IUT de Dijon-Auxerre
BUT GEA 1re année
Semestre 2
Année universitaire 2022-2023

R2.07 Outils Mathématiques de gestion

Notes et compléments de cours

Arnaud Rousselle
arnaud.rousselle@iut-dijon.u-bourgogne.fr

Préambule

Objectifs, contenus et structure :

L'objectif de ce cours est de consolider et étoffer lesoutils mathématiques pour la gestion. Dans la lignée du cours du premier semestre, il se veut donc être tourné vers les applications et la pratique, et il met en évidence les liens avec des problèmes concrets de gestion – rencontrés dans d'autres cours, SAÉ ou lors de stage –, tout en permettant une compréhension réelle et fine des concepts présentés. Pour ces raisons, les premières sections de chaque chapitre énoncent précisément les définitions, concepts et résultats clés et les illustrent par des exemples et exercices permettant de mettre en lumière les méthodes principales ; leur lecture est donc essentielle et les TD et TP permettront leurs mises en œuvre. Afin de rendre cette lecture la plus fluide possible, elles ne contiennent que peu de preuves de résultats théoriques, seulement les plus courtes. Pour donner un panorama plus complet et des pistes de réflexion utiles lors de certains projets, des sections de compléments sont intégrées et clairement repérées de certains chapitres.

Le premier chapitre s'inscrit dans la continuité du Chapitre 4 du cours du premier semestre. Il présente les outils d'étude des fonctions à une variable réelle (limites, continuité, dérivation, variations, optimisation) et donne une ouverture vers leurs applications en économie.

Le deuxième chapitre est, quant à lui, un prolongement du Chapitre 1 du cours du premier semestre. Il permet d'appréhender les aspects des statistiques lorsque plusieurs caractères sont étudiés sur une même population. On y présente les notions de corrélation, de régression et de distance du χ^2 allant vers les tests d'hypothèses du même nom qui pourront également être rencontrés en marketing ou méthodes d'enquêtes. Une introduction à l'inférence et à la prédiction y est faite.

Dans la continuité du deuxième chapitre, le troisième chapitre présente des méthodes d'analyse et de prédiction pour des séries chronologiques évoluant dans le temps selon une tendance générale et des effets périodiques. Il permet, entre autre, d'acquérir les compétences nécessaires pour la préparation d'un budget prévisionnel ou un plan d'investissement en connaissant les différentes recettes et dépenses des années précédentes.

Évaluations :

À préciser.

Avertissement :

Ces notes et compléments de cours ayant été fraîchement rédigés, quelques coquilles peuvent encore y figurer. Le fait de les signaler par courriel sera apprécié.

Table des matières

Chapitre 1

Analyse à une variable réelle

Le Chapitre 4 du cours du premier semestre est un prérequis indispensable au présent chapitre. Ce lui-ci le complète en donnant les outils d'étude des fonctions à une variable réelle (limites, continuité, dérivation, variations, optimisation) et une ouverture vers leurs applications en économie.

1.1 Limites

1.1.1 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 1.1. *On dit que f* tend *vers $+\infty$ (resp. $-\infty$) en $+\infty$ si pour tout $A > 0$, il existe $B > 0$ tel que :*

$$x \geq B \Longrightarrow f(x) \geq A \quad (\text{resp.} f(x) \leq -A).$$

On note alors :

$$\lim_{x\to+\infty} f(x) = +\infty \quad (\text{resp.} \lim_{x\to+\infty} f(x) = -\infty).$$

On définit de manière analogue les limites infinies en $-\infty$.

Définition 1.2. *On dit que f* tend *vers $+\infty$ (resp. $-\infty$) en $-\infty$ si pour tout $A > 0$, il existe $B > 0$ tel que :*

$$x \leq -B \Longrightarrow f(x) \geq A \quad (\text{resp.} f(x) \leq -A).$$

On note alors :

$$\lim_{x\to-\infty} f(x) = +\infty \quad (\text{resp.} \lim_{x\to-\infty} f(x) = -\infty).$$

Exemple 1.1. Vérifions que $\lim_{x\to-\infty} x^4 = +\infty$. Soit $A > 0$. Posons $B = A^{\frac{1}{4}}$. On a :

$$\begin{aligned} x \leq -B &\Longrightarrow x \leq -A^{\frac{1}{4}} \\ &\Longrightarrow x \geq A \end{aligned}$$

Ainsi,

$$\lim_{x\to-\infty} x^4 = +\infty.$$

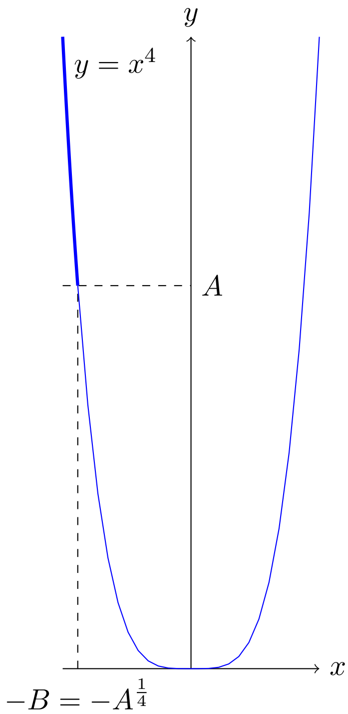

FIGURE 1.1 – Illustration du fait que $\lim_{x\to-\infty} x^4 = +\infty$. Pour tout $A > 0$, pour tout $x \leq -B = -A^{\frac{1}{4}}$, on a $x^4 \geq A$.

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.

Définition 1.3. *On dit que f admet pour* limite l *en $+\infty$ si pour tout $\varepsilon > 0$, il existe $A > 0$ tel que :*

$$x \geq A \Longrightarrow |f(x) - l| \leq \varepsilon.$$

On note alors :

$$\lim_{x\to+\infty} f(x) = l.$$

On définit de manière analogue les limites infinies en $-\infty$.

Définition 1.4. *On dit que f admet pour* limite l *en* $-\infty$ *si pour tout* $\varepsilon > 0$*, il existe* $A > 0$ *tel que :*

$$x \leq -A \Longrightarrow |f(x) - l| \leq \varepsilon.$$

On note alors :

$$\lim_{x \to -\infty} f(x) = l.$$

Exemple 1.2. Vérifions que $\lim_{x\to+\infty} \frac{1}{\sqrt{x}} = 0$. Soit $\varepsilon > 0$. Posons $A = \varepsilon^{-2}$. On a :

$$\begin{aligned} x \geq A &\Longrightarrow x \geq \varepsilon^{-2} \\ &\Longrightarrow \sqrt{x} \geq \varepsilon^{-1} \\ &\Longrightarrow \left| \frac{1}{\sqrt{x}} - 0 \right| = \frac{1}{\sqrt{x}} \leq \varepsilon. \end{aligned}$$

Ainsi,

$$\lim_{x \to +\infty} \frac{1}{\sqrt{x}} = 0.$$

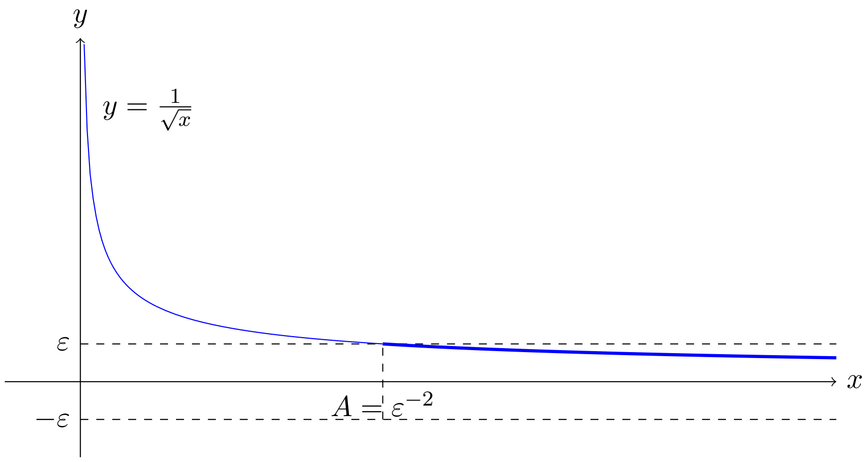

FIGURE 1.2 – Illustration du fait que $\lim_{x\to+\infty} \frac{1}{\sqrt{x}} = 0$. Pour tout $\varepsilon > 0$, pour tout $x \geq A = \varepsilon^{-2}$, on a $\left|\frac{1}{\sqrt{x}} - 0\right| = \frac{1}{\sqrt{x}} \leq \varepsilon$.

