D - Autres méthodes de Prévisions courantes: Méthodes de Lissage.

Les méthodes de lissage sont des techniques qui permettent de définir des prévisions sur le court terme d'une grandeur chronologique à travers un ajustement de la série par une forme fonctionnelle implicite ou explicite du temps. La prévision s'obtient après détermination d'une fonction $f(.)$ souvent sous la forme d'un polynôme en t , $Y_t = f(t) + \epsilon_t$,et après estimation des paramètres apparaissant dans la fonction $f(t)$, selon une méthode de moindres carrés accordant un poids plus important aux observations les plus récentes. Comme nous le verrons plus loin, cela revient à admettre que les termes d'erreur ϵ_t sont indépendants et vérifient les hypothèses suivantes,

$$E(\epsilon_t) = 0$$
$$et \quad Var(\epsilon_t) = \frac{\sigma^2}{W^{T-1}} \quad pour \quad t = 1, ..., T$$

Les conditions sur la variance signifient que les erreurs ϵ_t , et donc les Y_t eux mêmes, sont de plus en plus précis pour $t = 1, ..., T$ du fait que $W \in]0, 1[$. La prévision se ramène alors à extrapoler la forme fonctionnelle aux périodes futures $T + 1, T + 2, ...$

De ce fait, les propriétés des prévisions sont intimement reliées au degré de complexité des méthodes d'ajustement de la série Y_t pour $t = 1, ..., T$. Par prévision de court terme, on entend une prévision obtenue par extrapolation de la fonction $f(.)$ aux périodes futures $T+1, T+2, ...$ Cette extrapolation n'est valable que sur un horizon temporel ne permettant pas d'ajustements structurels majeurs du phénomène étudié par le biais de ses déterminants. Dans le cas où la période est suffisante pour générer des ajustements et des modifications majeures, on parle alors de moyenne ou de longue période.

Principes des Méthodes de Lissage

* "Lisser" les aléas afin d'observer la tendance;
* Ajuster une fonction du temps à des données et extrapoler;
* Donner plus de poids aux observations les plus récentes.
* Aussi, leur but est de retirer ou de réduire les fluctuations (cycliques ou non) de court terme des séries.

Les différentes méthodes de Lissage:

$*$ Méthodes de décomposition: Lissage par Moyenne Mobile _c'est une Méthode non paramétrique d'estimation du trend/des saisonnalités par moyenne mobile;

$*$ Lissage Exponentiel: Méthode paramétrique de prévision - par extrapolation, elle consiste à prolonger l'évolution passée.

$*$ Méthode de Holt-Winters.

$\Rrightarrow$ 1)–**Méthodes de décompositions (saisonnières) - Lissage par Moyenne Mobile:** Il n'y a pas de théorie statistique à proprement parler derrière les méthodes de décompositions, elles relèvent plutôt de l'intuition et de la pratique. Elle se base sur l'identification des diverses composantes de la série pour en tirer profit dans la phase prévisionnelle. En effet, dans un cadre non paramétrique, elle permet de décomposer la série initiale en fonction d'une tendance et d'une composante saisonnière pour la prévision.

- ## D'où, Ajustements saisonniers - i.e. Composantes ou Mouvements d'une chronique):

Il existe diverses méthodes de correction des variations saisonnières. Elles fonctionnent pour la plupart sur une décomposition / juxtaposition entre plusieurs composantes de nature différente qui caractérisent la dynamique de la série, tendance, cycle conjoncturel, variations saisonnières et variations accidentelles ou erreurs, de la forme :

- *Décomposition (modèle) additive:* modèle plus approprié pour les séries ayant une composante saisonnière constante dans le temps ($S_{\ t+s} = S_t$), i.e. que les paramètres décrivant la série ne changent pas au cours du temps. Il résulte d'une addition de la tendance et de la partie saisonnière où Lles composantes sont indépendantes les unes des autres.

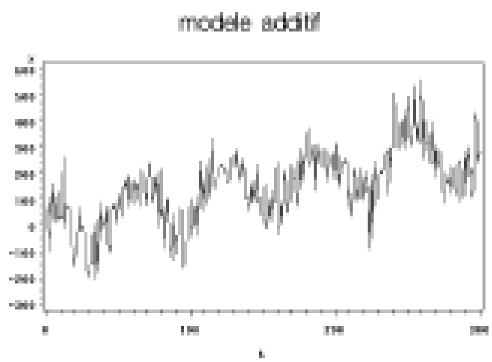
En général, on a

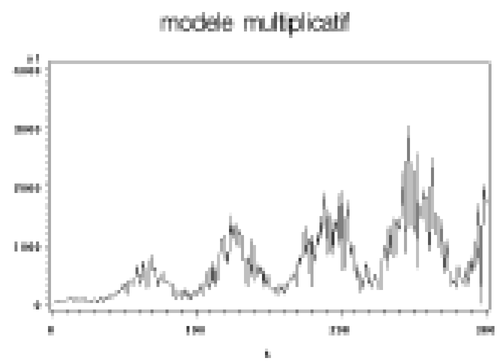
$$Yt = Tendance(t) + Cycle(t) + Saison(t) + Résidu(t)$$
$$i.e.,\quad Yt\ = T_t + C_t + S_t + E_t$$

