

## 3.1 Définitions des moyennes mobiles

### 3.1.1 Les moyennes mobiles simples

#### Définition (*Moyennes mobiles simples*)

On appelle **série des moyennes mobiles d'ordre  $k$**  de la série  $(t_i, y_i)_{1 \leq i \leq n}$ , la série des moyennes de  $k$  observations consécutives prenant ses valeurs aux dates moyennes correspondantes.

Plus précisément, cette série est définie par :

- la série des instants moyens  $(\bar{t}_i)_{1 \leq i \leq n-k+1}$  avec

$$\bar{t}_i = \frac{t_i + t_{i+1} + \cdots + t_{i+k-1}}{k} \quad \text{pour } i = 1, \dots, n - k + 1$$

- la série  $(\bar{y}_i)_{1 \leq i \leq n-k+1}$  des moyennes empiriques de  $k$  observations consécutives avec

$$\bar{y}_i = \frac{y_i + y_{i+1} + \cdots + y_{i+k-1}}{k} \quad \text{pour } i = 1, \dots, n - k + 1$$

## Remarque 1

On perd  $(k - 1)$  observations lorsqu'on construit la série des moyennes mobiles d'ordre  $k$ .

## Remarque 2

On limitera notre étude au cas où les instants d'observations sont équidistants. Par conséquent,

- si  **$k$  est impair** et vaut  $2m + 1$ , alors la série des moyennes mobiles d'ordre  $k$  est calculée aux instants d'observations

$$t_{m+1}, \dots, t_{n-m}, \text{ ie. } (t_j)_{j=m+1, \dots, n-m}$$

- si  **$k$  est pair** et vaut  $2m$ , alors la série des moyennes mobiles d'ordre  $k$  est calculée aux instants d'observations

$$\frac{t_m + t_{m+1}}{2}, \dots, \frac{t_{n-m} + t_{n-m+1}}{2}, \text{ ie. } \left( \frac{t_j + t_{j+1}}{2} \right)$$

## Définition (*Moyennes mobiles simples (bis)*)

On note  $MM(k)$  la série des moyennes mobiles d'ordre  $k$  de la série  $(y_j)_{j=1, \dots, n}$ , et on a

- **lorsque  $k$  est pair et vaut  $2m$  :**

$$MM(k)_j = \frac{\overbrace{y_{j-m+1} + \dots + y_j}^{m \text{ termes}} + \overbrace{y_{j+1} + \dots + y_{j+m}}^{m \text{ termes}}}{2m}$$

pour  $j = m, \dots, n - m$ .

- **lorsque  $k$  est impair et vaut  $2m + 1$  :**

$$MM(k)_j = \frac{y_{j-m} + \dots + y_j + \dots + y_{j+m}}{2m + 1}$$

pour  $j = m + 1, \dots, n - m$ .

### Remarque 3

Il est possible de calculer les moyennes mobiles simples d'ordre  $k$  de manière récursive. En effet,

- **lorsque  $k$  est pair et vaut  $2m$** , on calcule

- ▶  $MM(k)_m = \frac{y_1 + \cdots + y_{2m}}{2m}$

- ▶ puis, pour  $j = m, \cdots, n - m - 1$ ,

$$MM(k)_{j+1} = MM(k)_j + \frac{y_{j+m+1} - y_{j-m+1}}{2m}$$

- **lorsque  $k$  est impair et vaut  $2m + 1$** , on calcule

- ▶  $MM(k)_{m+1} = \frac{y_1 + \cdots + y_{2m+1}}{2m + 1}$

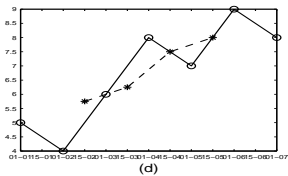
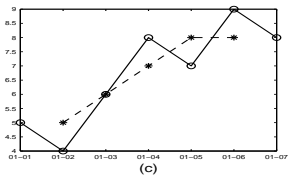
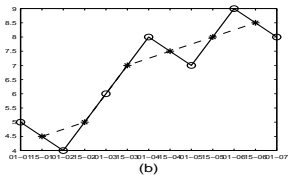
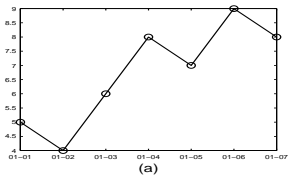
- ▶ puis, pour  $j = m + 1, \cdots, n - m - 1$ ,

$$MM(k)_{j+1} = MM(k)_j + \frac{y_{j+m+1} - y_{j-m}}{2m + 1}$$

## Exemple

Le tableau suivant présente les moyennes mobiles d'ordre 2, 3 et 4 d'une même série.

date $t_i$	date $\frac{t_i+t_{i+1}}{2}$	observ. $y_i$	$MM(2)$	$MM(3)$	$MM(4)$
01-01	15-01	5			
01-02	15-02	4			
01-03	15-03	6			
01-04	15-04	8			
01-05	15-05	7			
01-06	15-06	9			
01-07		8			



(a) la série observée ( $y_i, 1 \leq i \leq 7$ ), (b)  $MM(2)$ , (c)  $MM(3)$ , (d)  $MM(4)$

### Exercice (E.1)

Calculer les séries des moyennes-mobiles d'ordre 2, 3 et 4 de la série initiale ( $y_i$ ) suivante.

$t_i$	1	2	3	4	5	6	7	8
$y_i$	30	15	5	30	36	18	9	36

$t_i$	9	10	11	12	13	14	15	16
$y_i$	45	15	10	60	48	16	8	72

On complétera le tableau suivant.