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 1.5. *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 :*

$$|x - a| \leq \delta \Longrightarrow f(x) \geq A \quad (resp. f(x) \leq -A).$$

On note alors :

$$\lim_{x \to a} f(x) = +\infty \quad (resp. \lim_{x \to a} f(x) = -\infty).$$

Exemple 1.3. Vérifions que $\lim_{x\to 1} \frac{1}{x^2-2x+1} = +\infty$. Soit $A > 0$. Posons $\varepsilon = \frac{1}{\sqrt{A}}$. On a :

$$|x - 1| \leq \varepsilon = \frac{1}{\sqrt{A}} \Longrightarrow 0 \leq (x-1)^2 = x^2 - 2x + 1| \leq \frac{1}{A}$$
$$\Longrightarrow \frac{1}{x^2 - 2x + 1} \geq A$$

Ainsi,

$$\lim_{x \to 1} \frac{1}{x^2 - 2x + 1} = +\infty.$$

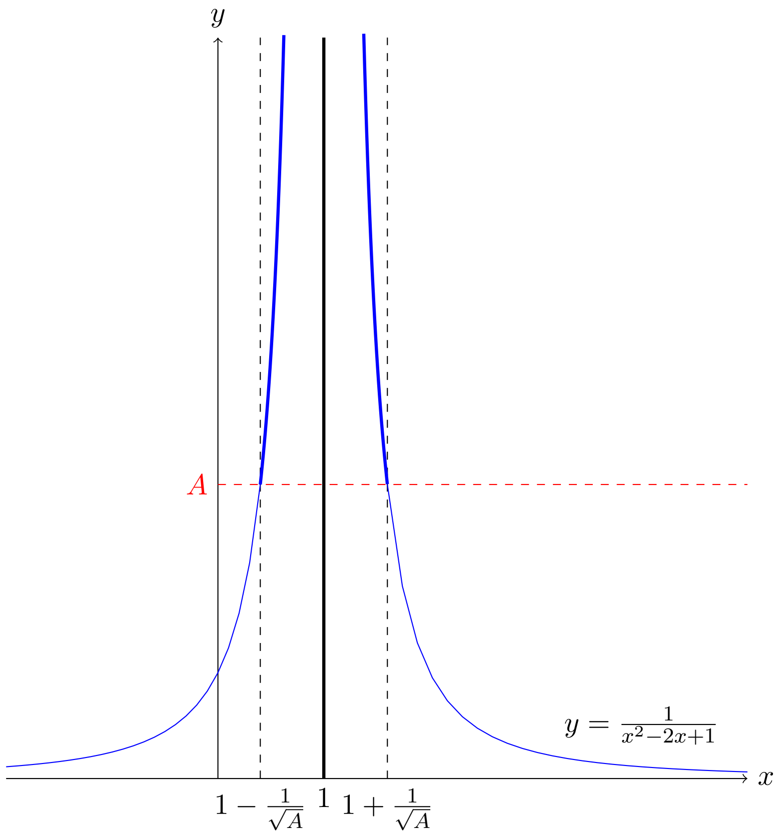

FIGURE 1.3 – Illustration du fait que $\lim_{x\to 1} \frac{1}{x^2-2x+1} = +\infty$. Pour tout $A > 0$, pour tout x vérifiant $|x - 1| \leq \frac{1}{\sqrt{A}}$, on a $\frac{1}{x^2-2x+1} \geq A$.

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 1.6. *On dit que f admet pour* limite *l en a si pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que :*

$$|x - a| \leq \delta \Longrightarrow |f(x) - l| \leq \varepsilon.$$

On note alors :

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

Exercice 1.1. Vérifier, en utilisant la définition, que $\lim_{x\to 3} \frac{-x^2+6x-9}{(x-5)(x-3)^2} = \frac{1}{2}$.

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. Ce constat conduit à la définition des limites à droite et à gauche de fonction en un point.

Définition 1.7. *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 :*

$$a < x \leq a + \delta \Longrightarrow |f(x) - l_+| \leq \varepsilon \quad \textit{(resp. } a - \delta \leq x < a \Longrightarrow |f(x) - l_-| \leq \varepsilon\textit{)}.$$

On note alors :

$$\lim_{\substack{x \to a \\ x > a}} f(x) = l_+ \quad \textit{(resp. } \lim_{\substack{x \to a \\ x > a}} f(x) = l_-\textit{)}.$$

Définition 1.8. *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 :*

$$a < x \leq a + \delta \Longrightarrow f(x) \geq A \quad \textit{(resp. } a - \delta \leq x < a \Longrightarrow f(x) \geq A\textit{)}.$$

On note alors :

$$\lim_{\substack{x \to a \\ x > a}} f(x) = +\infty \quad \textit{(resp. } \lim_{\substack{x \to a \\ x > a}} f(x) = +\infty\textit{)}.$$

Définition 1.9. *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 :*

$$a < x \leq a + \delta \Longrightarrow f(x) \leq -A \quad \textit{(resp. } a - \delta \leq x < a \Longrightarrow f(x) \leq -A\textit{)}.$$

On note alors :

$$\lim_{\substack{x \to a \\ x > a}} f(x) = -\infty \quad \textit{(resp. } \lim_{\substack{x \to a \\ x > a}} f(x) = -\infty\textit{)}.$$

Exemple 1.4. On peut voir que :

$$\lim_{\substack{x \to 0 \\ x > 0}} \frac{1}{x} = +\infty \quad \text{et} \quad \lim_{\substack{x \to 0 \\ x < 0}} \frac{1}{x} = -\infty$$

1.1.2 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 :

$$\lim_{x\to+\infty} x^n = +\infty, \quad \lim_{x\to-\infty} x^n = \begin{cases} +\infty & \text{si } n \text{ est pair} \\ -\infty & \text{si } n \text{ est impair} \end{cases}, \quad \lim_{x\to\pm\infty} \frac{1}{x^n} = 0,$$

$$\lim_{x\to 0} \frac{1}{x^n} = +\infty \quad \text{si } n \text{ est pair,}$$

$$\lim_{\substack{x\to 0\\ x>0}} \frac{1}{x^n} = +\infty \text{ et } \lim_{\substack{x\to 0\\ x<0}} \frac{1}{x^n} = -\infty \quad \text{si } n \text{ est impair.}$$

Puissances arbitraires Si $\alpha > 0$, alors

$$\lim_{x\to+\infty} x^\alpha = +\infty.$$

Si $\alpha < 0$, alors

$$\lim_{x\to+\infty} x^\alpha = 0 \quad \text{et} \quad \lim_{\substack{x\to 0\\ x>0}} x^\alpha = +\infty.$$

Exponentielles et logarithmes

$$\lim_{x\to-\infty} \exp(x) = 0, \quad \lim_{x\to+\infty} \exp(x) = +\infty,$$

$$\lim_{\substack{x\to 0\\ x>0}} \ln(x) = -\infty \quad \text{et} \quad \lim_{x\to+\infty} \ln(x) = +\infty.$$

1.1.3 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 1.1. *Soit $a, b \in \mathbf{R} \cup \{-\infty; +\infty\}$.*

1. *Si $\alpha \in \mathbf{R}$ et $\lim_{x\to a} f(x) = l \in \mathbf{R}$, alors $\lim_{x\to a} \alpha f(x) = \alpha l$.*
2. *Si $\alpha > 0$ et $\lim_{x\to a} f(x) = +\infty$ (resp. $-\infty$), alors $\lim_{x\to a} \alpha f(x) = +\infty$ (resp. $-\infty$).*
3. *Si $\alpha < 0$ et $\lim_{x\to a} f(x) = +\infty$ (resp. $-\infty$), alors $\lim_{x\to a} \alpha f(x) = -\infty$ (resp. $+\infty$).*
4. *Si $\lim_{x\to a} f(x) = l_1 \in \mathbf{R} \cup \{-\infty; +\infty\}$ et $\lim_{x\to a} g(x) = l_2 \in \mathbf{R} \cup \{-\infty; +\infty\}$, alors $\lim_{x\to a} f(x) + g(x) = l_1 + l_2$ avec $+\infty + \infty = +\infty$, $-\infty - \infty = -\infty$ et, pour $l \in \mathbf{R}$, $l + \infty = +\infty$ $l - \infty = -\infty$.*
5. *Si $\lim_{x\to a} f(x) = l_1 \in \mathbf{R}$ et $\lim_{x\to a} g(x) = l_2 \in \mathbf{R}$, alors $\lim_{x\to a} f(x)g(x) = l_1 l_2$.*
6. *Si $\lim_{x\to a} f(x) > 0$ et $\lim_{x\to a} g(x) = +\infty$ (resp. $-\infty$), alors $\lim_{x\to a} f(x)g(x) = +\infty$ (resp. $-\infty$).*
7. *Si $\lim_{x\to a} f(x) < 0$ et $\lim_{x\to a} g(x) = +\infty$ (resp. $-\infty$), alors $\lim_{x\to a} f(x)g(x) = -\infty$ (resp. $+\infty$).*

8. *Si* $\lim_{x\to a} f(x) = l_1 \in \mathbf{R}$ *et* $\lim_{x\to a} g(x) = l_2 \in \mathbf{R}^*$*, alors* $\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{l_1}{l_2}$*.*
9. *Si* $\lim_{x\to a} f(x) = l \in \mathbf{R}$ *et* $\lim_{x\to a} g(x) = \pm\infty$*, alors* $\lim_{x\to a} \frac{f(x)}{g(x)} = 0$*.*
10. *Si* $\lim_{x\to a} f(x) > 0$*,* $\lim_{x\to a} g(x) = 0$ *et g est positive au voisinage de a (resp. négative au voisinage de a), alors* $\lim_{x\to a} \frac{f(x)}{g(x)} = +\infty$ *(resp. $-\infty$).*
11. *Si* $\lim_{x\to a} f(x) < 0$*,* $\lim_{x\to a} g(x) = 0$ *et g est positive au voisinage de a (resp. négative au voisinage de a), alors* $\lim_{x\to a} \frac{f(x)}{g(x)} = -\infty$ *(resp. $+\infty$).*
12. *Si* $\lim_{x\to a} f(x) = b \in \mathbf{R} \cup \{-\infty; +\infty\}$ *et* $\lim_{x\to b} g(x) = l \in \mathbf{R} \cup \{-\infty; +\infty\}$*, alors* $\lim_{x\to a} (g \circ f)(x) = \lim_{x\to a} g(f(x)) = l$*.*

