Séries chronologiques
Corrigé des exercices

1. Bourse

Le tableau de calcul pour trouver l'équation de la tendance (droite des moindres carrés) est le suivant:

x_i	y_i	x_i^2	x_iy_i	T
1	135	1	135	135,8
2	143	4	286	140,3
3	140	9	420	144,8
4	154	16	616	149,3
5	152	25	760	153,8
Σx_i	Σy_i	Σx_i^2	Σx_iy_i	
15	724	55	2217	

$$\begin{cases} \overline{x} = 3 \quad \text{(17 mars)} \\ \overline{y} = 145 \ \text{F} \end{cases}$$

$$a = \frac{\sum x_i y_i - n\overline{x}\overline{y}}{\sum x_i^2 - n\overline{x}^2} \qquad \text{avec} \quad n = 5$$

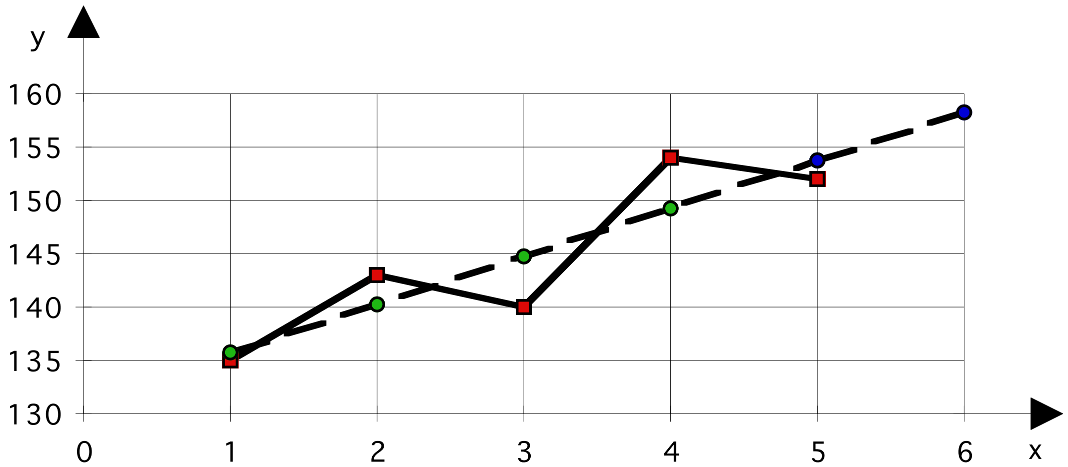
$$a = \frac{2217 - 5 \times 3 \times 145}{55 - 5 \times 3^2} = \frac{45}{10} = 4,5$$

Le cours de l'action a *tendance* à augmenter de 4,5 Francs par jour. L'équation de la droite de tendance est:

$$T = ax + b \qquad \text{avec} \qquad b = \overline{y} - a\overline{x}$$

$$T = 4,5x + 131,3$$

On trace le graphe:

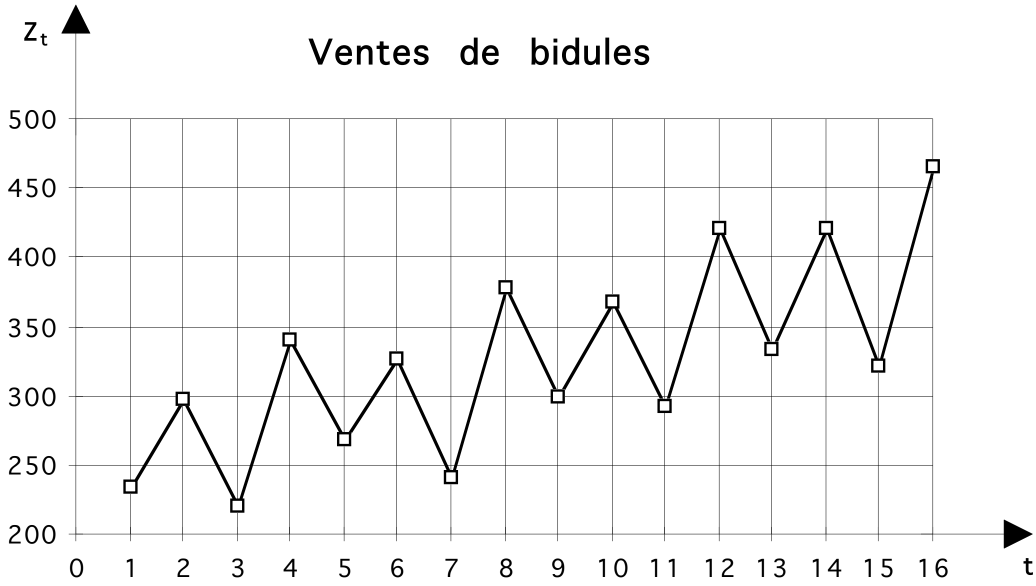


La prévision pour la date 6 s'obtient en remplaçant x par la valeur 6 dans l'équation de la tendance:

T(6) = 158 Francs

2. Bidules

On trace d'abord un graphique de la série:



La série manifeste une tendance régulière à la hausse avec des variations saisonnières assez marquées. Une méthode de décomposition semble donc bien adaptée pour son étude.

La tendance peut être déterminée par la méthode des moindres carrés. Il s'agit alors de la droite qui passe le plus près possible de l'ensemble des points. Plus précisément, si on appelle S la somme des carrés des écarts verticaux entre une droite quelconque et les points, la tendance est la droite pour laquelle la somme S est minimale.

L'effet saisonnier peut être évalué à l'aide de coefficients qui mesurent les écarts entre les ventes et la tendance.

Il reste enfin l'effet aléatoire. Dans cette approche, il s'agit d'un effet résiduel: ce sont les irrégularités qui restent inexpliquées une fois analysées la tendance et les saisons.

Les aléas peuvent être rendus plus visibles en effaçant les variations saisonnières. C'est l'intérêt de la courbe C.V.S. (Corrigée des Variations Saisonnières).

Les calculs principaux sont regroupés dans un tableau :

Années	dates t	série z_t	produits $t.z_t$	carrés t^2	Tendance T=at+b	rapports z_t / T	coef saison	CVS z_t /coef
	1	235	235	1	249	0,95	0,92	257
1	2	298	596	4	259	1,15	1,10	271
	3	221	663	9	270	0,82	0,81	273
	4	340	1360	16	280	1,21	1,17	290
	5	268	1340	25	290	0,92	0,92	293
2	6	327	1962	36	301	1,09	1,10	297
	7	242	1694	49	311	0,78	0,81	299
	8	378	3024	64	322	1,17	1,17	322
	9	300	2700	81	332	0,90	0,92	328
3	10	368	3680	100	343	1,07	1,10	334
	11	292	3212	121	353	0,83	0,81	361
	12	421	5052	144	364	1,16	1,17	359
	13	334	4342	169	374	0,89	0,92	365
4	14	421	5894	196	384	1,10	1,10	382
	15	322	4830	225	395	0,82	0,81	398
	16	465	7440	256	405	1,15	1,17	396
	136	5232	48024	1496	5232	16,00		

Point moyen
$$\begin{cases} \bar{t} = \dfrac{\sum t}{n} = \dfrac{136}{16} = 8,5 \\ \bar{z} = \dfrac{\sum z_t}{n} = \dfrac{5232}{16} = 327 \end{cases}$$

$$a = \frac{\sum tz_t - n\bar{t}\bar{z}}{\sum t^2 - n\bar{t}^2} = \frac{48024 - 16 \times 8,5 \times 327}{1496 - 16 \times 8,5^2} = \frac{3552}{340} = 10,447$$

Les ventes ont *tendance* à augmenter de 10 ou 11 milliers de bidules par trimestre. L'équation de la tendance s'écrit :

$T = at + b$ avec $b = \bar{z} - a\bar{t}$

$T = 10,447t + 238,20$

Le coefficient 238,5 (ordonnée à l'origine) correspondrait à des ventes fictives de 238 milliers de bidules au trimestre 0.