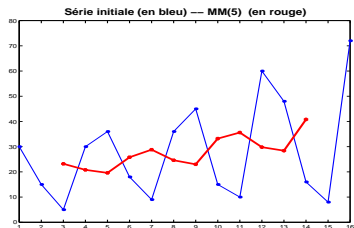
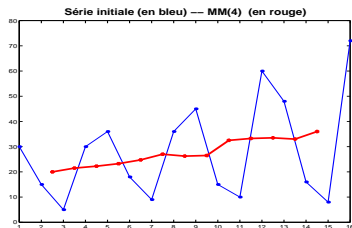
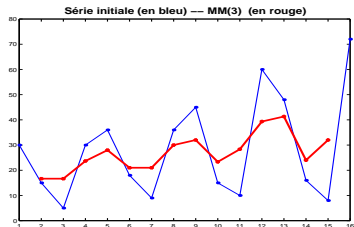
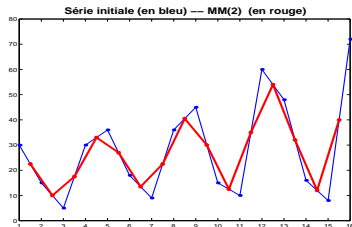
$(t_i)$	$(y_i)$	$\left(\frac{t_i+t_{i+1}}{2}\right)$	$MM(2)$	$MM(3)$	$MM(4)$
1	30	1.5			
2	15	2.5			
3	5				
4	30				
5	36				
6	18				
7	9				
8	36				
9	45				
10	15				
11	10				
12	60				
13	48				
14	16				
15	8				
16	72				



$(t_i)$	$(y_i)$	$\left(\frac{t_i+t_{i+1}}{2}\right)$	$MM(2)$	$MM(3)$	$MM(4)$
1	30	1.5	22.5		
2	15	2.5	10		
3	5	3.5	17.5		
4	30	4.5	33		
5	36	5.5	27		
6	18	6.5	13.5		
7	9	7.5	22.5		
8	36	8.5	40.5		
9	45	9.5	30		
10	15	10.5	12.5		
11	10	11.5	35		
12	60	12.5	54		
13	48	13.5	32		
14	16	14.5	12		
15	8	15.5	40		
16	72				

$(t_i)$	$(y_i)$	$\left(\frac{t_i+t_{i+1}}{2}\right)$	$MM(2)$	$MM(3)$	$MM(4)$
1	30	1.5	22.5		
2	15	2.5	10	16.67	
3	5	3.5	17.5	16.67	
4	30	4.5	33	23.67	
5	36	5.5	27	28	
6	18	6.5	13.5	21	
7	9	7.5	22.5	21	
8	36	8.5	40.5	30	
9	45	9.5	30	32	
10	15	10.5	12.5	23.33	
11	10	11.5	35	28.33	
12	60	12.5	54	39.33	
13	48	13.5	32	41.33	
14	16	14.5	12	24	
15	8	15.5	40	32	
16	72				

$(t_i)$	$(y_i)$	$\left(\frac{t_i+t_{i+1}}{2}\right)$	MM(2)	MM(3)	MM(4)
1	30	1.5	22.5		
2	15	2.5	10	16.67	20
3	5	3.5	17.5	16.67	21.5
4	30	4.5	33	23.67	22.25
5	36	5.5	27	28	23.25
6	18	6.5	13.5	21	24.75
7	9	7.5	22.5	21	27
8	36	8.5	40.5	30	26.25
9	45	9.5	30	32	26.5
10	15	10.5	12.5	23.33	32.5
11	10	11.5	35	28.33	33.25
12	60	12.5	54	39.33	33.5
13	48	13.5	32	41.33	33
14	16	14.5	12	24	36
15	8	15.5	40	32	
16	72				



### 3.1.2 Les moyennes mobiles centrées

Définition (*Moyennes mobiles centrées*)

On appelle **série des moyennes mobiles centrées d'ordre  $k$**  avec  $k = 2m$ , de la série  $(t_i, y_i)_{1 \leq i \leq n}$ , la série notée  $MMC(k)$  et définie par :

$$\begin{aligned} MMC(k)_j &= \frac{MM(k)_{j-1} + MM(k)_j}{2} \\ &= \frac{\frac{1}{2}y_{j-m} + y_{j-m+1} + \cdots + y_j + \cdots + y_{j+m-1} + \frac{1}{2}y_{j+m}}{2m} \end{aligned}$$

pour  $j = m + 1, \dots, n - m$ .

## Remarques

- ❶ Lorsque  $k = 2m$ , la série

$$\left( MMC(k)_j \right)_{j=m+1, \dots, n-m}$$

prend ses valeurs aux instants

$$(t_j)_{j=m+1, \dots, n-m}$$

alors que la série

$$\left( MM(k)_j \right)_{j=m, \dots, n-m}$$

prend ses valeurs aux instants

$$\left( \frac{t_j + t_{j+1}}{2} \right)_{j=m, \dots, n-m}$$

.

- ❷ La série  $MMC(k)$  comporte un terme de moins que la série  $MM(k)$ .

## Exercice (E.2)

Reprendre la série de l'exercice (E.1) et calculer les séries des moyennes-mobiles centrées d'ordre 2 et 4.