Remarque 1.1. On peut observer que, par exemple, on ne peut pas déterminer la limite de $f + g$ dans le point 4. lorsque $\lim_{x\to a} f(x) = -\infty$ et $\lim_{x\to a} g(x) = +\infty$. 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 1.5. 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 :

$$\lim_{x\to+\infty} x \ln(a) = -\infty \quad \text{si } 0 < a < 1$$

et

$$\lim_{x\to+\infty} x \ln(a) = +\infty \quad \text{si } a > 1.$$

Or, $\lim_{y\to-\infty} \exp(y) = 0$ et $\lim_{y\to+\infty} \exp(y) = +\infty$. On en déduit, par composition, que :

$$\lim_{x\to+\infty} \exp_a(x) = \lim_{x\to+\infty} \exp(x \ln(a)) = 0 \quad \text{si } 0 < a < 1$$

et

$$\lim_{x\to+\infty} \exp_a(x) = \lim_{x\to+\infty} \exp(x \ln(a)) = +\infty \quad \text{si } a > 1.$$

La limite de la fonction $x \longmapsto \exp(x) - x^2$ en $+\infty$ est une forme indéterminée. Nous verrons dans la Section 1.1.5 comment lever de telles indéterminations.

Exercice 1.2. Déterminer, pour $a > 0$ la limite en $+\infty$ de la fonction logarithme à base a.

1.1.4 Comparaison de limites

Le théorème de comparaison de limites suivant se déduit directement des définitions.

Théorème 1.1. *Soit $a \in \mathbf{R} \cup \{-\infty; +\infty\}$.*

1. *Si $f(x) \le g(x)$ dans un voisinage de a et* $\lim_{x\to a} f(x) = +\infty$*, alors* $\lim_{x\to a} g(x) = +\infty$*.*
2. *Si $f(x) \le g(x)$ dans un voisinage de a et* $\lim_{x\to a} g(x) = -\infty$*, alors* $\lim_{x\to a} f(x) = -\infty$*.*
3. *Si $f(x) \le g(x) \le h(x)$ dans un voisinage de a et* $\lim_{x\to a} f(x) = \lim_{x\to a} h(x) = l \in \mathbf{R}$*, alors* $\lim_{x\to a} g(x) = l$*.*

Exemple 1.6. 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 $\lim_{x\to+\infty} -\frac{1}{x} = \lim_{x\to+\infty} \frac{1}{x} = 0$ que :

$$\lim_{x\to+\infty} \frac{\cos(x)}{x} = 0.$$

1.1.5 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 1.2 (Admise)**.** *On a :*

$$\lim_{x\to+\infty} \frac{\exp(x)}{x} = +\infty, \qquad \lim_{x\to-\infty} x\exp(x) = 0,$$

$$\lim_{x\to+\infty} \frac{\ln(x)}{x} = 0 \quad \textit{et} \quad \lim_{\substack{x\to 0\\ x>0}} x\ln(x) = 0.$$

Plus généralement, on a, pour tout $\alpha > 0$ et tout $\beta \in \mathbf{R}$:

$$\lim_{x\to+\infty} \frac{\exp(\alpha x)}{x^\beta} = +\infty, \qquad \lim_{x\to-\infty} |x|^\beta \exp(\alpha x) = 0,$$

$$\lim_{x\to+\infty} \frac{(\ln(x))^\beta}{x^\alpha} = 0 \quad \textit{et} \quad \lim_{\substack{x\to 0\\ x>0}} x^\alpha |\ln(x)|^\beta = 0.$$

Remarque 1.2. 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 1.7. Déterminons la limite en $+\infty$ de la fonction $x \longmapsto \exp(x) - x^3$. Le terme dominant est $\exp(x)$. On a donc :

$$\lim_{x\to+\infty} \exp(x) - x^3 = \lim_{x\to+\infty} \exp(x)\left(1 - \frac{x^3}{\exp(x)}\right).$$

Or, $\lim_{x\to+\infty} \frac{x^3}{\exp(x)} = 0$ donc $\lim_{x\to+\infty} 1 - \frac{x^3}{\exp(x)} = 1$ et $\lim_{x\to+\infty} \exp(x) = +\infty$. On en déduit que :

$$\lim_{x\to+\infty} \exp(x) - x^3 = \lim_{x\to+\infty} \exp(x)\left(1 - \frac{x^3}{\exp(x)}\right) = +\infty.$$

Cette méthode permet d'affirmer que la limite en $\pm\infty$ de tout polynôme est celle de son monôme de plus haut degré.

Proposition 1.3. *Soit $a_n \in \mathbf{R}^*$ et $a_{n-1}, \ldots, a_0 \in \mathbf{R}$.*
On a :

$$\lim_{x\to\pm\infty} a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = \lim_{x\to\pm\infty} a_n x^n.$$

Preuve : On a :

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = a_n x^n \left(1 + \frac{a_{n-1}}{a_n x} + \cdots + \frac{a_1}{a_n x^{n-1}} + \frac{a_0}{a_n x^n}\right).$$

Or, chacun des termes $\frac{a_{n-1}}{a_n x}, \ldots, \frac{a_0}{a_n x^n}$ 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 1.8. Déterminons la limite en $-\infty$ de la fonction $x \longmapsto \frac{3x^3-2x+7}{2x^2+5x-1}$. Au numérateur, le terme dominant est celui du 3^{e} degré alors qu'au dénominateur, il s'agit de celui du 2^{nd} 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 :

$$\begin{aligned}\lim_{x\to-\infty} \frac{3x^3 - 2x + 7}{2x^2 + 5x - 1} &= \lim_{x\to-\infty}\left(x\frac{3 - \frac{2}{x^2} + \frac{7}{x^3}}{2 + \frac{5}{x} - \frac{1}{x^2}}\right) \\ &= \lim_{x\to-\infty} x \times \lim_{x\to-\infty} \frac{3 - \frac{2}{x^2} + \frac{7}{x^3}}{2 + \frac{5}{x} - \frac{1}{x^2}} = -\infty \times \frac{3}{2} = -\infty.\end{aligned}$$

Remarque 1.3. 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 1.9. Déterminons la limite de $x \longmapsto \exp(-2x)\left(2x^5 + 5x^2 - 9\right)$ lorsque x tend vers $+\infty$. On a :

$$\begin{aligned}
\lim_{x\to+\infty} \exp(-2x)\left(2x^5 + 5x^2 - 9\right) &= \lim_{x\to+\infty} \frac{1}{\exp(2x)}\left(2x^5 + 5x^2 - 9\right)\\
&= \lim_{x\to+\infty} \frac{2x^5 + 5x^2 - 9}{\exp(2x)}\\
&= \lim_{x\to+\infty} \frac{x^5}{\exp(2x)}\left(2 + \frac{5}{x^3} - \frac{9}{x^5}\right)\\
&= 2\lim_{x\to+\infty} \frac{x^5}{\exp(2x)}\\
&= 0.
\end{aligned}$$

1.2 Continuité

Définition 1.10. *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 :*

$$|x - x_0| \leq \delta \Longrightarrow |f(x) - f(x_0)| \leq \varepsilon.$$

On dit que f est continue *sur $I \subset \mathbf{R}$ si elle est continue en tout point de I.*

Remarque 1.4. 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 1.10. 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 1.4. *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 1.2 (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 1.3 (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 1.3. Vérifier que l'équation $\exp(2x+1) = 2$ admet une unique solution dans l'intervalle $[-1;3]$.

1.3 Dérivabilité

1.3.1 Définitions

Définition 1.11. *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 1.12. *Soit f une fonction continue au voisinage de x_0.*

Si, $\lim_{h\to 0} (\tau_f(x_0))\,(h)$ existe et est finie, on appelle nombre dérivé de f en x_0 *la quantité :*

$$f'(x_0) = \lim_{h\to 0} (\tau_f(x_0))\,(h).$$

On dit dans ce cas que f est dérivable en x_0.

Définition 1.13. *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 1.5. 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 :

$$\frac{\mathrm{d}}{\mathrm{d}\,s} f(s)$$

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 1.11. 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 :

$$C : Y \longmapsto C(Y) = C_0 + c(Y - T) = C_0 + c(Y - tY) = C_0 + c(1-t)Y.$$

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 :

$$\frac{\mathrm{d}}{\mathrm{d}\,Y} C(Y) = C'(Y) = c(1-t) \geq 0.$$

On en déduit que la consommation du ménage croit avec son revenu ; plus précisément, on en déduit que si le revenu du ménage croit d'une unité, sa consommation croit 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 1.3.4). Un 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.

