Chapitre I

GÉNÉRALITÉS SUR LES SÉRIES CHRONOLOGIQUES

1.1 Introduction

La théorie des séries temporelles(chronologiques) abordée dans ce travail est appliquée de nos jours dans des domaines aussi variés que l'économétrie, la médecine ou la démographie, pour n'en citer qu'une petite partie. On s'intéresse à l'évolution au cours du temps d'un phénomène, dans le but de décrire, expliquer puis prévoir ce phénomène dans le futur. Elles sont fréquentes aussi en gestion : surveillance du niveau des stocks, suivi des ratios d'une entreprise etc.. . . Leur particularité vient de l'introduction du temps dans l'analyse de ces données : on étudie une suite de couples de la forme (t, Y_t)), où Y_t est l'observation de la variable à l'instant t

L'étude des séries chronologiques (ou séries temporelles ou chroniques - time series- en anglais) permet de décrire, expliquer, contrôler, prévoir des phénomènes évoluant au cours du temps.

1.1.1 Objectifs

Les principaux objectifs de la modélisation des séries temporelles sont les suivants.

- ▶ En économétrie, détecter puis analyser les périodes de crises et croissances ;
- ▶ En reconnaissance vocale, reconnaitre les mots dans des signaux ;
- ▶ Dans le séquençage du génome, détecter les parties de l'ADN qui contiennent de l'information.
- ▶ Décrire l'évolution.

- Permettre l'explication des fluctuations.
- Faciliter la prévision (le passé peut expliquer le futur)
- Comparer deux séries temporelles. Par exemple, l'évolution démographique de deux régions ou deux séquences d'ADN.
- Prédire l'évolution future de la série temporelle à partir des valeurs qui ont été observées.

Par exemple, pour des raisons socio-économiques on veut prévoir le temps qu'il va faire, l'évolution des ventes d'un produit, la consommation d'électricité, etc.)

Comment prévoir : en s'appuyant sur le passé. Pour prédire les ventes de l'année $j+1$, on s'appuie sur l'évolution des ventes durant les années j, $j-1$, mais on tient compte aussi d'évènement (conjoncture, économique, crise, ...).

Peut-on prévoir parfaitement bien ?

1.1.2 Qu'est-ce qu'une série temporelle ?

Définition 1 *On appelle série chronologique la succession de valeurs que prend une variable (ou caractéristique) au cours du temps pour N périodes successives pour un même individu (ou cas).*

Donc, une série chronologique est une suite d'observations d'une variable statistique au cours du temps.

Une série chronologique peut être identifiée à une série statistique à deux variables t et Y_t, où t représentele temps. On la note (t, Y_t))

Exemple 2 ***1-*** *Nombre mensuel de vente de voitures neuves en France.*

2- *Nombre annuel de naissance au Maroc*

3- *Nombre trimestriel de chômeurs aux Etats-Unis*

4- *Consommation d'électricité mensuelle.*

1.1.3 Domaines d'application.

On trouve des exemples de séries chronologiques univariées dans de très nombreux domaines.

- **Finance et économétrie** : évolution des indices boursiers, des prix, des données économiques.

des entreprises, des ventes et achats de biens, des productions agricoles ou industrielles.

- **Assurance** : analyse des sinistres.
- **Médecine / biologie** : suivi des évolutions des pathologies, analyse d'électro-encéphalogrammes et d'électrocardiogrammes.
- **Sciences de la Terre et de l'Espace** : indices de marées, variations des Phénomènes physiques (Météorologie), évolution des taches solaires, phénomènes d'avalanches.
- **Traitement du signal** : signaux de communications, de radars, de sonars, analyse de la parole.
- **Démographie** : évolution de la population.

Exemple 3 *On s'intéresse ici au Chiffre d'Affaires trimestriel de l'entreprise Biomérieux entre* 2010 *et* 2012*. Deux façons équivalentes deprésenterce jeu de données :*

t	1	2	3	4	5	6	7	8	9	10	11	12
Y_t	306	344	333	373	327	345	347	406	362	387	382	437

1.1.4 Les deux présentations d'une série chronologique

Une série chronologique est présentée soit sous la forme d'un tableau à deux colonnes, contenant np lignes

t	y_t
1	y_2
2	y_3
.	.
.	.
.	.
np	y_{np}

Soit sous la forme d'un tableau contenant p colonnes et n lignes : une colonne par mois et une ligne par anné

	Mois					
	1	2	...	j	...	p
Année 1	y_{11}	y_{12}	...	y_{1j}	...	y_{1p}
Année 2	y_{21}	y_{22}	...	y_{2j}	...	y_{2p}
.	.	.	...	..		
.	.	.	...	.		
.	.	.	...			
i	y_{i1}	y_{i2}	...	y_{ij}	...	y_{ip}
.	.					
n	y_{n1}	y_{n2}	...	y_{nj}	...	y_{np}

Exemple 4 *Considérons la série trimestrielle du chiffre d'affaires en milliers de francs des ventes d'un magasin de* 1978 *à* 1982*.*

		T_1 $j=1$	T_2 $j=2$	$T3$ $j=3$	T_4 $j=4$
$i=1$	1978	2614	3010	2765	4856
$i=2$	1979	3010	3397	3168	5624
$i=3$	1980	3406	...		
...					

t	1	2	3	4	5	6	7	8	9	
y_t	2614	3010	2765	4856	3010	3397	3168	5624	3406	.....

1.1.5 Representation graphiques d'une série chronologique

a) **Graphe de la série chronologique**.

On représente graphiquement la série chronologique $\{y_t\}_{t\in T}$

1- En dessinant le nuage formé par les points $(t_j, y_j)_{1\leq j\leq n}$

2- En reliant les points entre eux par des segments de droite, pour indiquer la chronologie

On représente l'évolution de la grandeur considérée sur l'ensemble de la période observée.

1.1.6 Intérêt de la représentation graphique :

Essayer de repérer les caractéristiques de la séries chronologique, comme

- Une tendance
- Un cycle
- Un phénomène périodique
- Des variations accidentelles
- Des fluctuations (variations) irrégulières

Exemple 5

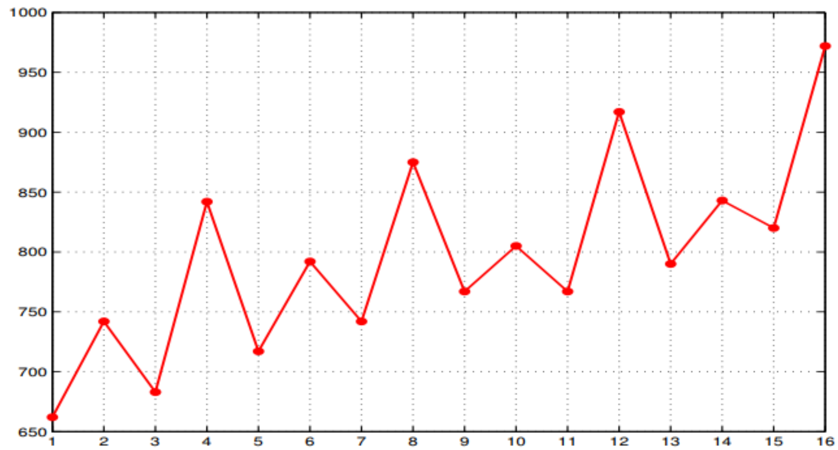


Fig. 1: Ventes trimestrielles (en Milliers d'Euros)

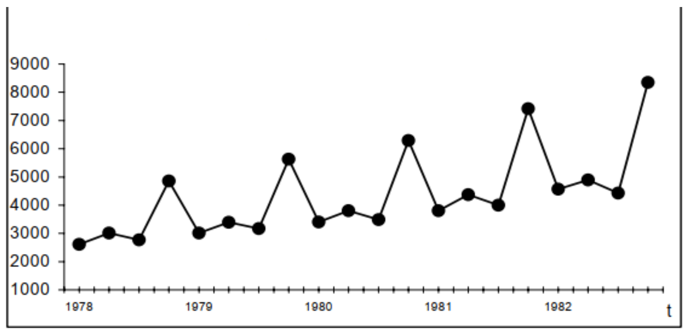


Fig. 2: Graphique de (Yt)

Exemple 6 *On représente les points* (t, Y_t), *que l'on relie par des segments de droites.*

On représente l'évolution de la grandeur considérée sur l'ensemble de la période observée.

b) Graphiques des courbes superposées.