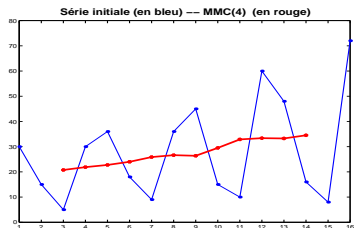
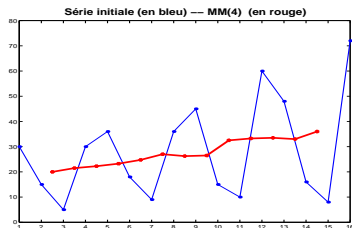
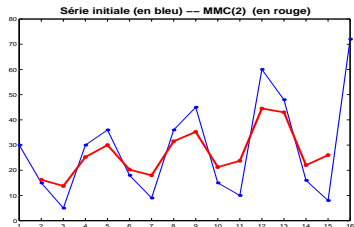
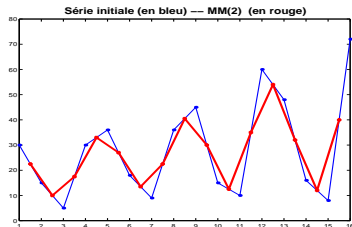
On complétera le tableau suivant.

$(t_i)$	$(y_i)$	$\left(\frac{t_i+t_{i+1}}{2}\right)$	MM(2)	MMC(2)	MM(3)	MM(4)	MMC(4)
1	30	1.5	22.5				
2	15	2.5	10		16.67	20	
3	5	3.5	17.5		16.67	21.5	
4	30	4.5	33		23.67	22.25	
5	36	5.5	27		28	23.25	
6	18	6.5	13.5		21	24.75	
7	9	7.5	22.5		21	27	
8	36	8.5	40.5		30	26.25	
9	45	9.5	30		32	26.5	
10	15	10.5	12.5		23.33	32.5	
11	10	11.5	35		28.33	33.25	
12	60	12.5	54		39.33	33.5	
13	48	13.5	32		41.33	33	
14	16	14.5	12		24	36	
15	8	15.5	40		32		
16	72						



$(t_i)$	$(y_i)$	$\left(\frac{t_i+t_{i+1}}{2}\right)$	MM(2)	MMC(2)	MM(3)	MM(4)	MMC(4)
1	30	1.5	22.5				
2	15	2.5	10	16.25	16.67	20	
3	5	3.5	17.5	13.75	16.67	21.5	
4	30	4.5	33	25.25	23.67	22.25	
5	36	5.5	27	30	28	23.25	
6	18	6.5	13.5	20.25	21	24.75	
7	9	7.5	22.5	18	21	27	
8	36	8.5	40.5	31.5	30	26.25	
9	45	9.5	30	35.25	32	26.5	
10	15	10.5	12.5	21.25	23.33	32.5	
11	10	11.5	35	23.75	28.33	33.25	
12	60	12.5	54	44.5	39.33	33.5	
13	48	13.5	32	43	41.33	33	
14	16	14.5	12	22	24	36	
15	8	15.5	40	26	32		
16	72						

$(t_i)$	$(y_i)$	$\left(\frac{t_i+t_{i+1}}{2}\right)$	MM(2)	MMC(2)	MM(3)	MM(4)	MMC(4)
1	30	1.5	22.5				
2	15	2.5	10	16.25	16.67	20	
3	5	3.5	17.5	13.75	16.67	21.5	20.75
4	30	4.5	33	25.25	23.67	22.25	21.875
5	36	5.5	27	30	28	23.25	22.75
6	18	6.5	13.5	20.25	21	24.75	24
7	9	7.5	22.5	18	21	27	25.875
8	36	8.5	40.5	31.5	30	26.25	26.625
9	45	9.5	30	35.25	32	26.5	26.375
10	15	10.5	12.5	21.25	23.33	32.5	29.5
11	10	11.5	35	23.75	28.33	33.25	32.875
12	60	12.5	54	44.5	39.33	33.5	33.375
13	48	13.5	32	43	41.33	33	33.25
14	16	14.5	12	22	24	36	34.5
15	8	15.5	40	26	32		
16	72						



## 3.2 Propriétés des moyennes mobiles

### 3.2.1 Elimination de la composante saisonnière

#### Propriété 1.

- 1 Si la série chronologique  $(y_i)_{i=1,\dots,n}$  possède une composante saisonnière de période  $p$ , alors l'application d'une moyenne mobile d'ordre  $p$  supprime cette saisonnalité.
- 2 La série  $MM(p)$  ou  $MMC(p)$  ne possède plus de composante saisonnière de période  $p$ .

#### Remarque

On se servira donc d'une moyenne mobile d'ordre  $p$  pour éliminer une composante saisonnière de période  $p$ .

## Vérification

Considérons une série purement périodique  $(s_i, 1 \leq i \leq n)$  de période  $p$ .

On a donc

- $s_1 + s_2 + \cdots + s_p = 0$ .
- pour tout  $j$ ,  $s_j = s_{j+p}$ .

Calculons maintenant la moyenne mobile d'ordre  $p$ .