1.3.2 Dérivées usuelles

Le tableau suivant récapitule les dérivées usuelles et domaines de dérivabilité des fonctions.

Fonction $f : x \mapsto f(x) = \dots$	Domaine de définition	Domaine de dérivabilité	Dérivée $f' : x \mapsto f'(x) =$
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^α, $\alpha > 0$	$\mathbf{R}_+$	$\mathbf{R}_+^*$	$\alpha x^{\alpha-1}$
x^α, $\alpha < 0$	$\mathbf{R}_+^*$	$\mathbf{R}_+^*$	$\alpha x^{\alpha-1}$
$\exp(x)$	$\mathbf{R}$	$\mathbf{R}$	$\exp(x)$
$\ln(x)$	$\mathbf{R}_+^*$	$\mathbf{R}_+^*$	$\frac{1}{x}$
$\ln\lvert x\rvert$	$\mathbf{R}^*$	$\mathbf{R}^*$	$\frac{1}{x}$

1.3.3 Opérations sur les dérivées

Proposition 1.5 (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 :*

$$(u+v)'(x) = u'(x) + v'(x) \quad \textit{et} \quad (u \times v)'(x) = u'(x)v(x) + u(x)v'(x).$$

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,

$$(\exp \circ u)'(x) = (\exp(u(x)))' = u'(x) \exp(u(x)),$$

$$((u(x))^n)' = nu'(x)(u(x))^{n-1}, \quad n \in \mathbf{N}^*,$$

et si de plus $u(x) > 0$,

$$(\ln \circ u)'(x) = (\ln(u(x)))' = \frac{u'(x)}{u(x)}.$$

Exemple 1.12.

1. La fonction $f : x \longmapsto x^2 \exp(x)$ 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 $g : x \longmapsto \frac{x^3-1}{x-2}$ 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 $h : x \longmapsto \ln(x^2 - 1)$ 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 :

$$h'(x) = u'(x)v'(u(x)) = 2x \times \frac{1}{x^2-1} = \frac{2x}{x^2-1}.$$

1.3.4 Tangente à une courbe en un point

Définition 1.14. *Soit f une fonction dérivable en x_0.*

La tangente à la courbe représentative de f en x_0 *est la droite d'équation :*

$$\mathrm{T}_{f,x_0} : y = f'(x_0)\,(x - x_0) + f(x_0).$$

Remarque 1.6. Il s'agit de la meilleure approximation locale de la fonction f par une fonction affine au voisinage de x_0.

Exemple 1.13. Reprenons la fonction $g : x \longmapsto \frac{x^3-1}{x-2}$ de l'Exemple 1.12. 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 1.4) est donc :

$$\mathrm{T}_{g,0} : y = g'(0)\,(x - 0) + g(x_0) = \frac{1}{4}x + \frac{1}{2}.$$

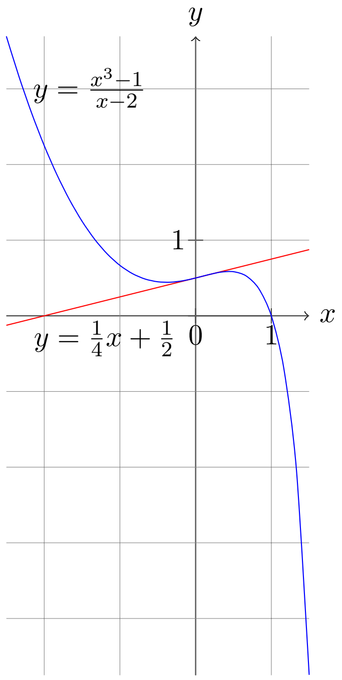

FIGURE 1.4 – Courbe représentative de la fonction $g : x \longmapsto \frac{x^3-1}{x-2}$ (bleu) et sa tangente en 0, d'équation $y = \frac{1}{4}x + \frac{1}{2}$ (rouge). Cette dernière est la meilleure approximation locale de la fonction g par une fonction affine au voisinage de 0.

1.3.5 Applications à l'étude des variations et à la recherche d'extrema

Définition 1.15. *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 1.16. *Soit f une fonction et D_f son ensemble de définition.*

1. *On dit qu'une fonction f admet un* maximum local *(rexp.* 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 *(rexp.* 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 *(rexp.* 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 1.7.

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 1.6. *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 1.8. 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 1.7. *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 1.9. 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 1.5.

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 1.14. 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 le nombre de produits 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 :

$$f(x) = 85x - (5x^2 + 5x + 100) = -5x^2 + 80x - 100, \qquad x \geq 0.$$

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 :

$$f'(x) = -10x + 80.$$

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 :

$$f(0) = -100 \quad \text{et} \quad \lim_{x \to +\infty} f(x) = -\infty.$$

Le tableau de variation de f prend la forme suivante :

x	0		8		∞
$f'(x)$		$+$	0	$-$	
f			220		
		$\nearrow$		$\searrow$	
	0				$-\infty$

1.4 Applications en économie

1.4.1 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 1.15. 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$^{\text{e}}$ unité est $CT(51) - CT(50) \simeq 0,995$. Il est (convenablement) approché par son coût marginal :

$$Cm(50) = CT'(50) = \frac{100}{\sqrt{200 \times 50}} = 1.$$

1.4.2 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.

1.4.3 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* $\mathrm{Rm} = f'(x)$.

1.4.4 Élasticité

Définition 1.17. *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 1.10. En utilisant la définition de la dérivée, on peut réécrire l'élasticité de f en x sous la forme :

$$\mathcal{E}_f(x) = \lim_{h\to 0} \frac{\frac{f(x+h)-f(x)}{f(x)}}{\frac{h}{x}}.$$

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 1.16. 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 $x = 20$€ 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 *luxe* 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

$$\mathcal{E}_D^{\text{revenu}}(Y_0) < 0 \qquad \text{(bien inférieur).}$$

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.*

1.5 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 1.8. *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 1.4 (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 1.6 assure que $f'(c) = 0$. Le Théorème de Rolle est donc établi. □

Le théorème suivant est essentiellement une conséquence du Théorème de Rolle et sera l'ingrédient principal de la preuve de la Proposition 1.8.

Théorème 1.5 (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 $x \longmapsto \frac{f(b)-f(a)}{b-a}x$ 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 :

$$\begin{aligned}
g(a) &= f(a) - \frac{f(b) - f(a)}{b - a}a \\
&= \frac{f(a)(b - a) - a\left(f(b) - f(a)\right)}{b - a} \\
&= \frac{bf(a) - af(a) - af(b) + af(a)}{b - a} \\
&= \frac{bf(a) - af(b)}{b - a} \\
&= \frac{bf(b) - af(b) - bf(b) + bf(a)}{b - a} \\
&= \frac{f(b)(b - a) - b\left(f(b) - f(a)\right)}{b - a} \\
&= f(b) - \frac{f(b) - f(a)}{b - a}b \\
&= g(b)
\end{aligned}$$

Ainsi, par le Théorème de Rolle (Théorème 1.4) 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 1.8 : 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 1.5), 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. □

Chapitre 2

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*).

2.1 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 correspondant dans le cours du premier semestre.

Pour chaque individu $i \in \{1, \ldots, N\}$, 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 $x_1, \ldots, x_l$ (resp. $y_1, \ldots, y_r$) 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`

Y X	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 2.1. *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 2.1. 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 :

Y X	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

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 2.1. 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 \ 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 2.1. 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 2.3.

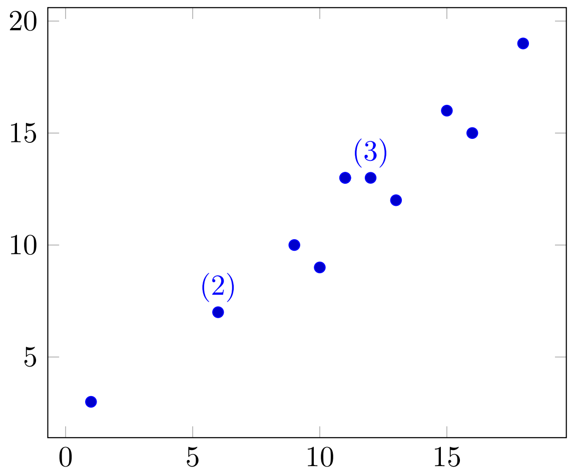

FIGURE 2.1 – Représentation de la série de l'Exemple 2.1 par un nuage de points pondérés.

2.2 Étude et mesure des liens entre les variables

2.2.1 Indépendance

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

Définition 2.2. *Soit (X, Y) une série statistique bivariée.*

On dit que X et Y sont indépendantes *si :*

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

Définition 2.3. *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 2.1. *Soit (X, Y) une série statistique.*

Les assertions suivantes sont équivalentes :

1. *X et Y sont indépendantes,*
2. *pour tout $i \in \{1, \ldots, l\}$, $f_{i|j}$ ne dépend pas de $j \in \{1, \ldots, r\}$ (et vaut $f_{i,\cdot}$),*
3. *pour tout $j \in \{1, \ldots, r\}$, $f_{j|i}$ ne dépend pas de $i \in \{1, \ldots, l\}$ (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), $f_{i,j} = f_{i,\cdot} f_{\cdot,j}$. En divisant par $f_{\cdot}$, on obtient que X et Y sont indépendantes si, et seulement si, pour tout (i, j)

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

□

Remarque 2.2.