On représente les points $(j, Y_{i,j})$ que l'on relie par des segments de droites, ceci pour chacune des années i.

On représente ainsi l'évolution annuelle de la grandeur au cours des mois (pour chacune des années). On peut ainsi comparer le même mois j des différentes années, mais on ne voit pas l'évolution globale.

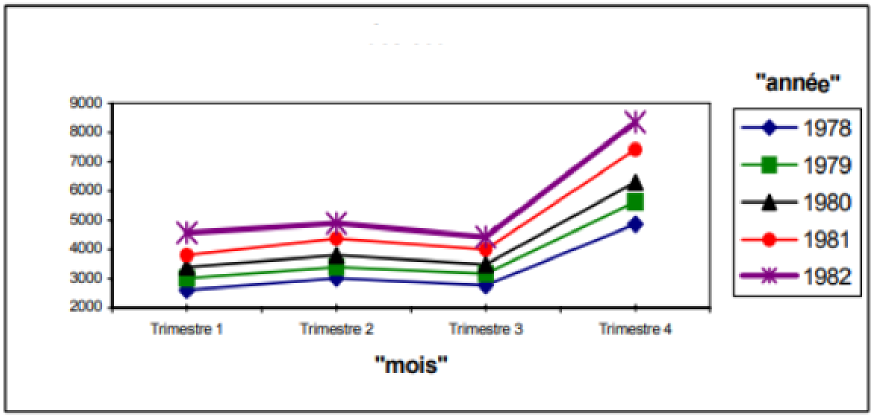


Fig. 3: Graphiques descourbes superposées.

Remarque 7 *La variable t sert à ordonner Y_t, ses valeurs sont séparées par une constante c'est-à-dire l'unité de temps (année, mois, semaine) qui caractérise la série chronlogique et en donne la périodicité. On peut ainsi parler de séries chronologiques annuelles, mensuelles ou trimestrielles. lles de temps réguliers. Dans la pratique, ce n'est pas toujours réalisé. Par exemple, pour des séries mensuelles, le nombre de jours varie d'un mois à l'autre, on utilise alors des données corrigées.*

1.2 Modéles de composition

Pour étudier (t, Y_t), on décompose la série chronologique en différents mouvements. On distingue en général trois composantes constitutifs d'une série chronologique

1.2.1 Les composantes fondamentales d'une série chronologique

– **La tendance à long terme ou "Trend"** : T_t

Définition 8 *La tendance ou trend : traduit généralement l'évolution générale du phénomène observè (croissance, d´ecroissance, stagnation,. . .). C'est une courbe (ici une droite) que l'on note T_t (fonction de t).*

Elle traduit l'aspect général (moyen) de la série. Par exemple, la série de la Figure 1.3 a tendance à augmenter de façon linéaire.

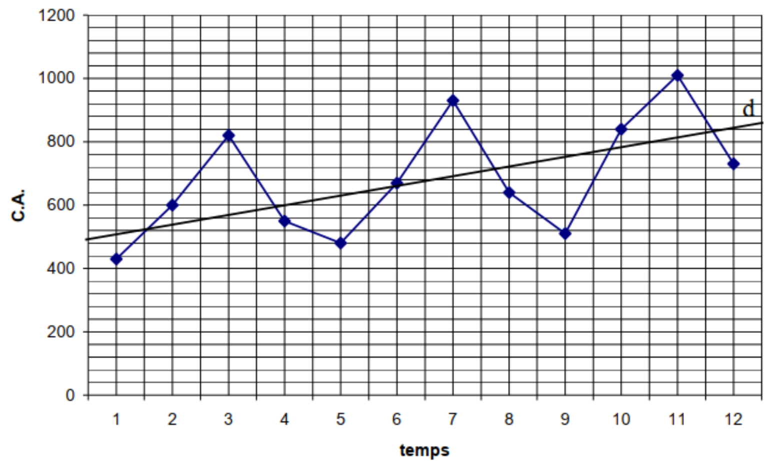


Fig. 4: Evolution du C.A. par trimestre

par exemple :

- *Tendance linéaire :* $T_t = at + b$
- *Tendance quadratique :* $T_t = at^2 + bT + c$
- *Tendance logarithmique :* $T_t = \log(t)$

– **La composante cyclique : $\mathbf{C_t}$**

Pour des séries trés longues, elle correspond à des variations connues régulières. Par exemple, en économie, des fluctuations correspondant à des périodes de prospérité ou de récession.

En général, la composante cyclique n'est pas détectable sur la période étudiée et on suppose alors que C_t n'existe pas

– **La composante saisonnière : $\mathbf{S_t}$**

Elle correspond à des variations régulières inconnues mais détectables sur des périodes courtes, généralement à l'intérieur d'une année : semaine, mois, trimestre. Par exemple, l'influence des congés annuels sur la production d'une entreprise. Sa période est égale à 12 pour les séries mensuelles, et à 4 pour les séries trimestrielles.

Si p désigne la période du mouvement saisonnier : $s_t = s_t + p = s_t + 2p = ...$

La composante saisonnière est donc totalement déterminée par p coefficients saisonniers :

$Cs_1, \ldots, Cs_j, ..., Cs_p$

– **La composante accidentelle** (résiduel) **: $\mathbf{A_t}$**

Le mouvement accidentel ou résiduel (composante accidentelle ou résiduelle) correspond à des variations accidentelles de forte amplitude dues à des phénomènes imprévisibles.

Par exemple, la grève, le risque de guerre, forte baisse des températures en été qui entraine une augmentation de l'éléctricité (chauffage) sur cette période). Cette partie est aléatoire.

1.3 Les modèles de décomposition déterministe

On étudiera deux modèles de décomposition déterministes :-

1. **Le modèle (ou Schéma) additif**
2. **Le modèle (ou Schéma) multiplicatif**

combinant chacun

– une tendance (T_t)

– une composante saisonnière (S_t) ;

– une composante résiduelle (A_t)

1.3.1 Le modèle additif

Il est défini par

$$Y_t = T_t + C_t + S_t + A_t,$$

si C_t, S_t et A_t sont indépendants de T_t

Dans ce modèle, le nuage de points a une enveloppe d'épaisseur plus ou moins constante

1.3.2 Le modèle multiplicatif :

- On emploie le modèle multiplicatif lorsque l'enveloppe du nuage de points « s'élargit » au fur et à mesure que la tendance générale croît (et est de plus en plus « resserrée » au fur et à mesure que le trend diminue tout en restant positif). Le terme Y_t est alors vu comme le produit de la tendance générale T_t, de la composante saisonnière S_t et de la composante aléatoire A_t :

$$Y_t = T_t \times S_t \times A_t$$

En général, le modèle le mieux adapté est le modèle multiplicatif.