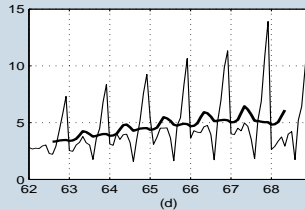
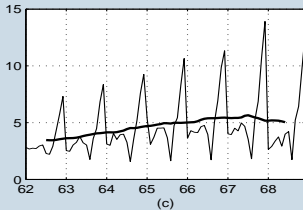
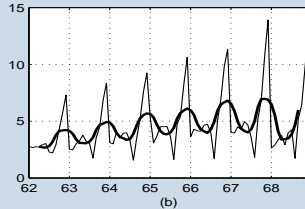
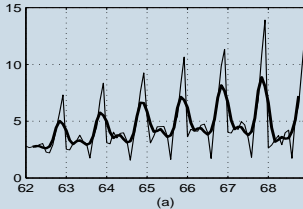
On a pour  $j = 1, \dots, n - p$  :

$$MM(p)_j = \frac{s_j + \cdots + s_{j+p-1}}{p}$$

et puisque la série est périodique de période  $p$ , alors (à une permutation circulaire près des  $p$  termes)

$$MM(p)_j = \frac{s_1 + \cdots + s_p}{p} = 0.$$

## Exemple (*Ventes mensuelles de champagne de 1962 à 1968*)



La série des Ventes mensuelles de Champagne de 1962 à 1968, avec en

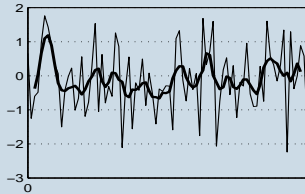
- (a), MM(4)
- (b), MM(6)
- (c), MM(12)
- (d), MM(15)

## 3.2.2 Atténuation des fluctuations irrégulières

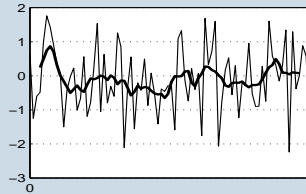
### Propriété 2.

- 1 Une moyenne mobile atténue l'amplitude des fluctuations irrégulières d'une chronique.
- 2 Plus l'ordre de la moyenne mobile est élevé, et plus cette atténuation est importante.
- 3 Donc en appliquant une moyenne mobile sur une série chronologique on obtient **un lissage** de la série.

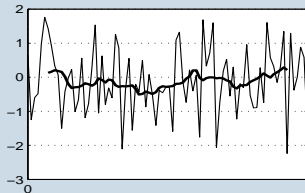
## Exemple (*Moyennes mobiles sur une série de fluctuations irrégulières*)



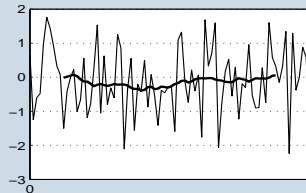
(a)



(b)



(c)



(d)

La série des fluctuations irrégulières, avec en

- (a),  $MM(4)$
- (b),  $MM(6)$
- (c),  $MM(12)$
- (d),  $MM(20)$



## Interprétation (*Probabiliste*)

- Soit  $(\varepsilon_i)_{i=1,\dots,n}$  une suite de variables aléatoires indépendantes de moyenne nulle et de variance  $\sigma^2$ .

Appliquons une  $MM(k)$  à cette suite : pour  $i = 1, \dots, n - k$ ,

$$MM(k)_i = \frac{1}{k} \sum_{j=1}^k \varepsilon_{i+j}$$

$$\text{et donc } \text{Var}(MM(k)_i) = \frac{1}{k^2} \sum_{j=1}^k \text{Var}(\varepsilon_{i+j}) = \frac{\sigma^2}{k}.$$

La variance est donc divisée par un facteur  $k$ , c'est-à-dire que l'écart-type est divisé par un facteur  $\sqrt{k}$ .

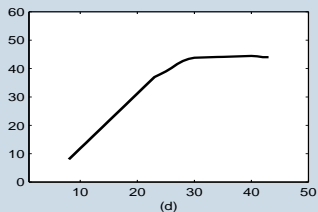
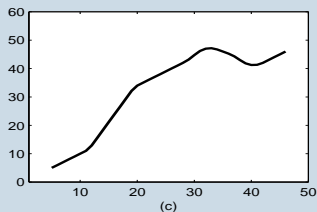
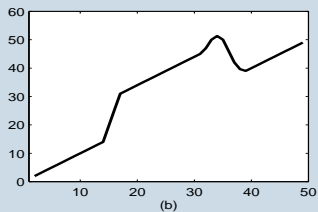
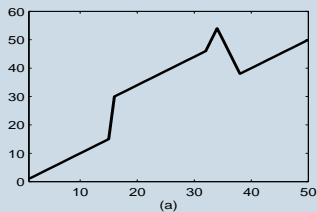
- Donc, si l'on considère une suite de fluctuations irrégulières  $(e_i)$  comme la réalisation d'une suite de variables aléatoires  $(\varepsilon_i)$  indépendantes, centrées et de même variance,  $MM(k)$  réduit l'amplitude de ces fluctuations irrégulières d'un facteur  $\sqrt{k}$ .

### 3.2.3 Effet d'une moyenne mobile sur la tendance

#### Propriété 3.

- 1 Une moyenne mobile (d'ordre quelconque) ne modifie pas une tendance constante.
- 2 Une moyenne mobile simple (d'ordre impair quelconque) ou centrée (d'ordre quelconque) ne modifie pas une tendance linéaire.
- 3 Lorsque la tendance n'est pas aussi régulière, les détails disparaîtront d'autant plus que l'ordre de la moyenne mobile sera grand.

## Exemple (*Application de moyennes mobiles sur une tendance non régulière*)



- (a), La tendance
- (b),  $MM(3)$
- (c),  $MM(8)$
- (d),  $MM(15)$

## 3.3 Choix pratique de l'ordre d'une moyenne mobile

**Rappel.** Le but d'un lissage par moyenne mobile est de faire apparaître l'allure de la tendance.

Lorsque des fluctuations périodiques et/ou irrégulières sont présentes, on aimerait par conséquent réaliser les trois objectifs suivants :

- ❶ Supprimer la composante périodique ;
- ❷ Réduire le plus possible l'amplitude des fluctuations irrégulières ;
- ❸ Et bien sûr, ne pas trop modifier la tendance !