1. Il découle de la Proposition 2.1 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*.

2.2.2 Coefficient de Cramer

On souhaite déterminer si les composantes X et Y d'une série statistique bivariée sont inépendantes (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 Ysont effectivement indépendantes.

Définition 2.4. *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 $T_{i,j} = \frac{n_{i,\cdot} n_{\cdot,j}}{N}$ est l'effectif théorique de la caractéristique (x_i, y_j) en faisant l'hypothèse que X et Ysont effectivement indépendantes.

Proposition 2.2. *Soit (X, Y) une série statistique bivariée.*
On a :

$$\chi^2 \in \left[0, \chi^2_{max}\right]$$

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

Remarque 2.3.

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

Définition 2.5. *Soit (X, Y) une série statistique bivariée.*
On appelle coefficient de Cramer *(ou* V de Cramer*) de la série statistique (X, Y), la quantité :*

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

Proposition 2.3. *Soit (X, Y) une série statistique bivariée.*
On a :

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

Remarque 2.4. 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 2.2. 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 étés 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) :

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

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

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

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

2.2.3 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 2.2.2. 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 2.3. 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 155 habitants sur leur département de résidence X et le fait Y qu'ils aient vu la publicité ou non. Les résultats ont étés consignés dans le tableau suivant.

X \ 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	7	8

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 2.3, 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 perųe sur l'ensemble du territoire.

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

Cette étape est analogue à ce que l'on a présenté dans la Section 2.2.2. 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 :

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

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 2.2.2, on complète le tableau des effectifs observés avec les effectifs marginaux comme suit.

X \ 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	7	8	15
$n_{\cdot,j}$	76	79	$N = 155$

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	$\frac{28\times72}{155} = 13,73$	14,27	28
Doubs	14,22	14,78	29
Jura	6,862	7,14	14
Nièvre	5,40	5,60	11
Haute-Saône	5,88	6,12	12
Saône-et-Loire	13,73	14,27	28
Yonne	8,83	9,17	18
Territoire de Belfort	7,35	7,65	15
$n_{\cdot,j}$	76	79	$N = 155$

Observons qu' il n'y a pas d'effectif théorique est inférieur à 5.

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

On calcule, comme dans la Section 2.2.2, 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 :

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

Dans le contexte de l'Exemple 2.3, on obtient :

$$\chi^2_{\text{obs}} \simeq 21,2640.$$

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 2.3, on a 8 lignes et 2 colonnes donc :

$$\text{ddl} = (8 - 1) \times (2 - 1) = 7,$$

et la statistique de test suit la loi $\chi^2(7)$.

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 tord H_0, autrement dit :

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

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* $\chi_c^2 = \chi^2_{c,\alpha,\mathrm{ddl}}$ dépendant de α et ddl telle que si χ_2 suit la loi $chi_2(\text{ddl})$, 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

KHIDEUX.INVERSE(α;ddl).

Dans le contexte de l'Exemple 2.3, en choisissant le risque $\alpha = 1\% = 0,01$ et en utilisant que ddl=7, on obtient :

$$\chi_c^2 \simeq 18,4753.$$

Étape 6 : Règles de décision :

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

Ces règles de décisions ce comprennent comme suit. La valeur du Khi-2 observée χ^2_{obs} 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 2.3, les règles de décision sont :
— Si $\chi^2_{\text{obs}} > \chi_c^2 \simeq 18,4753$, on rejette H_0 ;
— Si $\chi^2_{\text{obs}} \leq \chi_c^2 \simeq 18,4753$, 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 2.3, on a :

$$\chi^2_{\text{obs}} \simeq 21,2640 > \chi^2_c \simeq 18,4753$$

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 dons de risquer de fournir un travail supplémentaire inutile !).

2.2.4 Coefficient de corrélation linéaire

Définition 2.6. *Soit (X, Y) une série statistique bivariée avec X et Y quantitatives.*

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

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

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

Remarque 2.5. 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 2.4. *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 :*
$$\text{Cov}(X, X) = \text{V}[X].$$
3. *Si X et Y sont indépendantes, alors $\text{Cov}(X, Y) = 0$.*
4. *On a :*
$$\text{Cov}(X, Y) = \frac{1}{N}\sum_{i=1}^{l}\sum_{j=1}^{r} n_{i,j}x_iy_j - \overline{X}\overline{Y}.$$

Définition 2.7. *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 2.5. *Soit (X, Y) une série statistique bivariée avec X et Y quantitatives.*

1. *On a :*

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

2. *On a :*

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

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

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

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

Interprétation : On dira qu'il y a :
— une *forte corrélation* entre X et Y si $|\,\mathrm{Cor}(X,Y)| \geq 0,8$;
— une *corrélation médiocre* entre X et Y si $0,5 < |\,\mathrm{Cor}(X,Y)| < 0,8$;
— une *mauvaise corrélation* entre X et Y si $|\,\mathrm{Cor}(X,Y)| \leq 0,5$.

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 2.4. 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 :

$$\overline{X} = \frac{20 + \cdots + 32}{10} = 57,2$$

$$\overline{Y} = \frac{47 + \cdots + 45}{10} = 111,6$$

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

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

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

et donc

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

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 2.2. 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.

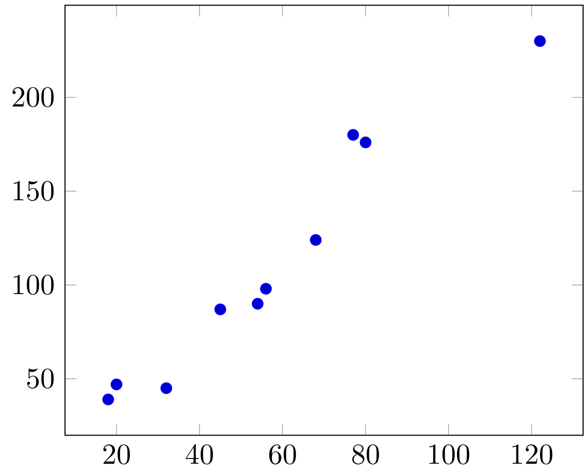

FIGURE 2.2 – Représentation des prix relevés des appartements en fonction de leurs surfaces (Exemple 2.4).

2.3 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 *regression* (ou d'*ajustement*) de Y en X. Quand on cherche à prédire X en supposant Y connue on parle de *regression* (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 (models exponentiels, logarithmiques...). Nous restreindrons notre attention au cas simple où l'on cherche une relation entre X et Y de la forme

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

ou

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

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 $\mathrm{Cor}(X, Y)$ 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 2.8. *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 2.6.

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 $\mathrm{Cor}(X, Y) = \pm 1$). Dans le cas contraire, les droites $D_{Y|X}$ et $D_{X|Y}$ se coupent en $G(\overline{X}, \overline{Y})$ 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 2.6. *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 $a = \frac{\mathrm{Cov}(X,Y)}{\mathrm{V}[X]}$ et $b = \overline{Y} - a\overline{X}$.
2. *L'équation de la droite de régression de X en Y est :*
$$D_{X|Y} : x = a'y + b',$$
avec $a' = \frac{\mathrm{Cov}(X,Y)}{\mathrm{V}[Y]}$ et $b' = \overline{X} - a'\overline{Y}$.

Exemple 2.4 (suite)**.** Reprenons l'Exemple 2.4. On a vu que $|\mathrm{Cor}(X, Y)| \geq 0,8$ 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
$$a = \frac{\mathrm{Cov}(X, Y)}{\mathrm{V}[X]} \simeq 2 \quad \text{et} \quad b = \overline{Y} - a\overline{X} \simeq -2,8.$$

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

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

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

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

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

Ces droites de régressions sont représentées dans la Figure 2.3.

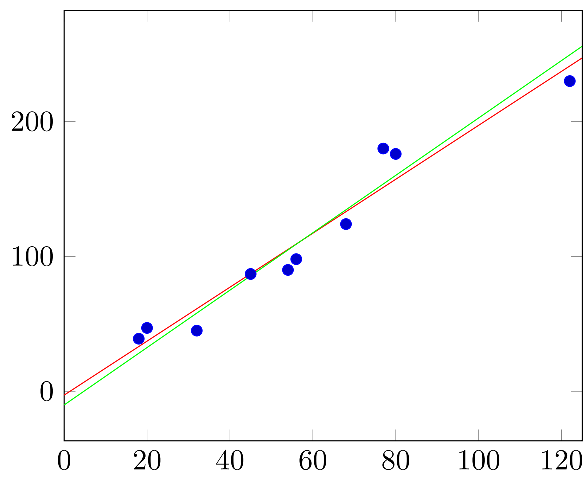

FIGURE 2.3 – Représentation des prix relevés des appartements en fonction de leurs surfaces (Exemple 2.4) et des droites de régression de Y en X (rouge) et de X en Y (vert).

Chapitre 3

Séries chronologiques

3.1 Premières définitions, exemples et motivations

Définition 3.1. *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 3.1. *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 3.1. 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	70
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 $y_1 = 20$, $y_2 = 25$, $y_3 = 50$, $y_4 = 70$, $y_5 = 35$, $\ldots$, $y_{16} = 170$.