Exemple 9 *Le tableau ci-dessous indique le nombre de naissances par trimestre d'un Land allemand, au cours des dernières années.*

Trimestre / Années.	2004	2005	2006	2007	2008	2009	2010	2011
trimestre 1	7684	7437	7311	7221	7148	7105	7067	7062
trimestre 2	7899	7705	7616	7471	7336	7189	7146	7128
trimestre 3	7320	7208	7093	7008	6970	7043	6983	7008
trimestre 4	7683	7450	7298	7184	7231	7206	7185	7088

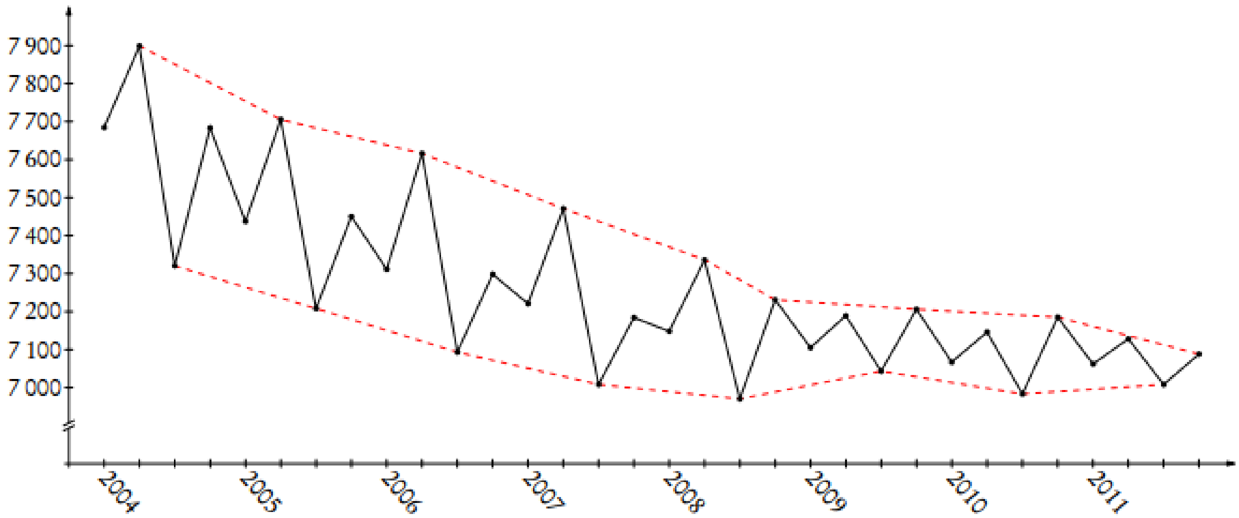
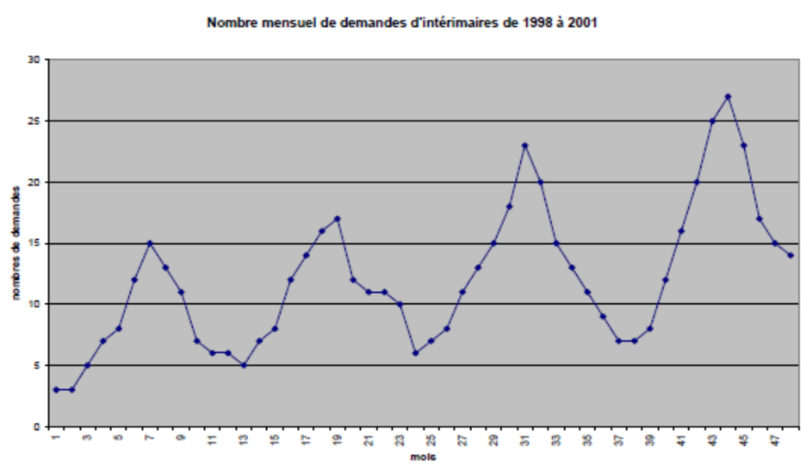


Fig. 5: Nombre trimestriel de naissances

Exemple 10



1.3.3 Schéma additif ou multiplicatif

Afin de faire la distinction entre schéma multiplicatif et additif, on peut se baser sur une méthode graphique.

On représente graphiquement les droites passant respectivement par les minimas et les maximas. On obtient deux cas de figures :

Schéma additif : (courbes parallèles).

Graphiquement, l'amplitude des variations est constante autour de la tendance

Schéma multiplicatif : courbes qui s'écartent (ou se rejoignent).

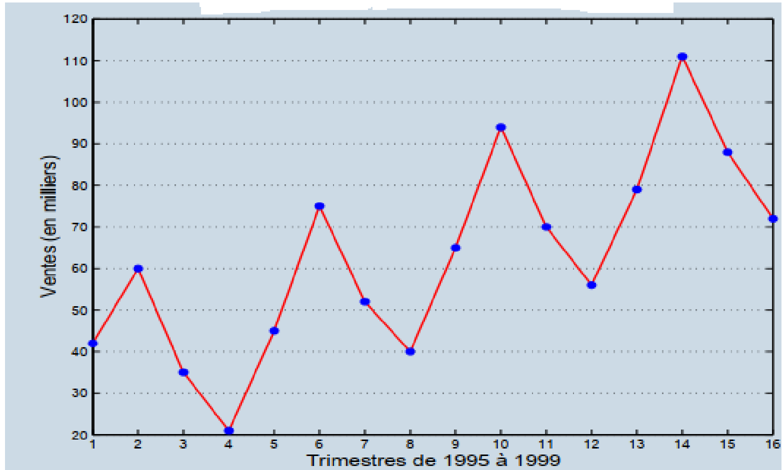


Fig. 6: Modèle Additif - Vente d'un Produit P

Les variations saisonnières augmentent (ou diminuent) avec la saison

1.3.4 Transformatios des modèles multiplicatifs en modèles additifs

Dans le cas d'une série (Y_t) à valeurs positives, le modèle multiplicatif se ramène à un modèle additif.

$$\begin{aligned} Y_t &= T_t \times S_t \times A_t \Rightarrow \log(Y_t) = \log(T_t \times S_t \times A_t) \\ &\Leftrightarrow \log(Y_t) = \log(T_t) + \log(S_t) + \log(A_t). \end{aligned}$$

Si on écrit $\log(Y_t) = Y'_t$, $\log(T_t) = T'_t$, $\log(S_t) = S'_t$, $\log(A_t) = A'_t$, le modèle s'écrit de façon additive

$$Y'_t = T'_t + S'_t + A'_t.$$

Donc, le travaille sur les logarithmes suppose que le modèle est additif.

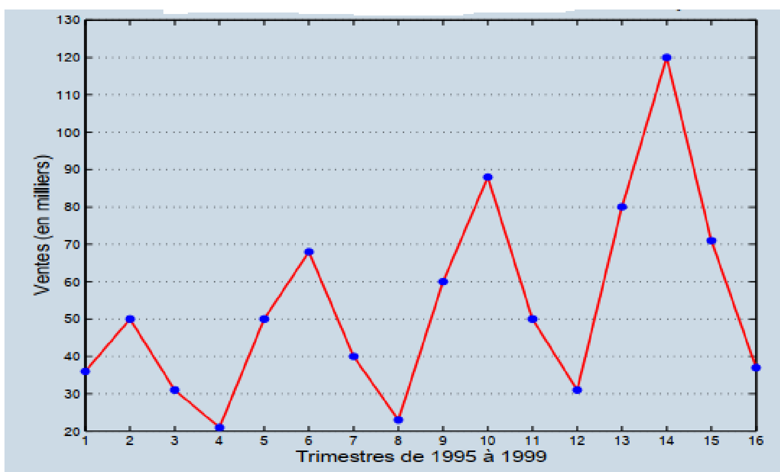


Fig. 7: Modèle Multiplicatif - Vente d'un Produit Q

Chapitre II

ESTIMATION DE LA TENDANCE.

2.1 Analyse de la tendance à long terme

On dispose d'une serie chronologique $(y_i)_{i=1,\ldots,n}$

Objectifs : trouver une fonction simple du temps qui modélise au mieux la tendance de la serie $(y_i)_{i=1,\ldots,n}$

Types de méthodes

- Par la méthode graphique
- Par la méthode non paramétrique (**Moyennes mobiles**)
- Par la méthode des **semi moyennes**
- Par la méthode paramétrique **(La méthode des moindes carrés).**

2.1.1 Méthode graphique

On représente tout d'abord sur un graphique les livraisons trimestrielles (sur l'axe des ordonnées) en fonction des différents trimestres (sur l'axe des abscisses)

L'objectif de l'analyse graphique est d'exhiber du graphe un mouvement, une tendance générale afin de préciser la maniére dont les livraisons progressent au cours du temps. La période d'observation est ici de quatre années.