## Remarque

Il est bien rare que ces trois objectifs puissent être parfaitement atteints simultanément. En effet, chacun de ces objectifs conduit à des choix différents de moyenne mobile :

- ❶ On supprime la composante périodique de période  $p$  avec une moyenne mobile d'ordre  $p$  ;
- ❷ On réduit l'amplitude des fluctuations irrégulières avec une moyenne mobile d'ordre élevé ;
- ❸ On préserve les détails de la tendance avec une moyenne mobile d'ordre faible.

**En pratique**, on doit trouver le meilleur compromis pour le choix de l'ordre du lissage optimal.

La solution consiste donc :

- à appliquer  $MM(k)$  ou  $MMC(k)$  pour supprimer la composante périodique dans un premier temps,
- puis à composer éventuellement différentes moyennes mobiles d'ordre peu élevé, jusqu'à l'obtention d'une série suffisamment lissée, mais telle que les détails intéressants de la tendance ne soient pas effacés.

## 4. Décomposition d'une Série Chronologique

On considèrera deux types de décomposition :

- la décomposition à partir d'un modèle additif
- la décomposition à partir d'un modèle multiplicatif

La démarche consistera à :

- identifier les coefficients saisonniers
- désaisonnaliser la série initiale, pour pouvoir ensuite ajuster une courbe de tendance
- construire la série lissée des prédictions en vue de faire de la prévision

## 4.1. Décomposition additive

### Objectif

Soit  $(y_i)_{i=1,\dots,n}$  une série chronologique.

On veut ajuster à cette série un modèle de la forme

$$y_i = f_i + s_i + e_i, \quad i = 1, \dots, n.$$

constitué

- de deux composantes déterministes :
  - ▶ la tendance ( $f_i$ )
  - ▶ la composante saisonnière ( $s_i$ ) périodique de période  $p$  telle que

$$\sum_{j=1}^p s_j = 0$$

- et d'une composante accidentelle ou bruit ( $e_i$ ) correspondant à un aléa.



## 4.1.1 Les étapes de la décomposition additive

On procèdera en suivant les étapes suivantes :

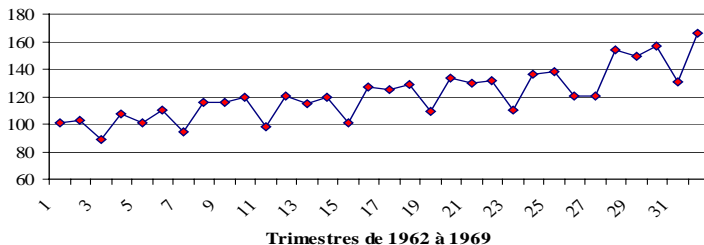
- ➊ La désaisonnalisation (5 étapes)
  - ➊ Lissage par moyennes mobiles
  - ➋ Construction de la série des différences
  - ➌ Calcul des coefficients saisonniers non centrés
  - ➍ Centrage des coefficients saisonniers
  - ➎ Construction de la série corrigée des variations saisonnières
- ➋ La série lissée des prévisions (2 étapes)
  - ➊ Ajustement d'une tendance
  - ➋ Construction de la série lissée des prévisions

## Exemple (*Fil rouge*)

On veut étudier la série chronologique de l'**Indice trimestriel de la production industrielle base 100 en 1962**. Source INSEE.

Années	Trimestre 1	Trimestre 2	Trimestre 3	Trimestre 4
1962	101.3	102.9	88.4	107.3
1963	101	109.8	94.1	116.1
1964	115.6	119.2	97.7	120.3
1965	115.1	119.5	101.1	127.4
1966	124.8	129	109.3	133.6
1967	129.4	131.8	110.2	136.4
1968	138.5	120.1	120.8	154.4
1969	149.5	157.1	130.8	166.5

**Indice trimestriel de la production industrielle base 100  
en 1962. Source INSEE.**



## 4.1.2 La désaisonnalisation : les différentes étapes.

### Etape 1. (*Lissage par Moyennes Mobiles*)

On effectue **un lissage** par la méthode des Moyennes Mobiles afin d'obtenir une première évaluation de la tendance de la série.

On notera  $(\tilde{f}_j)$  la série obtenue.

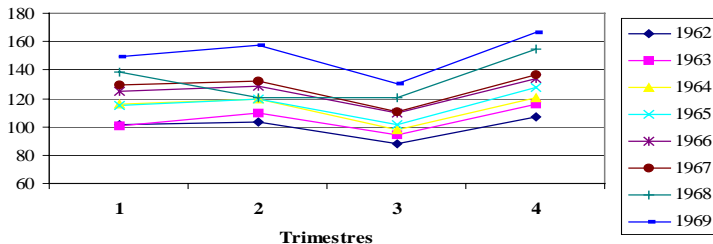
Si  $p$  désigne la période de la composante saisonnière, on aura

$$\tilde{f}_j = \begin{cases} MMC(p)_j & \text{si } p \text{ est pair} \\ MM(p)_j & \text{si } p \text{ est impair} \end{cases}$$

**Remarque.** Dans de nombreuses applications, la période  $p$  est paire et vaut  $2m$  (série mensuelle, trimestrielle, bi-mensuelle, semestrielle, etc). Avec le lissage, on perd  $m$  observations au début de la série et  $m$  observations à la fin, soit au total  $p = 2m$  observations.

