que $F(0,66) \simeq 0,7454$ et $F(0,67) \simeq 0,7486$. En utilisant la méthode d'interpolation linéaire (voir Chapitre 2), on en déduit que :

$$\begin{aligned} \frac{F\left(\frac{2}{3}\right) - F(0,66)}{F(0,67) - F(0,66)} &= \frac{\frac{2}{3} - 0,66}{0,67 - 0,66} \\ &= 100 \left(\frac{2}{3} - 0,66\right) \end{aligned}$$

puis que

$$\begin{aligned} F\left(\frac{2}{3}\right) &= 100 \left(\frac{2}{3} - 0,66\right) (F(0,67) - F(0,66)) + F(0,66) \\ &\simeq 100 \left(\frac{2}{3} - 0,66\right) (0,7486 - 0,7454) + 0,7454 \\ &\simeq 0,7475. \end{aligned}$$

ANNEXE A. TABLE DE LA FONCTION DE RÉPARTITION DE LA LOI NORMALE CENTRÉE RÉDUITE

Table de la fonction de répartition de la loi normale centrée réduite $\mathcal{N}(0; 1)$.

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000