La procédure est ensuite la suivante :

- On joint les maxima M_i à la règle,
- On joint les minima m_i à la règle,
- On définit les projections orthogonales M_i' des maxima M_i sur les segments $[m_i,\ m_{i+1}]$,
- On définit les projections orthogonales m_i' des minima m_i sur les segments

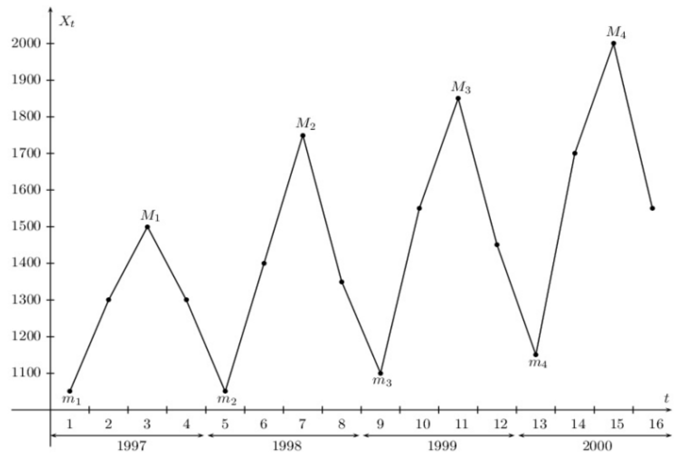


Fig. 8: Représentation graphique

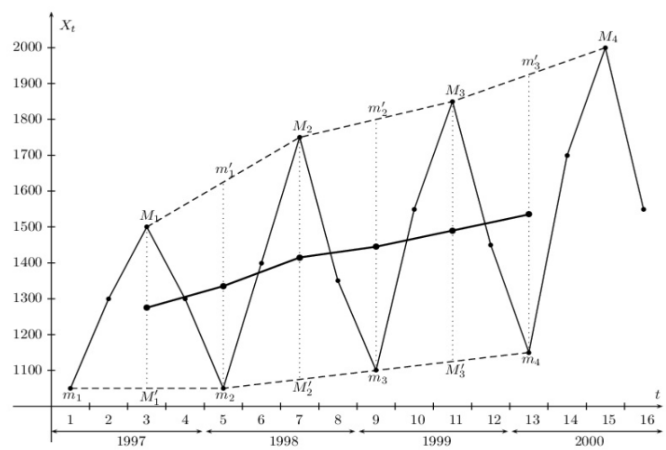


Fig. 9: Tendance - Méthode graphique

$[M_i, M_{i+1}]$,

- On trace les segments verticaux $[M_i, M'_i]$ et $[m_i, m'_i]$.

La ligne brisé qui joint les milieux des segments $[M_i, M'_i]$ et $[m_i, m'_i]$ constitue une tendance (voir figure 2.2)

2.2 Méthode de la moyenne mobile

– Les moyennes mobiles d'ordre k, k désignant un entier supérieur ou égal à 2, sont les moyennes (arithmétiques) :

▶ des cycles de k observations consécutives, lorsque k est impair ;

▶ des cycles de $(k+1)$ observations consécutives, lorsque k est pair, avec une pondération moitié pour les valeurs extrême

t	y_t	moyenne mobile d'ordre 2	moyenne mobile d'ordre 3	moyenne mobile d'ordre 4
1	y_1			
2	y_2	$\frac{1}{2}\left(\frac{y_1}{2}+y_2+\frac{y_3}{2}\right)$	$\frac{1}{3}\left(y_1+y_2+y_3\right)$	
3	y_3	$\frac{1}{2}\left(\frac{y_2}{2}+y_3+\frac{y_4}{2}\right)$	$\frac{1}{3}\left(y_2+y_3+y_4\right)$	$\frac{1}{4}\left(\frac{y_1}{2}+y_2+y_3+y_4+\frac{y_5}{2}\right)$
4	y_4	$\frac{1}{2}\left(\frac{y_3}{2}+y_4+\frac{y_5}{2}\right)$	$\frac{1}{3}\left(y_3+y_4+y_5\right)$	$\frac{1}{4}\left(\frac{y_2}{2}+y_3+y_4+y_5+\frac{y_6}{2}\right)$
5	y_5	$\frac{1}{2}\left(\frac{y_4}{2}+y_5+\frac{y_6}{2}\right)$	$\frac{1}{3}\left(y_4+y_5+y_6\right)$	
6	y_6			

2.3 Estimation de la tendance par la méthode de la moyenne mobile .

2.3.1 Méthode des moyennes mobiles

Lorsque la tendance est linéaire, la méthode la plus simple et la plus utilisée est le lissage par moyenne mobile.

La moyenne mobile centrée d'ordre p est notée $M_p(t)$

– Si p est impair (c'est-à-dire $p = 2m+1$), alors :

$$M_p(t) = \frac{1}{p}\left[y(t-m) + y(t-(m-1)) + ... + y(t-(m+1)) + y(t+m)\right].$$

– Si p est pair (c'est-à-dire $p = 2m$), alors :

$$M_p(t) = \frac{1}{p}\left[\frac{1}{2}y(t-m) + y(t-(m-1)) + ... + y(t-(m+1)) + \frac{1}{2}y(t+m)\right]$$

Illustration numérique :

t	1	2	3	4	5	6
y_t	84	123	165	108	103	137

Calculons : $M_3(t) = \frac{1}{3}[y(t-1) + y(t) + y(t+1)]$

Pour $t = 1$, on a $M_3(1)$ est non définie car $y(0)$ est non disponible

Pour $t = 2$, on a $M_3(2) = \frac{1}{3}[y(1) + y(2) + y(3)] = \frac{1}{3}(84 + 123 + 165) = 124$

Pour $t = 3$, on a $M_3(3) = \frac{1}{3}[y(2) + y(3) + y(4)] = \frac{1}{3}(123 + 165 + 108) = 132$

Pour $t = 4$, on a $M_3(4) = \frac{1}{3}[y(3) + y(4) + y(5)] = \frac{1}{3}(165 + 108 + 103) = \frac{376}{3} = 125.33$

Pour $t = 5$, on a $M_3(5) = \frac{1}{3}[y(4) + y(5) + y(6)] = \frac{1}{3}(108 + 103 + 137) = 116$

Pour $t = 6$, on a :$M_3(6)$ est non-définie car $y(7)$ est non disponible

Calculons : $M_4(t) = \frac{1}{4}\left[\frac{1}{2}y(t-2) + y(t-1) + y(t) + y(t+1) + \frac{1}{2}y(t+2)\right]$, $t \in \{3.4\}$

Pour $t = 3$, on a $M_4(3) = \frac{1}{4}\left[\frac{1}{2}y(1) + y(2) + y(3) + y(4) + \frac{1}{2}y(5)\right] = 122.38$

Pour $t = 4$, on a

$$M_4(4) = \frac{1}{4}\left[\frac{1}{2}y(2) + y(3) + y(4) + y(5) + \frac{1}{2}y(6)\right] = 126.5$$

On a alors :

t	1	2	3	4	5	6
y_t	84	123	165	108	103	137
$M_3(t)$	–	124	132	125.3	116	–
$M_4(t)$	–	–	122.38	126.5	–	–

Exemple 11 *Premier cas : $k = 3$*

t	$y(t)$	$M_3(t)$
1	306	$\times$
2	344	**326.7**
3	333	**350**
4	373	**344.3**
5	327	**348.3**
6	345	**339.7**
7	347	**366**
8	406	**371.7**
9	362	**385**
10	387	**377**
11	**382**	**402**
12	437	$\times$