On détermine ensuite les coefficients saisonniers en calculant les moyennes des rapports entre les ventes et la tendance pour une même saison.

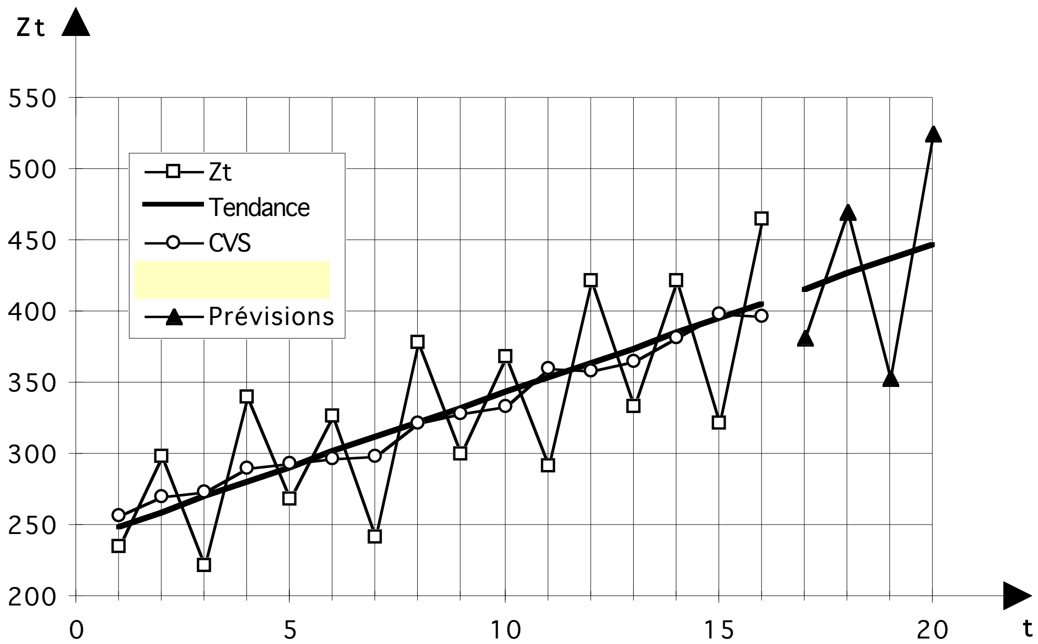
	années				coefficients saisonniers	Par rapport à la tendance
	1	2	3	4		
1° trimestre	0,95	0,92	0,90	0,89	0,92	8% au dessous
2° trimestre	1,15	1,09	1,07	1,10	1,10	10% au dessus
3° trimestre	0,82	0,78	0,83	0,82	0,81	19% au dessous
4° trimestre	1,21	1,17	1,16	1,15	1,17	17% au dessus

Les prévisions pour l'année 5 s'obtiennent en prolongeant la tendance et en appliquant chaque trimestre le coefficient saisonnier adéquat:

année	dates	Tendance	coef	Prévisions
5	17	416	0,92	381
	18	426	1,10	470
	19	437	0,81	354
	20	447	1,17	525