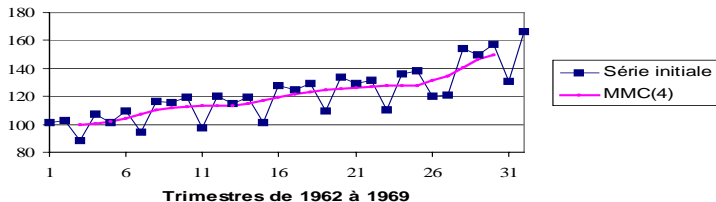
# Application sur l'exemple

**Indice trimestriel de la production industrielle base 100  
en 1962. Source INSEE.**



Date	Y_j	MM(4)	MMC(4)
1	101.3		
2	102.9		
3	88.4	99.975	99.9375
4	107.3	99.9	100.7625
5	101	101.625	102.3375
6	109.8	103.05	104.15
7	94.1	105.25	107.075
8	116.1	108.9	110.075
9	115.6	111.25	111.7
10	119.2	112.15	112.675
11	97.7	113.2	113.1375
12	120.3	113.075	113.1125
13	115.1	113.15	113.575
14	119.5	114	114.8875
15	101.1	115.775	116.9875
16	127.4	118.2	119.3875
17	124.8	120.575	121.6
18	129	122.625	123.4
19	109.3	124.175	124.75
20	133.6	125.325	125.675
21	129.4	126.025	126.1375
22	131.8	126.25	126.6
23	110.2	126.95	128.0875
24	136.4	129.225	127.7625
25	138.5	126.3	127.625
26	120.1	128.95	131.2
27	120.8	133.45	134.825
28	154.4	136.2	140.825
29	149.5	145.45	146.7
30	157.1	147.95	149.4625
31	130.8	150.975	
32	166.5		

**Indice trimestriel de la production industrielle  
base 100 en 1962. Source INSEE.**



## Etape 2. (*Série des différences*)

On calcule **la série des différences**, soit

$$D_i = y_i - \tilde{f}_i$$

pour les valeurs de  $i$  disponibles, soit

$$i = m + 1, \dots, n - m.$$



Date	Y_j	MMC(4)	D_j
1	101.3		
2	102.9		
3	88.4	99.9375	-11.5375
4	107.3	100.7625	6.5375
5	101	102.3375	-1.3375
6	109.8	104.15	5.65
7	94.1	107.075	-12.975
8	116.1	110.075	6.025
9	115.6	111.7	3.9
10	119.2	112.675	6.525
11	97.7	113.1375	-15.4375
12	120.3	113.1125	7.1875
13	115.1	113.575	1.525
14	119.5	114.8875	4.6125
15	101.1	116.9875	-15.8875
16	127.4	119.3875	8.0125
17	124.8	121.6	3.2
18	129	123.4	5.6
19	109.3	124.75	-15.45
20	133.6	125.675	7.925
21	129.4	126.1375	3.2625
22	131.8	126.6	5.2
23	110.2	128.0875	-17.8875
24	136.4	127.7625	8.6375
25	138.5	127.625	10.875
26	120.1	131.2	-11.1
27	120.8	134.825	-14.025
28	154.4	140.825	13.575
29	149.5	146.7	2.8
30	157.1	149.4625	7.6375
31	130.8		
32	166.5		

### Etape 3. (*Calcul des coefficients saisonniers non centrés*)

On calcule les  $p$  coefficients saisonniers non centrés ( $\tilde{s}_j$ ).

Pour cela, on effectue des moyennes avec les valeurs de la série  $(D_j)_{m+1 \leq j \leq n-m}$ .

Pour simplifier l'écriture, supposons que  $n$  soit multiple de la période  $p$  de la saisonnalité de telle sorte que l'on dispose de  $K_0$  périodes complètes d'observations, ie.  $n = K_0 * p$ .

Par exemple si la saisonnalité est annuelle  $p = 12$  et que la série est observée sur  $n = 36$  mois, on a  $K_0 = 3$  périodes complètes.

- Pour  $j \geq m$ , 
$$\tilde{s}_j = \frac{1}{K_0 - 1} \sum_{k=1}^{K_0-1} D_{j+p(k-1)}$$
- Pour  $j \leq m$ , 
$$\tilde{s}_j = \frac{1}{K_0 - 1} \sum_{k=1}^{K_0-1} D_{j+pk}$$

## Application sur l'exemple

Ici,  $n = 32$  et  $p = 4$ , donc  $K_0 = 8$  et  $m = 2$ .

Par conséquent,

$$\tilde{s}_3 = \frac{1}{7} \sum_{k=1}^7 D_{3+4(k-1)} = \frac{1}{7} (D_3 + D_7 + D_{11} + D_{15} + D_{19} + D_{23} + D_{27})$$

$$\tilde{s}_4 = \frac{1}{7} \sum_{k=1}^7 D_{4+4(k-1)} = \frac{1}{7} (D_4 + D_8 + D_{12} + D_{16} + D_{20} + D_{24} + D_{28})$$

et

$$\tilde{s}_1 = \frac{1}{7} \sum_{k=1}^7 D_{1+4k} = \frac{1}{7} (D_5 + D_9 + D_{13} + D_{17} + D_{21} + D_{25} + D_{29})$$

$$\tilde{s}_2 = \frac{1}{7} \sum_{k=1}^7 D_{2+4k} = \frac{1}{7} (D_6 + D_{10} + D_{14} + D_{18} + D_{22} + D_{26} + D_{30})$$