Démarche : pour $t = 11,$ *on obtient*

$\mathbf{\frac{1}{3}(387 + 382 + 437) = 402}$

Autrement dit,

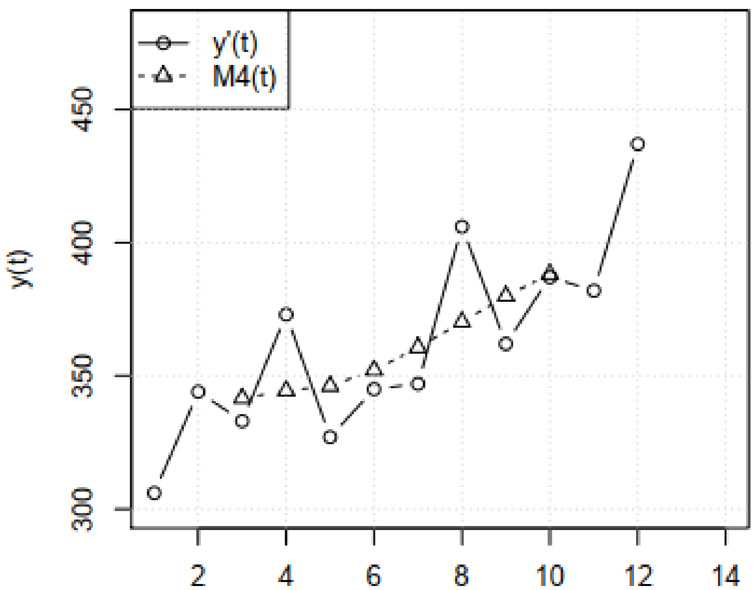
pour tout $t = 2, ..., 11$

$$M_3(t) = \frac{1}{3}[y(t-1) + y(t) + y(t+1)]$$

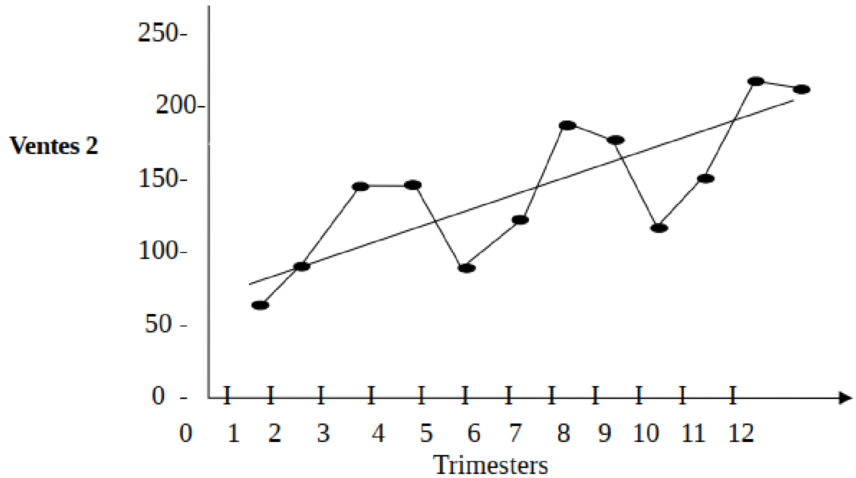
Second cas : $k = 4$

pour tout $t = 3, ..., 10,$

$$M_4(t) = \frac{1}{4}\left[\frac{1}{2}y(t-2) + y(t-1) + y(t) + y(t+1) + \frac{1}{2}y(t+2)\right]$$



t	$y(t)$	$M_4(t)$
1	306	$\times$
2	344	$\times$
3	333	**341.6**
4	373	**344.4**
5	327	**346.2**
6	345	**352.1**
7	347	**360.6**
8	406	**370.2**
9	362	**379.9**
10	387	**388.1**
11	382	$\times$
12	437	$\times$



Exercise 12 *Observons l'évolution des ventes d'une entreprise sur douze trimestres consécutifs :*

	Trimestre 1	*Trimestre 2*	*Trimestre 3*	*Trimestre 4*
Année $N-2$	66	96	145	144
Année $N-1$	92	131	195	189
Année N	120	167	246	239

Le graphique ci-dessous montre clairement une tendance générale à la hausse représentée par une droite de tendance (ou « trend ») avec des variations au-dessus ou en-dessous de cette tendance en fonction des saisons. Ainsi les premiers et seconds trimestres se situent plutôt endessous de la tendance alors que les troisième et quatrième se situent au-dessus.

2.4 Ajustement linéaire

Objectif

On utilise la méthode des moindres carrés pour ajuster la série chronologique (Yt, t) , avec la fonction

2.4.1

$$y(t) = at + b, \;\; a \in \mathbb{R}, \;\; b \in \mathbb{R}$$

La méthode des moindres carrés (Estimation paramétrique)

Cette méthode consiste à choisir les coefficients a et b de sorte que la quantité

$$\sum (y_i - (at_i + b))^2 = f(a, b)$$

soit minimale.

On cherche donc a et b solutions du système

$$S(a, b) \quad \begin{cases} \frac{\partial f}{\partial a}(a, b) = 0 \\ \frac{\partial f}{\partial b}(a, b) = 0 \end{cases}$$

On démontre que le couple solution est $(\hat{a}, \hat{b})$ avec

$$\hat{a} = \frac{\displaystyle\sum_{i=1}^{n} y_i t_i - n\bar{t}\bar{y}}{\displaystyle\sum_{i=1}^{n} t_i^2 - n\bar{t}^2}, \ \hat{b} = \bar{y} - \hat{a}\bar{t}.$$

oú

$$\bar{t} = \frac{\displaystyle\sum_{i=1}^{n} t_i}{n}, \ \bar{y} = \frac{\displaystyle\sum_{i=1}^{n} y_i}{n}.$$

La droite d'équation $y = \hat{a}t + \hat{b}$ s'appelle la droite des moindres carrés

Chapitre III

DÉCOMPOSITION D'UNE SÉRIE CHRONOLOGIQUE

3.1 Série désaisonnalisée ou série CVS

3.1.1 Schéma (ou modèle) additif

1. La série se décompose en composantes indépendantes les unes des autres. La composante saisonnière de la série est indépendante du mouvement de longue période (Variationsd'amplitudes égales)

$$Y_t = T_t + S_t.$$

On suppose que la variation résiduelle n'existe pas ou, ce qui revient au même, est intégrée dans le trend.

2. La tendance s'écrit $T(t) = at + b$ (seul l'ajustement linéaire est étudié).

Donc

$$Y_t = at + b + S_t.$$

Les coefficients a et b de l'équation du trend sont calculés par la méthode des moindres carré

3. Le calcul des variations saisonnières :

$$S_t = Y_t - T_t$$

– Les Y_t sont les **valeurs observées** (série brute), ou **série initiale**

– Les T_t sont **les valeurs calculées** à partir de l'équation du trend.

Exemple : Si la série comporte des données mensuelles sur 3 ans, on calcule $3 \times 12 = 36$ valeurs de S_t, si les données sont **trimestrielles**, $3 \times 4 = 12$ valeurs de

S_t.

4. Les coefficients saisonniers S_j

• On retient 12 valeurs de S_j (de S_1à S_{12}) si la série est mensuelle.

• On retient 4 valeurs de S_j (de S_1 à S_4) si la série est trimestrielle.

On obtient les coefficients S_j en calculant les **moyennes** des différences $Y_t - T_t$ correspondant à une même

saison

On calcule ensuite la moyenne $\bar{C}$ de ces coefficients S_j. Si le modèle est bien choisi, la moyenne

$$\bar{C} = \frac{\sum_{\mathbf{j=1}}^{\mathbf{p}} \mathbf{S_j}}{\mathbf{p}}$$

doit être proche de 0. Dans le cas contraire

5.On calcule les corrections (les coefficients saisonniers corrigés S'_j)

$$S'_j = S_j - \bar{C}$$

3.1.2 Série désaisonnalisée

Nous appelons série désaisonnalisée ou **série corrigée des variations saisonnières** notée série **CVS**, la série chronologique (Y_t^*, t) à laquelle on a enlevé les variations saisonnières.

Dans le cas du modèle additif :

$$Y_t^* = Y_t - S'_t$$

3.1.3 Schéma (ou modèle) multiplicatif

1.La série se décompose en composantes dépendantes les unes des autres. La composante saisonnière de la série est proportionnelle au mouvement de longue période.

$$Y_t = T_t \times S_t.$$

On suppose que la variation résiduelle n'existe pas ou, ce qui revient au même, est intégrée dans

le trend.