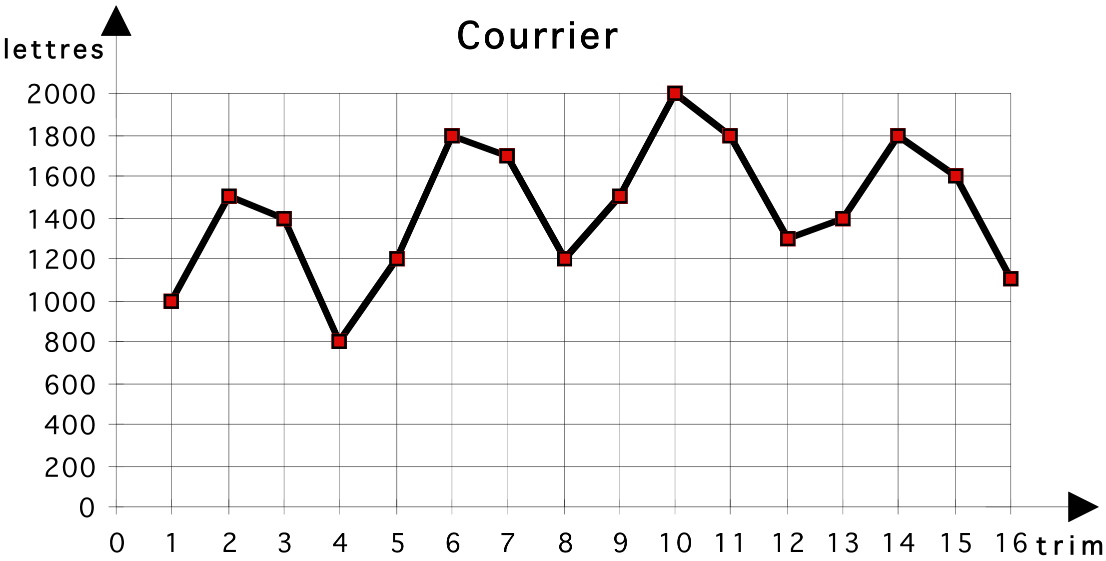
La série C.V.S. se calcule en divisant les ventes de bidules par le coefficient saisonnier du trimestre correspondant. On obtient ainsi le niveau de ventes fictif qui aurait été atteint en l'absence d'effet saisonnier. Cette série fait bien ressortir les aléas qui étaient masqués par l'effet saisonnier.

Finalement, on peut tracer les différentes courbes sur un même graphique:



3. *Un homme de lettres*

L'évolution de la masse du courrier reçu par le chanteur est la suivante:



Pour une série saisonnière, il faut lisser sur une année entière avec un nombre impair de points:

$$M_t = \frac{0,5z_{t-2} + z_{t-1} + z_t + z_{t+1} + 0,5z_{t+2}}{4} \qquad 3 \leq t \leq 14$$

Les calculs principaux sont regroupés dans un tableau:

trimestres	lettres	moy mob	rapports	coef	CVS	
1	1000	xxx	xxx	0,91	1099	
2	1500	xxx	xxx	1,23	1221	
3	1400	1200	1,17	1,13	1239	
4	800	1263	0,63	0,74	1087	*
5	1200	1338	0,90	0,91	1318	
6	1800	1425	1,26	1,23	1466	
7	1700	1513	1,12	1,13	1504	
8	1200	1575	0,76	0,74	1630	
9	1500	1613	0,93	0,91	1648	
10	2000	1638	1,22	1,23	1628	
11	1800	1638	1,10	1,13	1593	
12	1300	1600	0,81	0,74	1766	*
13	1400	1550	0,90	0,91	1538	
14	1800	1500	1,20	1,23	1466	
15	1600	xxx	xxx	1,13	1416	
16	1100	xxx	xxx	0,74	1495	