## Application sur l'exemple (*suite*)

On trouve :

- $\tilde{s}_3 = -14.7428571$

- $\tilde{s}_4 = 8.27142857$

et

- $\tilde{s}_1 = 3.46071429$

- $\tilde{s}_2 = 3.44642857$

Date	Y_j	MMC(4)	D_j	Stild_j
1	101.3			3.46071429
2	102.9			3.44642857
3	88.4	99.9375	-11.5375	-14.7428571
4	107.3	100.7625	6.5375	8.27142857
5	101	102.3375	-1.3375	
6	109.8	104.15	5.65	
7	94.1	107.075	-12.975	
8	116.1	110.075	6.025	
9	115.6	111.7	3.9	
10	119.2	112.675	6.525	
11	97.7	113.1375	-15.4375	
12	120.3	113.1125	7.1875	
13	115.1	113.575	1.525	
14	119.5	114.8875	4.6125	
15	101.1	116.9875	-15.8875	
16	127.4	119.3875	8.0125	
17	124.8	121.6	3.2	
18	129	123.4	5.6	
19	109.3	124.75	-15.45	
20	133.6	125.675	7.925	
21	129.4	126.1375	3.2625	
22	131.8	126.6	5.2	
23	110.2	128.0875	-17.8875	
24	136.4	127.7625	8.6375	
25	138.5	127.625	10.875	
26	120.1	131.2	-11.1	
27	120.8	134.825	-14.025	
28	154.4	140.825	13.575	
29	149.5	146.7	2.8	
30	157.1	149.4625	7.6375	
31	130.8			
32	166.5			

#### Etape 4. (*Centrage des coefficients saisonniers.*)

On calcule la moyenne arithmétique des  $p$  coefficients saisonniers obtenus à l'étape 3 :

$$\bar{s} = \frac{1}{p} \sum_{j=1}^p \tilde{s}_j.$$

Puis on centre les coefficients saisonniers  $(\tilde{s}_j)_{j=1,\dots,p}$  en retranchant la moyenne  $\bar{s}$  et on obtient :

$$\hat{s}_j = \tilde{s}_j - \bar{s}, \quad \text{pour } j = 1, \dots, p.$$

## Application sur l'exemple

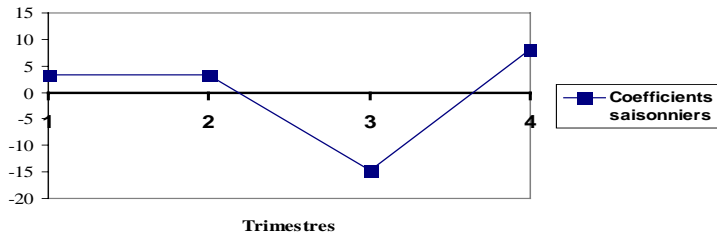
Sur l'exemple, on a

$$\begin{aligned}\bar{s} &= \frac{1}{4} \sum_{j=1}^4 \tilde{s}_j = \frac{1}{4} (\tilde{s}_1 + \tilde{s}_2 + \tilde{s}_3 + \tilde{s}_4) \\ &= 0.10892857\end{aligned}$$

et on trouve :

- $\hat{s}_1 = 3.35178571$
- $\hat{s}_2 = 3.3375$
- $\hat{s}_3 = -14.8517857$
- $\hat{s}_4 = 8.1625$

**Indice trimestriel de la production industrielle base 100  
en 1962. Source INSEE.**





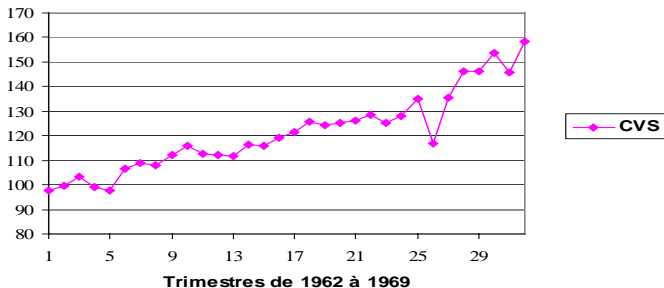
## Etape 5. (*La série corrigée des variations saisonnières*)

La série corrigée des variations saisonnières ou la série désaisonnalisée notée  $(CVS_i)_{1 \leq i \leq n}$  s'obtient en retranchant à la série initiale  $(y_i)_{1 \leq i \leq n}$ , la suite des coefficients saisonniers centrés obtenus à l'étape 4 :

$$CVS_j = y_j - \hat{s}_j, \quad \text{pour } j = 1, \dots, n.$$

Date	Y	Schap_j	CVS
1	101.3	3.35178571	97.9482143
2	102.9	3.3375	99.5625
3	88.4	-14.8517857	103.251786
4	107.3	8.1625	99.1375
5	101	3.35178571	97.6482143
6	109.8	3.3375	106.4625
7	94.1	-14.8517857	108.951786
8	116.1	8.1625	107.9375
9	115.6	3.35178571	112.248214
10	119.2	3.3375	115.8625
11	97.7	-14.8517857	112.551786
12	120.3	8.1625	112.1375
13	115.1	3.35178571	111.748214
14	119.5	3.3375	116.1625
15	101.1	-14.8517857	115.951786
16	127.4	8.1625	119.2375
17	124.8	3.35178571	121.448214
18	129	3.3375	125.6625
19	109.3	-14.8517857	124.151786
20	133.6	8.1625	125.4375
21	129.4	3.35178571	126.048214
22	131.8	3.3375	128.4625
23	110.2	-14.8517857	125.051786
24	136.4	8.1625	128.2375
25	138.5	3.35178571	135.148214
26	120.1	3.3375	116.7625
27	120.8	-14.8517857	135.651786
28	154.4	8.1625	146.2375
29	149.5	3.35178571	146.148214
30	157.1	3.3375	153.7625
31	130.8	-14.8517857	145.651786
32	166.5	8.1625	158.3375