2.On détermine **la tendance** T_t par la droite de **régression** ou par la méthode des **moyennes mobiles**,

Donc

$$Y_t = (at + b) \times S_t$$

Les coefficients a et b de l'équation du trend sont calculés par la méthode des moindres carrés.

3.$\forall t$ on calcule les variations saisonnières (appelé rapport de la tendance) :

$$S_t = \frac{Y_t}{T_t}.$$

- Les Y_t sont les valeurs observées (série brute),
- Les T_t sont les valeurs calculées à partir de l'équation du trend.

4. On calcule **Les coefficients saisonniers** S_j

Si le modèle est bien choisi, les **moyennes** $\bar{\mathbf{C}} = \frac{\sum_{\mathbf{j=1}}^{\mathbf{p}} \mathbf{S_j}}{\mathbf{p}} = \mathbf{1}$ où p est le nombre de coefficients saisonniers.

5. Dans le cas contraire Les **coefficients saisonniers corrigés** S'_j $\left(\bar{\mathbf{C}} \neq \mathbf{1}\right)$ sont calculées par

$$S'_j = \frac{S_j}{\bar{\mathbf{C}}}$$

3.1.4 Série désaisonnalisée

Nous appelons série **désaisonnalisée** ou **série corrigée des variations saisonnières** notée série **CVS**, la série chronologique (Y_t^*, t) définit par

$$Y_t^* = \frac{Y_t}{S'_j}$$

3.1.5 Application aux des saionnalisations .

Soit les coefficients saisonniers suivants :

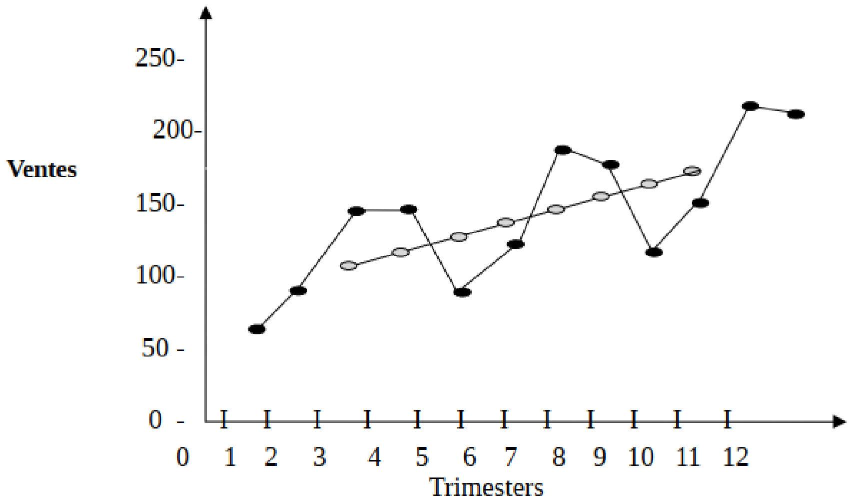
Trimestre 1	Trimestre 2	Trimestre 3	Trimestre 4
$0,7$	$0,9$	$1,25$	$1,15$

L'établissement de la série corrigée des variations saisonnières se présenterait ainsi :

Trimestre x	1	2	3	4	5	6
Données observées : y_i	66	96	145	144	92	131
Coefficient saisonnier : C_t	0.7	0.9	1.25	1.15	0.7	0.9
Série corrigée des variations saisonnières $y'_i = \frac{y_i}{C_t}$	94	107	116	125	131	146
Trimestre x	7	8	9	10	11	12
Données observées : y_i	195	189	120	167	246	239
Coefficient saisonnier : C_t	1.25	1.15	0.7	0.9	1.25	1.15
Série corrigée des variations saisonnières $y'_i = \frac{y_i}{C_t}$	156	164	171	186	197	208

En reportant cette série corrigée des variations saisonnières sur un graphique, on voit qu'elle présente une évolution beaucoup plus régulière que les données brutes.

En procédant à un ajustement sur cette série, on pourra avoir une prévision de la tendance des ventes et passer de là aux prévisions par trimestre.



APPLICATION AUX PREVISIONS.

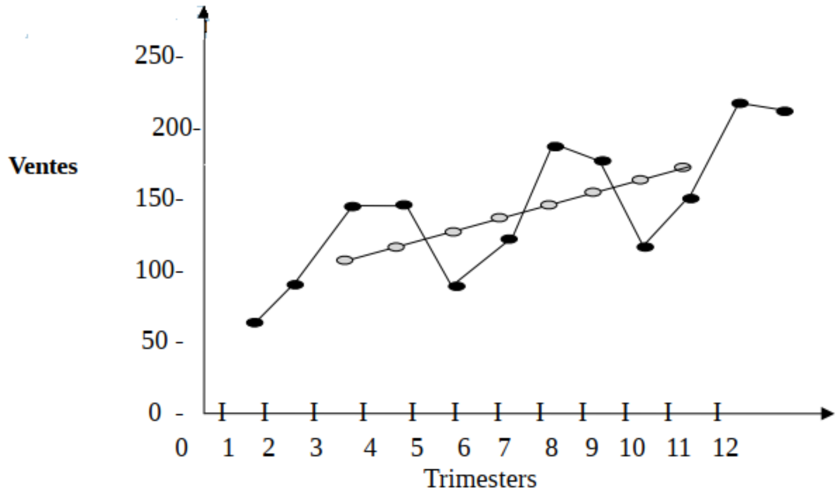
La droite d'ajustement de y' a pour équation $y' = 10,095x + 84,72$.

Afin de prévoir les ventes de l'année $N+1$, donc des trimestres $13, 14, 15, 16$, on procédera aux calculs suivants :

Trimestre x	13	14	15	16
Prévision de la tendance $y' = 10,095x + 84,72$.	216	226	236	246
Coefficient saisonnier : C_t	$0,7$	$0,9$	$1,25$	$1,15$
Prévision des ventes par trimestre $y = y' \times C_t$	151	203	295	283

3.1.6 Détermination de la tendance à partir des moyennes mobiles

Afin d'éliminer l'effet des variations saisonnières, on peut remplacer chaque valeur de la série par une moyenne calculée sur des valeurs représentant un ensemble correspondant à une durée d'une année centrée sur la date de chaque valeur. Ainsi pour des données trimestrielles, on calculera :



$$y_i^{,} = \frac{\frac{y_{i-2}}{2} + y_{i-1} + y_{i+1} + \frac{y_{i+2}}{2}}{4}$$

En coefficientant y_{i-2} et y_{i+2} par $1/2$, chaque trimestre de l'année est représenté dans le calcul avec la même pondération et la valeur calculée est centrée sur la période i.

– **APPLICATION A L'EXEMPLE**

Trimestre	1	2	3	4	5	6	7	8	9	10	11	12
Données observées : y_i	66	96	145	144	92	131	195	189	120	167	246	239
Moyenne mobile : y_i'			116	124	134	146	155	163	174	187		

Par exemple :

$$y_3' = \frac{\frac{66}{2} + 96 + 145 + 144 + \frac{92}{2}}{4} = 116$$

Les moyennes mobiles ainsi calculées permettent de représenter la tendance.