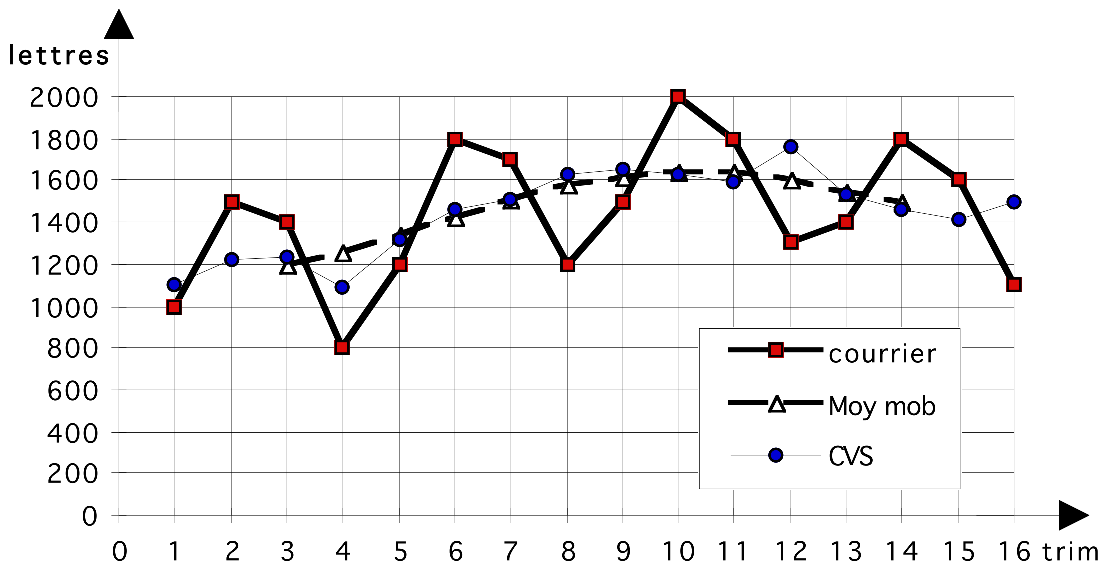
Les coefficients saisonniers sont les moyennes des rapports entre le nombre de lettres reçues et la tendance pour chaque saison:

	années				coefficients saisonnier
	1	2	3	4	
1° trimestre	xxx	0,90	0,93	0,90	0,91
2° trimestre	xxx	1,26	1,22	1,20	1,23
3° trimestre	1,17	1,12	1,10	xxx	1,13
4° trimestre	0,63	0,76	0,81	xxx	0,74

Les coefficients saisonniers s'interprètent par rapport à la tendance:

- au 1° trimestre: 9% de moins en moyenne,
- au 2° trimestre: 23% de plus,
- au 3° trimestre: 13% de plus,
- au 4° trimestre: 26% de moins.

Enfin la courbe CVS (Corrigée des Variations Saisonnières) s'obtient en divisant le volume trimestriel de courrier par le coefficient saisonnier:



Deux points retiennent l'attention en s'écartant sensiblement de la tendance: ils correspondent au courrier reçu aux trimestres 4 et 12. Un spécialiste pourrait sans doute rattacher ces perturbations à des événements marquants dans la carrière du chanteur...

4. Lissages

La série des températures du mois de février n'est pas saisonnière. Pour l'étudier, on peut faire appel au lissage par moyennes mobiles ou au lissage exponentiel. Les calculs se prêtent particulièrement bien à l'utilisation d'un tableur (EXCEL):