Exemple 3.2. 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

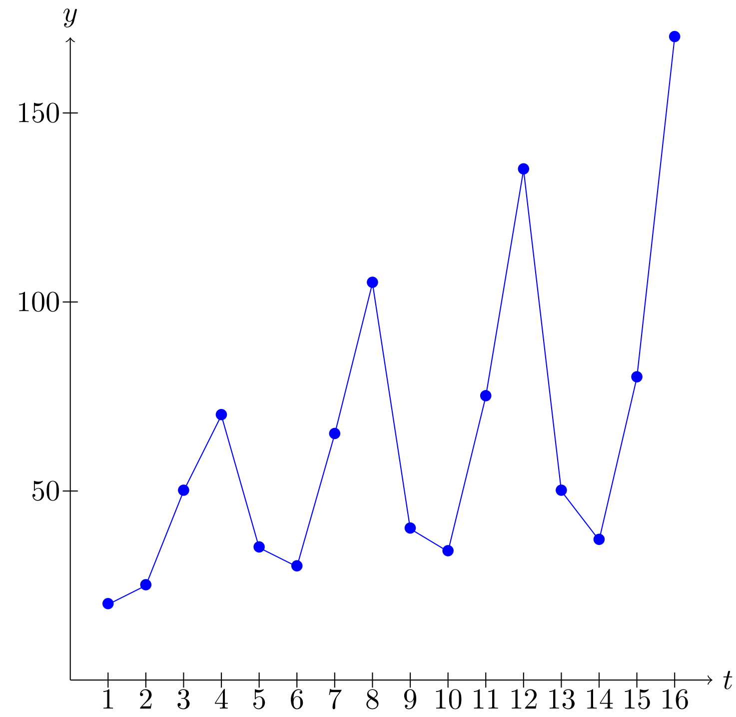

FIGURE 3.1 – Chiffres d'affaires y_t (en K€) de la t^e période (Exemple 3.1).

Ici, le nombre d'années est $n = 3$, le nombre de périodes par année est $p = 12$ et on a $y_1 = 1$, $y_2 = 1,5$,..., $y_{12} = 0,5$, $y_{13} = 3,5$, ..., $y_{36} = 6$.

3.1.1 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 3.1 et 3.2). 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 3.2 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 3.2) ;

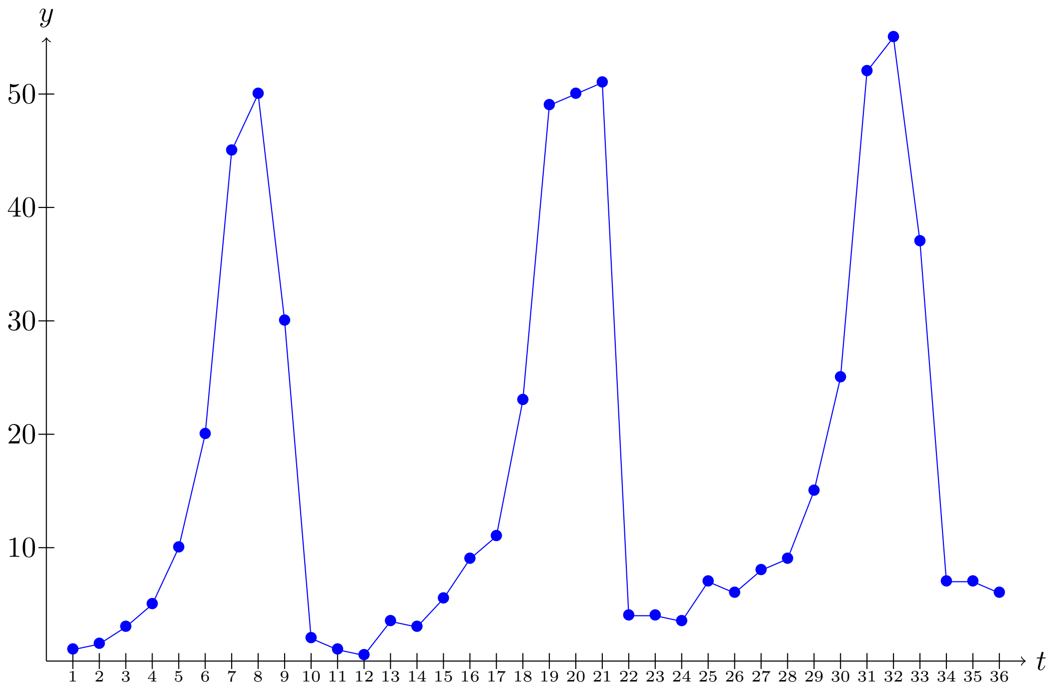

FIGURE 3.2 – Consommation d'eau y_t (en Hm^3) au cours du t^e mois (Exemple 3.2).

— d'estimer les composantes de la série chronologique (voir Section 3.3) ;
— prédire les valeurs futures de la série (voir Section 3.4).

3.2 Analyse d'une série chronologique

3.2.1 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 **g**énérale » de la série au temps t,
— s_t désigne la composante « **s**aisonnière » de la série au temps t,
— a_t désigne la composante « **a**léatoire » 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 3.2.2) et *multiplicatif* (voir Sous-section 3.2.3).

Exemple 3.1 (suite)**.** Reprenons l'exemple des chiffres d'affaires (Exemple 3.1).

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 supposerons par la suite que celle-ci est négligeable.

3.2.2 Modèle additif

Définition 3.2. *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* **g***énérale », s_t désigne la composante* **s***aisonnière et* a_t *désigne la composante* **a***léatoire de la série au temps* t.

Remarque 3.1. 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.* $s_1 + s_2 + \cdots + s_p = 0$ pour le modèle additif). La composante aléatoire est supposée négligeable (*i.e.* $a_t \simeq 0$, pour tout t, pour le modèle additif).

Critère 3.1. *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 3.2 (suite)**.** Le modèle additif est adapté pour l'Exemple 3.2 (voir Figure 3.3).

3.2.3 Modèle multiplicatif

Définition 3.3. *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* **g***énérale », s_t désigne la composante* **s***aisonnière et* a_t *désigne la composante* **a***léatoire de la série au temps* t.

Remarque 3.2. 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.* $s_1 \times s_2 \times \cdots \times s_p = 1$ pour le modèle multiplicatif). La composante aléatoire est supposée négligeable (*i.e.* $a_t \simeq 1$, pour tout t, pour le modèle multiplicatif).

Critère 3.2. *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 3.1 (suite)**.** Le modèle multiplicatif est adapté pour l'Exemple 3.1 (voir Figure 3.4).

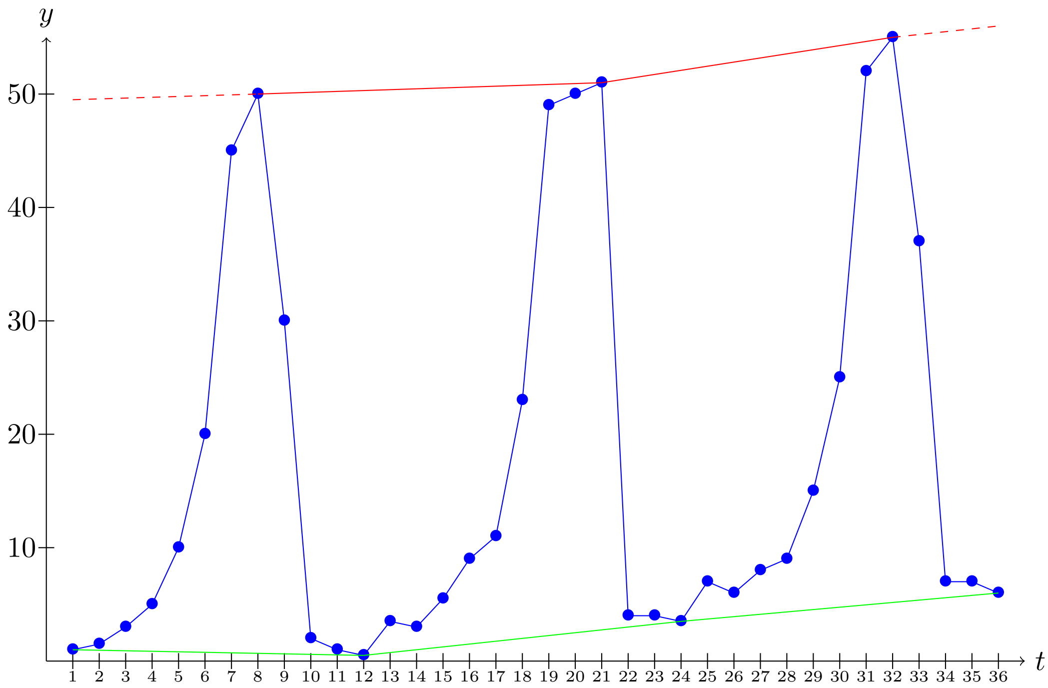

FIGURE 3.3 – Consommation d'eau y_t (en Hm^3) au cours du t^{e} mois de l'Exemple 3.2 (bleu) et les lignes polygonales passant par les pics « positifs » (rouge) et « négatifs » (vert).