En ajustant cette tendance, par exemple par la méthode des moindres carrées, on pourra comme précédemment l'utiliser à des fin de prévisions

3.1.7 Application aux coefficients saisonniers

Trimestre	1	2	3	4	5	6	7	8	9	10	11
Données observées : y_i	66	96	145	144	92	131	195	189	120	167	246
Moyenne mobile : $y'i$			116	124	134	146	155	163	174	187	
Coefficients saisonniers : $\frac{y_i}{y'i}$			1.25	1.16	0.69	0.90	1.26	1.16	0.69	0.98	

On peut rapprocher les coefficients saisonniers relatifs à un même trimestre sur plusieurs années afin de déterminer les valeurs retenues définitivement. On retiendra pour chaque trimestre la moyenne des valeurs trouvées en procédant à une approximation éventuelle pour que la somme des valeurs soit bien égale à 4

	Trimestre 1	Trimestre 2	Trimestre 3	Trimestre 4
Année N-2			$1,25$	$1,16$
Année N-1	$0,69$	$0,90$	$1,26$	$1,16$
Année N	$0,69$	$0,89$		
Valeur retenue	$0,69$	$0,90$	$1,25$	$1,16$

Exercise 13 *L'entreprise CAROT vous fournit les informations suivantes relatives à ses ventes au cours des périodes* 1 *à* 7*.*

x_i	y_i
1	120
2	155
3	182
4	202
5	220
6	235
7	240

Solution 14 (Calcul des paramètres) *a **et** b de l'équation d'ajustement par la méthode des moindres carrés :*

Calcul de $\bar{x}$ et de $\bar{y}$

x_i	y_i	$X_i = x_i - \bar{x}$	$Y_i = y_i - \bar{y}$	$X_i Y_i$	X_i^2
1	120	-3	$-73,42$	$220,26$	9
2	155	-2	$-38,42$	$76,84$	4
3	182	-1	$-11,42$	$11,42$	1
4	202	0	8.5	0	0
5	220	1	$26,58$	$26,58$	1
6	235	2	$41,58$	$83,16$	4
7	7240	3	$46,58$	$139,74$	9
$\sum x_i = 28$	$\sum y_i = 1354$	0	$0,06$	$\sum X_i Y_i = 558$	$\sum X_i^2 = 28$

– *D'où Soit une droite d'ajustement (la tendance) :*

$$\mathbf{y = 19,93x + 113,71}$$

– *La prévision pour la période 8 est :*

$$\mathbf{y_8 = 19,93(8) + 113,71 = 273,15}$$

3.1.8 Les Moyennes Mobiles

La méthode des moyennes mobiles est une technique de lissage des données. Son principe est de substituer une série de valeurs observées par leur moyenne. Cette moyenne est calculée en prenant par exemple, trois valeurs (nous dirons qu'il s'agit de moyennes mobiles d'ordre **3**), quatre valeurs (moyennes mobiles d'ordre 4) etc.

Lorsque la tendance est linéaire, la méthode la plus simple et la plus utilisée est le lissage par moyenne mobile.

D'où le modèle additif devient :

$$Y_t = mm_t + S_t + A_t. \ \forall t \in \{1, ..., T\}$$

Ou bien d'une manière générale

$$Y_{i,j} = mm_{i,j} + S_{i,j} + A_{i,j}. \forall i \in \{1, ..., n\}, \forall j \in \{1, ..., p\} \ .$$

Remarque 15 *Le terme $S_{i,j}$ caractérise la variation saisonnière à l'instant j de chaque période i : du trimestre j*

dans des données trimestrielles $(p = 4)$, du mois j dans des données mensuelles $(p = 12)$ etc...

Illustrons le principe de cette méthode grâce à l'exemple suivant :

Exercise 16 *Soit le chiffre d'affaires réalisé par l'entreprise Mail*

Trimestres	T_1	T_2	T_3	T_4
Année 1	100	125	135	110
Année 2	105	135	150	125
Année 3	115	160	175	140
Année 4	120	165	180	150

On vous demande de calculer :

1. *Les moyennes mobiles d'ordre* 4 $(MM4)$.
2. *De représenter les données brutes et les moyennes mobiles sur un même graphique*

Solution 17 (Calcul des moyennes mobiles :) *Si on choisit une périodicité d'ordre* 4, *on obtient les valeurs ajustées suivantes :*

Rang du trimestre x_i	*Chiffre d'affaires* y_i *en* €	*Moyennes mobiles d'ordre* 4 ($MM4$)
1	100	–
2	125	–
3	135	$117,5$
4	110	$118,75$
5	105	$121,25$
6	135	125
7	150	$128,75$
8	125	$131,2$
9	115	$137,5$
10	160	$143,75$
11	175	$147,5$
12	140	$148,75$
13	120	150
14	165	$151,25$
15	180	–
16	150	–

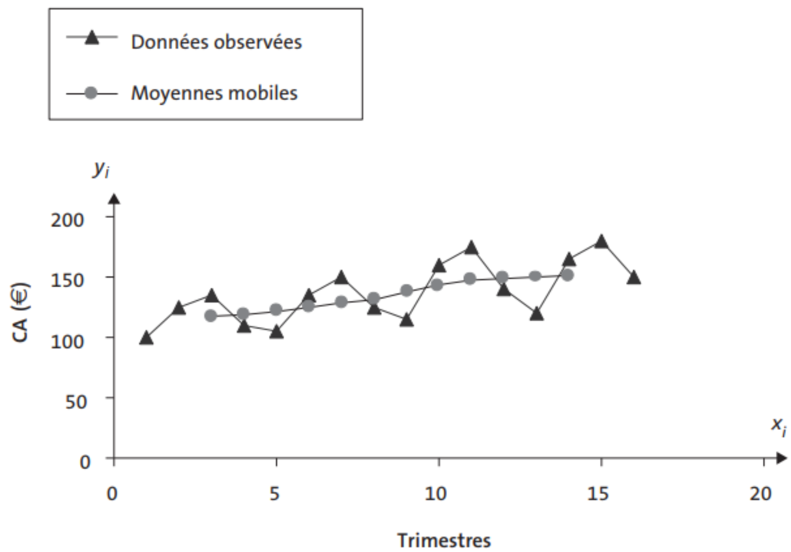


Fig. 10: Représentation des données ajustées :

Exercise 18 *On se donne les livraisons trimestrielles d'essence en super sans plomb* 98 *(en millions de m^3) dans un hypermarchéde la région pendant quatre années :*

Années / Trimestres	1	2	3	4
1997	1050	1300	1500	1300
1998	1050	1400	1750	1350
1999	1100	1550	1850	1450
2000	1150	1700	2000	1550

3.1.9 Ajustement linéaire par la méthode des moindes carrés

Recherchons l'équation de régression de Y_t dans le cadre de l'exemple . On utilisera donc le tableau de valeurs de la page suivante oú t_i représente le rang d'un trimestre et $Y_t(i)$ la livraison d'essence correspondante : On récupère les valeurs suivant

t_i	$Y(t)_i$	t_i^2	$Y(t)^2$	$Y(t)_i \times t_i$
1	1050	1	102500	1050
2	1300	4	1690000	2600
3	1500	9	2250000	4500
4	1300	16	1690000	5200
5	1050	25	1102500	5250
6	1400	36	1960000	840
7	1750	49	3062500	12250
8	1350	64	1822500	10800
9	1100	81	1210000	9900
10	1550	100	2402500	15500
11	1850	121	3422500	20350
12	1450	144	2102500	17400
13	1150	169	1322500	14950
14	1700	196	2890000	23800
15	2000	225	4000000	30000
16	1550	256	2402500	24800
136	23050	1496	34432500	206750

1. $\bar{t} = \frac{136}{16} = 8.5$
2. $X\bar{(t)} = \frac{23050}{16} = 1440;625$
3. $V(t) = \frac{1496}{16} - (8.5)^2 = 21.25$
4. $V(Y(t)) = \frac{34432500}{16} - (1440.625)^2 = 76630;859$
5. $Cov(t.Y(t)) = \frac{206750}{16} - (8.5 \times 1440.625) = 676.562$

$$
\begin{aligned}
a &= \frac{Cov(t, X(t))}{V(t)} \\
b &= \bar{Y}(t) - a\bar{t}
\end{aligned}
$$

soit

$$a = \frac{676.5625}{21.25} = 31.838$$

et

$$b = 1440.625 - 31.838 \times 8.5 = 1170.002.$$

Finalement, la droite admet pour équation

$$Y(t) = 31.838t + 1170.002$$

3.1.10 Correction des variations saisonnières ($Y_t = T_t \times S_t$)

t	1	2	3	4	5	6	7	8
Y_t	1050	1300	1500	1300	1050	1400	1750	1350
T_t	1201.84	1233.68	1265.52	1297.35	1329.19	1361.03	1392.87	1424.71
$s_t = \frac{Y_t}{T_t}$	0.87	1.05	1.19	1.00	0.79	1.03	1.26	0.95
t	9	10	11	12	13	14	15	16
Y_t	1100	1550	1850	1450	1150	1700	2000	1550
T_t	1456.54	1488.38	1520.22	1552.06	1583.90	1615.73	1647.57	1679.41
$s_t = \frac{Y_t}{T_t}$	0.76	1.04	1.22	0.93	0.73	1.05	1.21	0.92