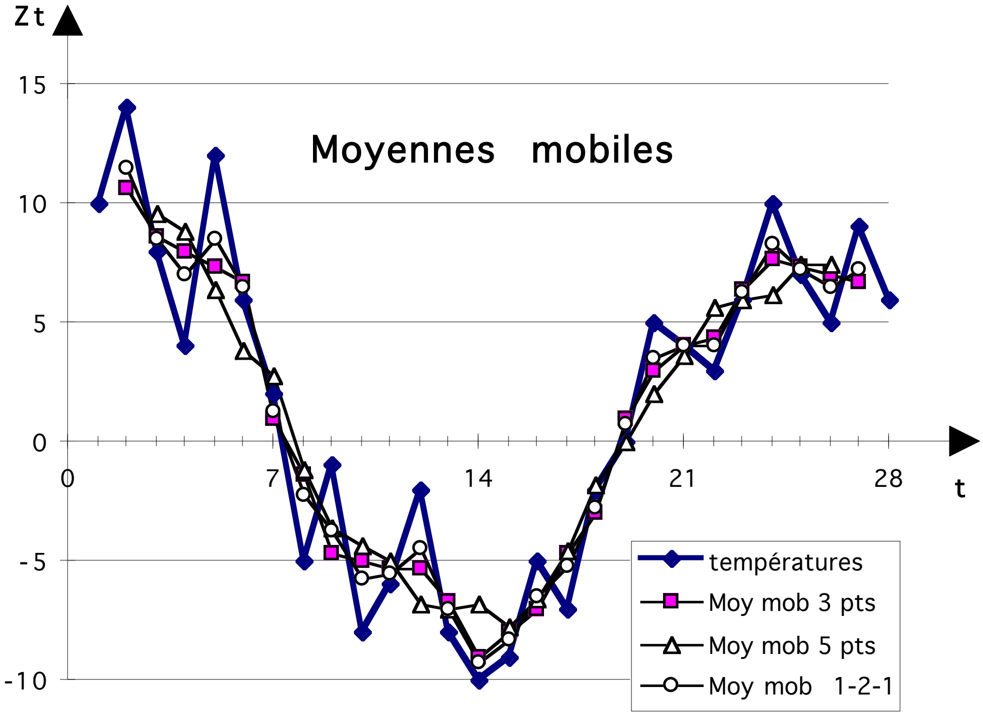
dates	Tempé-ratures	Moy mob 3 pts	Moy mob 5 pts	Moy mob 1-2-1	exp 0,3	exp 0,5	exp 0,8
1	10	xxx	xxx	xxx	10,0	10,0	10,0
2	14	10,7	xxx	11,5	11,2	12,0	13,2
3	8	8,7	9,6	8,5	10,2	10,0	9,0
4	4	8,0	8,8	7,0	8,4	7,0	5,0
5	12	7,3	6,4	8,5	9,5	9,5	10,6
6	6	6,7	3,8	6,5	8,4	7,8	6,9
7	2	1,0	2,8	1,3	6,5	4,9	3,0
8	-5	-1,3	-1,2	-2,3	3,0	-0,1	-3,4
9	-1	-4,7	-3,6	-3,8	1,8	-0,5	-1,5
10	-8	-5,0	-4,4	-5,8	-1,1	-4,3	-6,7
11	-6	-5,3	-5,0	-5,5	-2,6	-5,1	-6,1
12	-2	-5,3	-6,8	-4,5	-2,4	-3,6	-2,8
13	-8	-6,7	-7,0	-7,0	-4,1	-5,8	-7,0
14	-10	-9,0	-6,8	-9,3	-5,9	-7,9	-9,4
15	-9	-8,0	-7,8	-8,3	-6,8	-8,4	-9,1
16	-5	-7,0	-6,6	-6,5	-6,3	-6,7	-5,8
17	-7	-4,7	-4,6	-5,3	-6,5	-6,9	-6,8
18	-2	-3,0	-1,8	-2,8	-5,1	-4,4	-3,0
19	0	1,0	0,0	0,8	-3,6	-2,2	-0,6
20	5	3,0	2,0	3,5	-1,0	1,4	3,9
21	4	4,0	3,6	4,0	0,5	2,7	4,0
22	3	4,3	5,6	4,0	1,2	2,8	3,2
23	6	6,3	6,0	6,3	2,7	4,4	5,4
24	10	7,7	6,2	8,3	4,9	7,2	9,1
25	7	7,3	7,4	7,3	5,5	7,1	7,4
26	5	7,0	7,4	6,5	5,4	6,1	5,5
27	9	6,7	xxx	7,3	6,4	7,5	8,3
28	6	xxx	xxx	xxx	6,3	6,8	6,5

Pour les moyennes mobiles, les formules utilisées sont les suivantes:

$$L_t = \frac{z_{t-1} + z_t + z_{t+1}}{3} \qquad 2 \leq t \leq 27$$

$$M_t = \frac{z_{t-2} + z_{t-1} + z_t + z_{t+1} + z_{t+2}}{5} \qquad 3 \leq t \leq 26$$

$$N_t = \frac{z_{t-1} + 2z_t + z_{t+1}}{4} \qquad 2 \leq t \leq 27$$



On peut classer les courbes dans l'ordre suivant, de la plus agitée à la plus lisse:

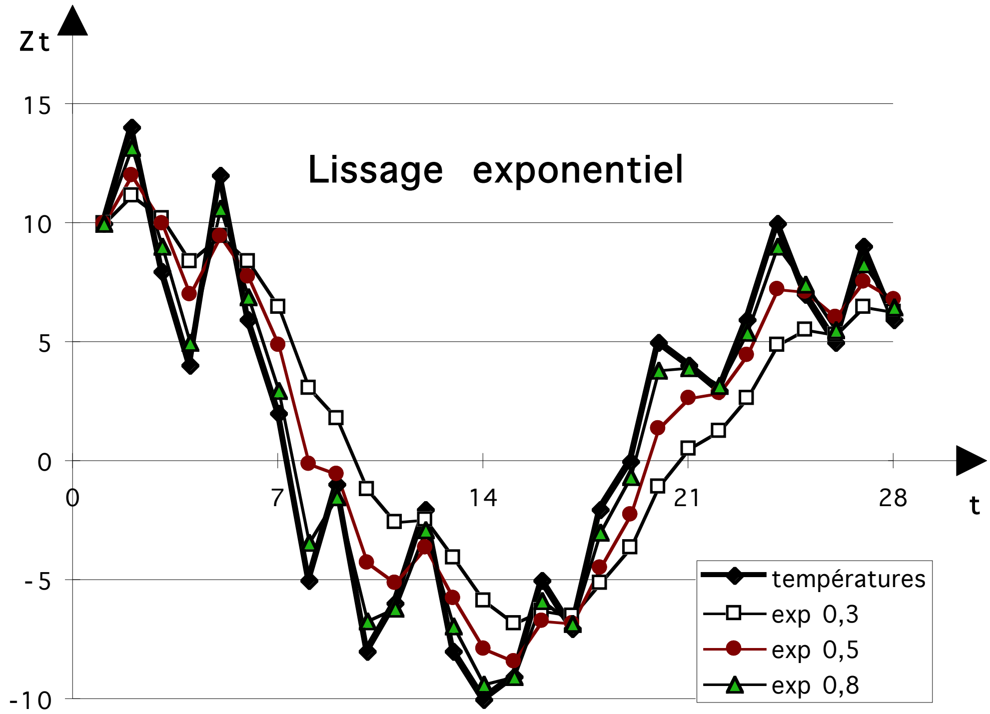
zt : série brute,
Nt : moyennes mobiles sur 3 points avec les poids 1-2-1,
Lt : moyennes mobiles sur 3 points,
Mt : moyennes mobiles sur 5 points.

On peut aussi essayer le lissage exponentiel avec différentes valeurs de la constante de lissage α.

$$E_t = \alpha z_t + (1-\alpha)E_{t-1} \qquad \begin{cases} 0 \le \alpha \le 1 \\ 2 \le t \le 28 \end{cases}$$

Pour démarrer la récurrence, il faut choisir une première valeur, par exemple $E_1=z_1$. De toutes façons, l'influence de ce choix s'atténue assez vite en progressant dans la série.

On constate facilement que les courbes sont d'autant plus lisses que la constante de lissage est plus faible.



Avec une constante forte (proche de 1, par exemple 0,8), la courbe est proche de la série brute. La série brute correspond à un lissage limite avec α=1.

On peut vérifier que la constante de lissage définit une sorte de "mémoire" du processus:

- avec a=0,8 la mémoire est faible. La série lissée suit très rapidement les changements de direction.

- avec a=0,3 la mémoire est forte. La série lissée suit avec beaucoup de retard les changements de tendance comme si elle gardait plus longtemps en mémoire les valeurs anciennes.

La valeur lissée E_t sert également de prévision à l'horizon 1 (prévision faite à la date t pour la date t+1):

$$\hat{z}_t(1) = E_t$$

Ce type de modèle (lissage exponentiel simple) ne convient qu'à des séries stationnaires, c'est-à-dire avec une tendance à long terme horizontale. Dans le cas d'une série non stationnaire, la série lissée serait constamment en retard sur la série étudiée.

Il existe des modèles plus sophistiqués (lissages multiples) qui permettent de rattraper ce retard...