3.2.4 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 correspondants 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, ... nécéssaires à une bonne compréhension du graphique et pour le choix du modèle (pensez au `clic droit` pour accéder aux menus d'édition).

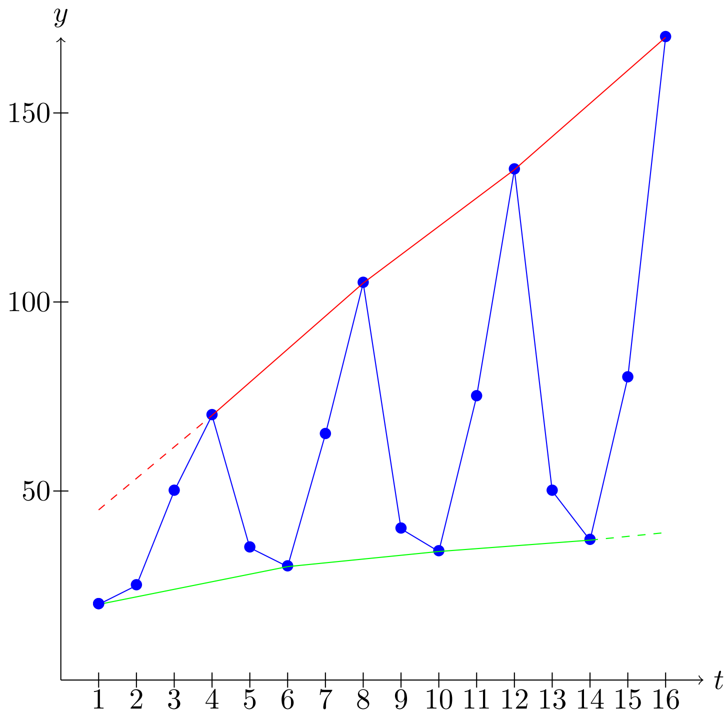

FIGURE 3.4 – Chiffres d'affaires y_t (en K€) du t^e trimestre de l'Exemple 3.1 (bleu) et les lignes polygonales passant par les pics « positifs » (rouge) et « négatifs » (vert).

3.3 Estimation des composantes d'une série chronologique

3.3.1 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 2). 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 3.4. *Soit* $\mathbf{y}$ *une série chronologique portant sur* n *années décomposées en* p *saisons et soit* $k \in \{2, \ldots, np\}$*. On appelle* moyenne mobile (ou glissante) d'ordre k au temps $t \in \{1 + \lfloor \frac{k}{2} \rfloor, \ldots, np - \lfloor \frac{k}{2} \rfloor\}$ *la quantité :*

$$\mathrm{M}_t^{(k)} = \begin{cases} \frac{1}{k}\left(y_{t-\frac{k-1}{2}} + \cdots + y_t + \cdots + y_{t+\frac{k-1}{2}}\right) & \textit{si } k \textit{ est impair} \\ \frac{1}{k}\left(\frac{y_{t-\frac{k}{2}}}{2} + y_{t-\frac{k}{2}+1} + \cdots + y_t + \cdots + y_{t+\frac{k}{2}-1} + \frac{y_{t+\frac{k}{2}}}{2}\right) & \textit{si } k \textit{ est pair} \end{cases}.$$

Exemple 3.3. Les moyennes mobiles d'ordre 2, 3, 4 et 5 sont données par :

$$\mathrm{M}_t^{(2)} = \frac{1}{2}\left(\frac{y_{t-1}}{2} + y_t + \frac{y_{t+1}}{2}\right),$$

$$\mathrm{M}_t^{(3)} = \frac{1}{3}\left(y_{t-1} + y_t + y_{t+1}\right),$$

$$\mathrm{M}_t^{(4)} = \frac{1}{4}\left(\frac{y_{t-2}}{2} + y_{t-1} + y_t + y_{t+1} + \frac{y_{t+2}}{2}\right)$$

et

$$\mathrm{M}_t^{(5)} = \frac{1}{5}\left(y_{t-2} + y_{t-1} + y_t + y_{t+1} + y_{t+2}\right).$$

Remarque 3.3.

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 jeux 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 3.1 ($k = p = 4$). Dans l'Exemple 3.2, 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 3.1. *Soit* $\mathbf{y}$ *une série chronologique dont la tendance générale est notée* g_t*.*
Alors, un estimateur de g_t *est donné par :*

$$\hat{g}_t = \mathrm{M}_t^{(k)}.$$

Autrement dit, on peut considérer que :

$$g_t \simeq \hat{g}_t = \mathrm{M}_t^{(k)}.$$

Remarque 3.4. L'estimation de la tendance générale ne dépend pas du fait que l'on utilise un modèle additif ou multiplicatif.

Exemple 3.1 (suite)**.** L'estimation de la tendance des chiffres d'affaires trimestriels de l'Exemple 3.1 par des moyennes mobiles d'ordre 4 est donnée dans le tableau suivant et représentée dans la Figure 3.5.

Année \ Trimestre	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		

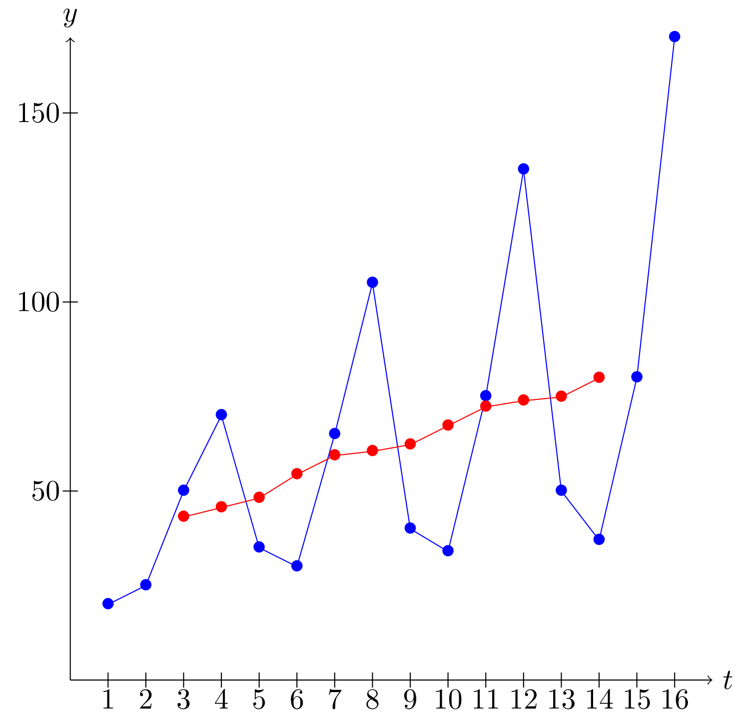

FIGURE 3.5 – Chiffres d'affaires y_t (en K€) du t^{e} trimestre (bleu, Exemple 3.1) et estimation de la tendance g_t par $\hat{g}_t = \mathrm{M}_t^{(4)}$ (rouge).

Exemple 3.2 (suite)**.** Les valeurs approchées à une décimale de l'estimation de la tendance la consommation mensuelle d'eau de l'Exemple 3.2 par des moyennes mobiles d'ordre 7 sont données dans le tableau suivant et représentées dans la Figure 3.6.

	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			

3.3.2 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 :

$$y_t = g_t + s_t + a_t$$

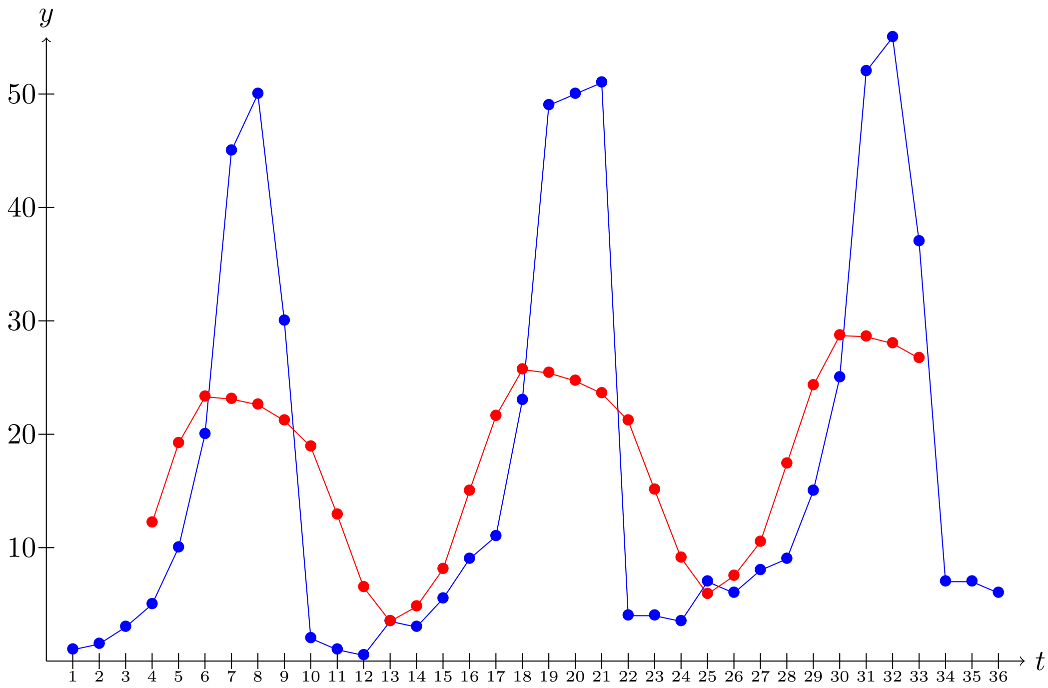

FIGURE 3.6 – Consommation d'eau y_t (en Hm^3) au cours du t^{e} mois (bleu, Exemple 3.1) et estimation de la tendance g_t par $\hat{g}_t = \mathrm{M}_t^{(7)}$ (rouge).