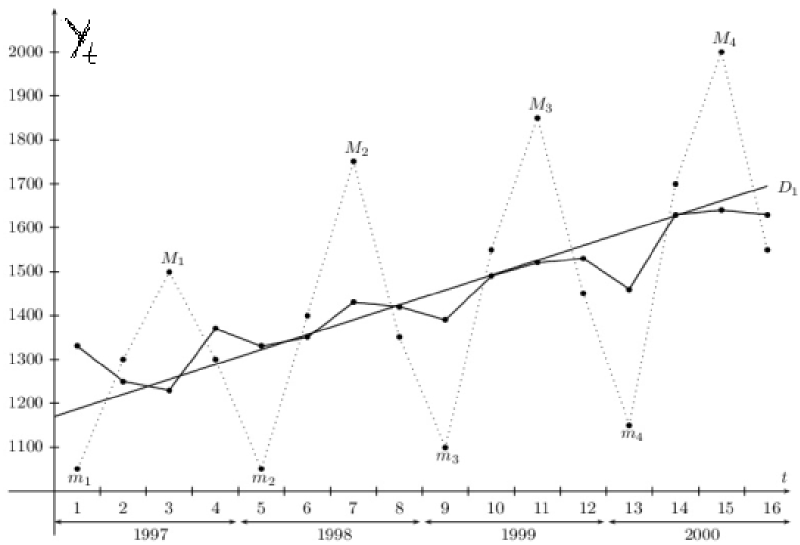
On calcule ensuite les coefficients saisonniers

$$\begin{aligned}
Coff_1 &= C_1 = \frac{1}{4}(s_1 + s_5 + s_9 + s_{13}) = \frac{1}{4}(0.87 + 0.79 + 0.76 + 0.73) = 0.79 \\
Coff_2 &= C_2 = \frac{1}{4}(s_2 + s_6 + s_{10} + s_{14}) = \frac{1}{4}(1.05 + 1.03 + 1.04 + 1.05) = 1.04 \\
Coff_3 &= C_3 = \frac{1}{4}(s_3 + s_7 + s_{11} + s_{15}) = \frac{1}{4}(1.19 + 1.26 + 1.22 + 1.21) = 1.22 \\
Coff_4 &= C_4 = \frac{1}{4}(s_4 + s_8 + s_{12} + s_{16}) = \frac{1}{4}(1.00 + 0.95 + 0.93 + 0.92) = 0.95
\end{aligned}$$

Comme $\bar{C} = \frac{1}{4}(C_1 + C_2 + C_3 + C_4) = \frac{1}{4}(0.79 + 1.04 + 1.22 + 0.95) = 1.$ le modèle est bien choisi. On se sert ensuite des coefficients saisonniers pour désaisonnaliser la série. On obtient le tableau suivant (Série désaisonnalisée) :

t	1	2	3	4	5	6	7	8
Y_t	1050	1300	1500	1300	1050	1400	1750	1350
$Y_t^\star$	1329.11	1250	1229.51	1368.42	1329.11	1346.15	1434.43	1421.05
t	9	10	11	12	13	14	15	16
Y_t	1100	1550	1850	1450	1150	1700	2000	1550
$Y_t^\star$	1392.41	1490.38	1516.39	1526.32	1455.70	1634.62	1639.34	1631.58

Voici pour terminer les représentations graphiques de la série originale, de la droite de régression et de la série désaisonnalisée



Exemple 19 (cas additif) *Considérons la série statistique des chiffres d'affaires trimestriels (en milliers d'euros) d'une entreprise.*

Trimestre	1	2	3	4	5	6	7	8	9	10	11	12	13	14	1
Chiffre d'affaires	120	181	71	119	128	190	73	124	140	196	84	133	145	206	9

Le modèle de la série statistique est bien additif, pour s'en convaincre, il suffit d'employer la méthode de la bande et de constater que la droite qui passe par les maxima et celle qui passe par les minima sont parallèles

Calculons les moyennes mobiles T_t *d'ordre* 4 *(puisque nous sommes en présence* *d'une série dont la périodicité est de 4 trimestres). Ensuite nous d´eterminons les différences saisonnières $d_t = Y_t - T_t$, c'est-à-dire les différences entre les valeurs de la série statistique de départ et les valeurs des moyennes mobiles correspondantes. Ces différences vont nous servir pour obtenir les coefficients saisonniers non corrigés trimestriels*

t	1	2	3	4	5	6	7	8
Y_t	120	181	71	119	128	190	73	124
T_t	$-$	$-$	$123,75$	$125,88$	$127,25$	$128,13$	$130,25$	$132,50$
d_t	$-$	$-$	$-52,75$	$-6,88$	$0,75$	$61,88$	$-57,25$	$-8,50$
t	9	10	11	12	13	14	15	16
Y_t	140	196	84	133	145	206	96	142
Tt	$134,63$	$137,13$	$138,88$	$140,75$	$143,50$	$146,13$	$-$	$-$
$d_t = Y_t - T_t$	$5,38$	$58,88$	$-54,88$	$-7,75$	$1,50$	$59,88$		

Les coefficients saisonniers correspondent aux moyennes des différences saisonnières pour chacun des trimestres. On obtient :

- $C_1 = \frac{1}{3}(d_5 + d_9 + d_{13}) = \frac{1}{3}(0.75 + 5.38 + 1.50) = 2.54$
- $C_2 = \frac{1}{3}(d_6 + d_{10} + d_{14}) = \frac{1}{3}(61.88 + 58.88 + 59.88) = 60.21$
- $C_3 = \frac{1}{3}(d_3 + d_7 + d_{11}) = \frac{1}{3}(-52.75 - 57.25 - 54.88) = -54.96$
- $C_4 = \frac{1}{3}(d_4 + d_8 + d_{12}) = \frac{1}{3}(-6.88 - 8.50 - 7.75) = -7.71$

Nous supposons que la composante saisonnière est strictement périodique. L'effet net de la composante saisonnière sur une période doit être nul car il est repris dans la tendance générale de la série chronologique.

Or ce n'est pas le cas dans cet exemple puisque

$$\bar{C} = \frac{C_1 + C_2 + C_3 + C_4}{4} = \frac{1}{4}(2.54 + 60.21 - 54.96 - 7.71) = 0.02$$

Ceci nous amène donc à rectifier les coefficients saisonniers non corrigés en leur retranchant la moyenne des coefficients saisonniers pour toutes les périodes. On a alors :

$C'_1 = C_1 - \bar{C} = 2.54 - 0.02 = 2.52$

$C'_2 = C_2 - \bar{C} = 60.21 - 0.02 = 60.19$

$C'_3 = C_3 - \bar{C} = -54.96 - 0.02 = 54.98$

$C'_4 = C_4 - \bar{C} = -7.71 - 0.02 = -7.73$

Disposant maintenant des coefficients saisonniers corrigés, nous pouvons désaisonnaliser la série chronologique en retranchant à chacune des valeurs initiales de la série la valeur du coefficient saisonnier correspondant.

t	1	2	3	4	5	6	7	8
Y_t	120	181	71	119	128	190	73	124
d_t	$117,48$	$120,81$	$125,98$	$126,73$	$125,48$	$129,81$	$127,98$	$131,73$
t	9	10	11	12	13	14	15	16
Yt	140	196	84	133	145	206	96	142
$Y'_t = Yt - C'_t)$	$137,48$	$135,81$	$138,98$	$140,73$	$142,48$	$145,81$	$150,98$	$149,73$