**Indice trimestriel de la production industrielle  
base 100 en 1962. Source INSEE.**



## 4.1.3 La série lissée des prédictions

Etape 1. (*Ajustement d'une courbe de tendance*)

Maintenant que la série est corrigée des variations saisonnières, on peut ajuster une courbe de tendance de la forme

$$y = f_{\theta}(t).$$

Par exemple, si on décide d'ajuster une **tendance linéaire** de la forme :

$$f_{\theta}(t) = a t + b$$

alors on aura :

$$\hat{a} = \frac{\text{Cov}(\text{CVS}; (t_i)_{1 \leq i \leq n})}{\text{Var}((t_i)_{1 \leq i \leq n})} \quad \text{et} \quad \hat{b} = \overline{\text{CVS}} - \hat{a} \bar{t}$$

## Application sur l'exemple

Sur l'exemple, on va ajuster un polynôme de degré 2.

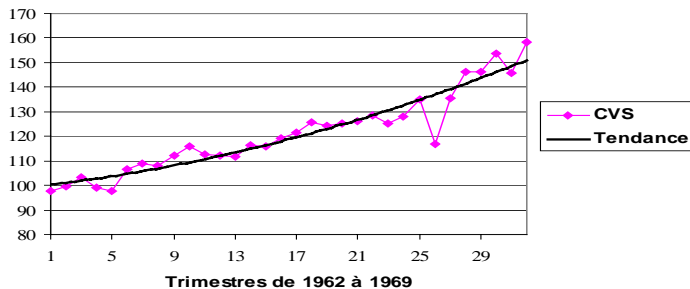
On trouve

$$\hat{f}(t) = 0.0287 t^2 + 0.6873 t + 99.621$$

et donc pour  $i = 1, \dots, 32$ ,

$$\hat{f}_i = 0.0287 t_i^2 + 0.6873 t_i + 99.621$$

**Indice trimestriel de la production industrielle  
base 100 en 1962. Source INSEE.**



## Etape 2. (*Construction de la série lissée des prédictions*)

La série lissée des prédictions  $(\hat{y}_i)_{1 \leq i \leq n}$ , utile pour la prévision, est obtenue en additionnant la tendance  $(\hat{f}_i)_{1 \leq i \leq n}$  et la composante saisonnière  $(\hat{s}_i)_{1 \leq i \leq n}$ , ie.

$$\hat{y}_i = \hat{f}_i + \hat{s}_i \quad \text{pour } i = 1, \dots, n$$

On peut alors définir l'erreur de prévision associée

$$\hat{e}_i = y_i - \hat{y}_i \quad \text{pour } i = 1, \dots, n$$

On pourra étudier la qualité de la modélisation de la chronique

- à l'aide de représentations graphiques
- à l'aide de critères numériques, comme l'erreur quadratique moyenne

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

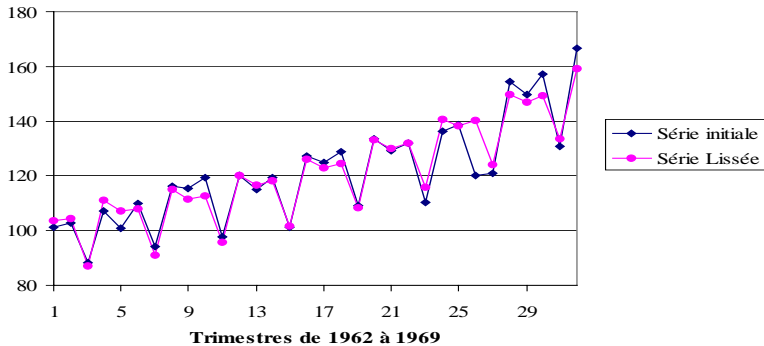
ou bien encore l'erreur absolue moyenne (MAE)

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

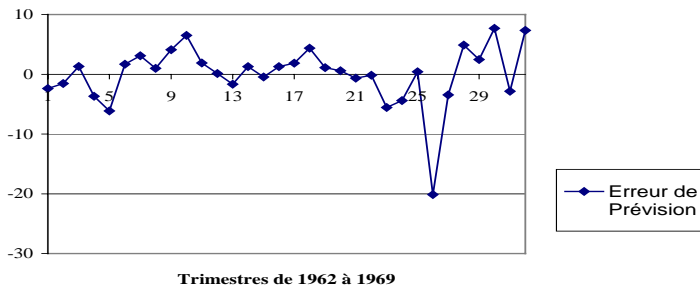
qui sont des **critères d'erreurs** permettant de comparer différentes modélisations.



**Indice trimestriel de la production industrielle base 100 en 1962.**  
**Source INSEE.**



**Indice trimestriel de la production industrielle base 100  
en 1962. Source INSEE.**



On trouve une **erreur quadratique moyenne** de :

$$\sum_{i=1}^{32} (y_i - \hat{y}_i)^2 = 24.58$$