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 $\hat{g}_t = \mathrm{M}_t^{(k)}$. 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 $\bar{e}_t$: $\bar{e} = \frac{\bar{e}_1 + \cdots + \bar{e}_p}{p}$;
4. pour $t \in \{1, \dots, p\}$, on calcule les *coefficients saisonniers normalisés* $\hat{s}_t = \bar{e}_t - \bar{e}$.

Remarque 3.5.

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 3.2. *Dans le cadre d'un modèle additif, une estimation de la composante saisonnière est donnée par* $\hat{s}_t$ *défini ci-dessus.*

Exemple 3.2 (suite)**.** Reprenons l'Exemple 3.2. Le tableau suivant donne les valeurs approchées à une décimale près des écarts $e_t = y_t - \hat{g}_t$ obtenus pour chaque période $t \in \{4, \ldots, 33\}$, celles des moyennes $\bar{e}_t$ des écarts correspondant à chacun des mois d'une année :

$$\bar{e}_1 = \frac{e_{13} + e_{25}}{2}, \ldots, \bar{e}_3 = \frac{e_{15} + e_{27}}{2}, \bar{e}_4 = \frac{e_4 + e_{16} + e_{28}}{3}, \ldots, \bar{e}_9 = \frac{e_9 + e_{21} + e_{33}}{3},$$

$$\bar{e}_{10} = \frac{e_{10} + e_{22}}{2}, \ldots, \bar{e}_{12} = \frac{e_{12} + e_{24}}{2},$$

ainsi que l'estimation de la composante saisonnière $\hat{s}_t = \bar{e}_t - \bar{e}$ où $\bar{e} = \frac{\bar{e}_1 + \bar{e}_2 + \cdots + \bar{e}_{12}}{12} \simeq 0,05$.

	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 $\hat{g}_t = \mathrm{M}_t^{(k)}$. 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 $t \in \{1 + \lfloor \frac{k}{2} \rfloor, \ldots, np - \lfloor \frac{k}{2} \rfloor\}$, 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 $t \in \{1, \ldots, p\}$, 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 $\bar{r}_t$: $\bar{r} = \sqrt[p]{\bar{r}_1 \times \cdots \times \bar{r}_p}$;

4. pour $t \in \{1, \dots, p\}$, on calcule les *coefficients saisonniers normalisés* $\hat{s}_t = \frac{\bar{r}_t}{\bar{r}}$.

Remarque 3.6.

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 3.3. *Dans le cadre d'un modèle multiplicatif, une estimation de la composante saisonnière est donnée par $\hat{s}_t$ défini ci-dessus.*

Exemple 3.1 (suite)**.** Reprenons l'Exemple 3.1. 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

3.4 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 $\hat{g}_t = \mathrm{M}_t^{(k)}$, $t \in \{1, \dots, np\}$ (voir Sous-section 3.3.1).
2. On estime la composante saisonnière par $\hat{s}_t$ comme détaillé dans la Sous-section 3.3.2.
3. On effectue une régression linéaire sur les $\hat{g}_t$ pour prédire leurs valeurs futures (voir Chapitre 2).
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 3.1 (suite)**.** Reprenons l'Exemple 3.1. 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 3.7.

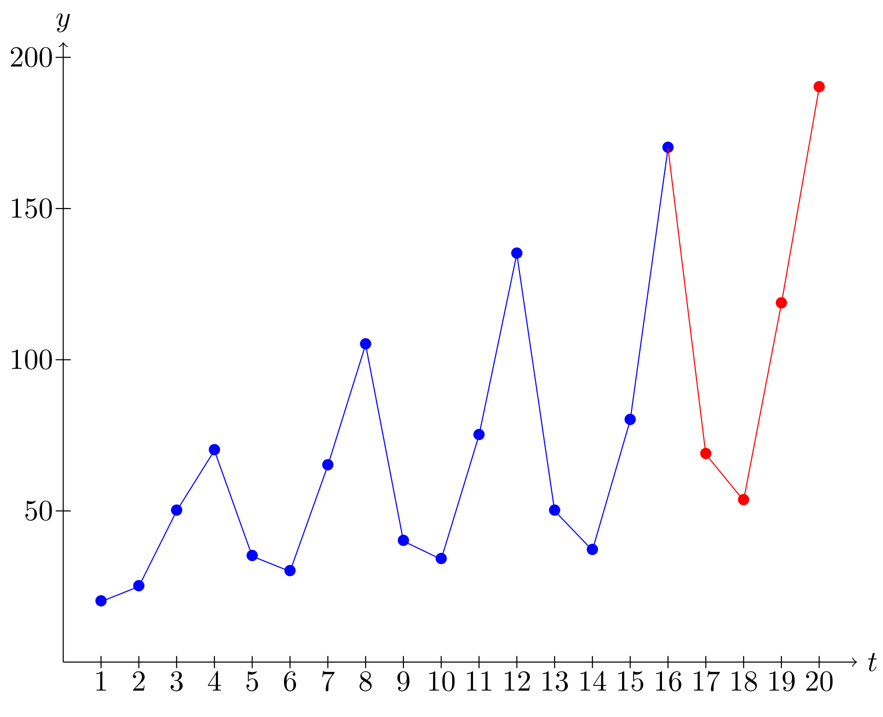

FIGURE 3.7 – Chiffres d'affaires y_t (en K€) de la t^e période de l'Exemple 3.1 (bleu) et prédictions pour l'année 2016 (rouge).

Exemple 3.2 (suite)**.** Reprenons l'Exemple 3.2. 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 $t \in \{37, \dots, 48\}$,

$$y_t \simeq 0,04t + 17,1 + \hat{s}_t.$$

Les prédictions sont les suivantes et sont représentées dans la Figure 3.7 :

$$y_{37} \simeq 19,08, \quad y_{38} \simeq 16,92, \quad y_{39} \simeq 16,06, y_{40} \simeq 11,4, \quad y_{41} \simeq 8,94, \quad y_{42} \simeq 15,88,$$

$$y_{43} \simeq 41,72, \quad y_{44} \simeq 45,36, \quad y_{45} \simeq 27,65, \quad y_{46} \simeq 1,84, y_{47} \simeq 7,48, \quad y_{48} \simeq 13,19.$$

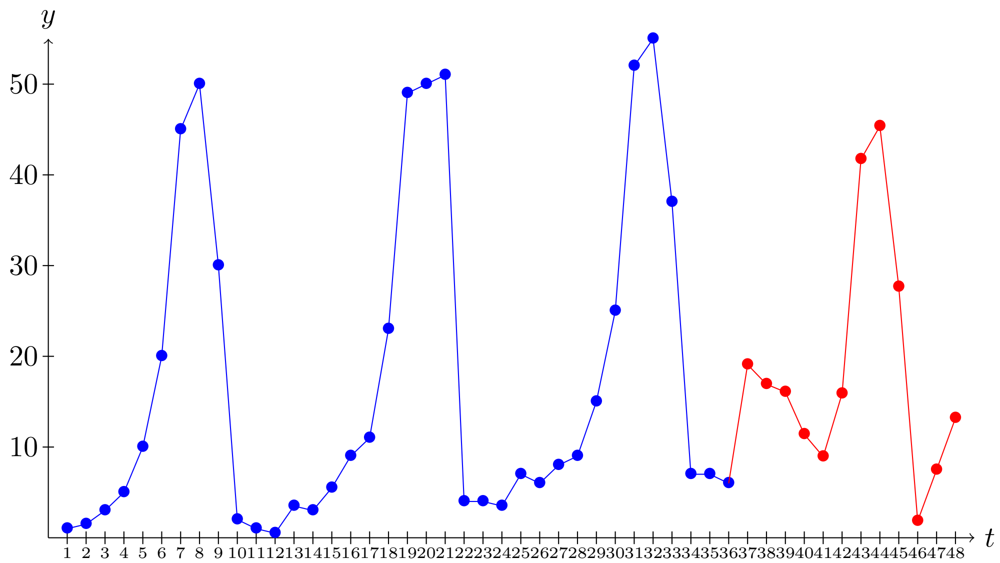

FIGURE 3.8 – Consommation d'eau y_t (en Hm^3) au cours du t^{e} mois de l'Exemple 3.2 (bleu) et prévisions pour l'année 2016 (rouge).

Remarque 3.7. Les prévisions obtenues dans l'Exemple 3.2 sont moins satisfaisantes que celles obtenues dans l'Exemple 3.2. On peut observer sur la Figure 3.6 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.

3.5 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, ... nécéssaires à une bonne compréhension du graphique et pour le choix du modèle (pensez au `clic droit` pour accéder aux menus d'édition).