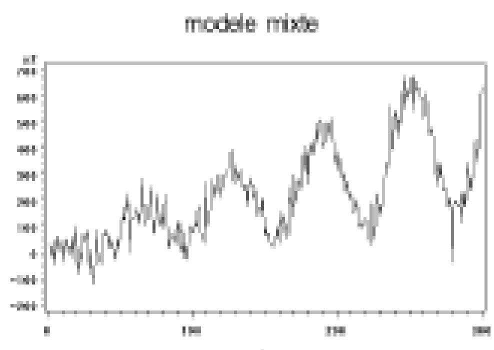
- *Décomposition (modèle) multiplicative:* convient aux séries comportant une partie saisonnière qui décroît ou croit dans le temps.Et, les composantes dépendent les unes des autres.

$$Yt = Tendance(t) * Cycle(t) * Saison(t) * Résidu(t)$$
$$i.e.,\quad Yt\ = T_t * C_t * S_t * E_t$$

- Aussi, on trouve une *Décomposition (modèle) Mixte :* $Yt\ = T_t{*}C_t{*}S_t{+}E_t$


modele additif


modele multiplicatif


modele mixte

La $Tendance$, T_t, donne l'allure générale de la série, c'est à dire les variations, l'évolution sur un horizon de long terme du phénomène étudié (i.e. la série) Yt. L'évolution régulière de la série Yt se fait alors autour de cette direction tendancielle appelée également tendance séculaire. Cette dernière peut être croissante ou décroissante ou alors croissante pour certaines périodes successives et décroissantes pour d'autres.La tendance peut être sous la forme,

- affine: $T_t = at + b$
- polynômiale: $T_t = a_p t^p + a_{p-1} t^{p-1} + ... + a_1 t + a_0 \ \ avec\ p \in N^*$
- logistique: $T_t = a\left(1 + e^{-bt}\right)^{-1},\ b \succeq 0$
- de Gompertz / Exponentielle: $T_t = a - be^{ct}$

Le $cycle\ conjoncturel, C_t$, (ou dit aussi cycles autour de la tendance) est un élément de l'évolution de la grandeur qui se réalise d'une manière à la fois subite, importante et souvent non répétitive. Il regroupe les variations autours de la tendance avec des alternances de phases d'expansion, de contraction et de récession. Il peut traduire une évolution résultant d'une action involontaire telle une crise économique, un incident majeur ou une catastrophe naturelle. Aussi, il peut également découler d'une évolution volontaire telle une action promotionnelle lancée par un grand magasin ou une intervention délibérée des autorités publiques sur un marché en vue d'atteindre un objectif commercial, financier, monétaire, etc.

Les $variations\ saisonnières$ ou $mouvements\ saisonniers, S_t$, traduisent des effets périodiques de période connue p jouant sur l'évolution de la série Yt dans un même sens pour les périodes correspondantes aux mêmes saisons (i.e.ils se reproduisent de façon plus ou moins identique d'une période sur l'autre). Les effets saisonniers correspondent aux fluctuations périodiques, ils peuvent être favorables ou défavorables sur l'évolution de la grandeur étudiée selon la nature de la saison. Ces effets se manifestent souvent durant les jours de fête, les vacances, la période de récoltes, la rentrée scolaire, etc.

La chronique correspondante, également déterministe, notée $S_t\ où\ t = 1, ..., T$, est généralement supposée rigoureusement périodique, $S_{t+p} = S_t$ et les valeurs $S_j = S_{ij}\ où\ j = 1, ..., p$ d'une période sont appelées les coefficients / effets saisonniers. Le bilan de l'effet saisonnier sur une période doit être nul car il est pris en compte dans la tendance. La composante saisonnière permet simplement de distinguer à l'intérieur d'une même période une répartition stable dans le temps d'effets positifs ou négatifs qui se compensent sur l'ensemble de la période.

Enfin, les $variations\ accidentelles,\ aléatoire\ ou\ résiduelles,\ E_t$, correspondent à la partie non structurée du phénomène. Elles résultent de multiples causes. Elles sont aléatoires et de courte durée. Elles décrivent une partie de l'évolution de la grandeur Yt souvent imprévue, exceptionnelle et dont l'origine n'est souvent pas parfaitement identifiée. L'absence de l'élément aléatoire, E_t, rendrait l'évolution de la chronique Yt de nature déterministe, ce qui est irréel. Elle est modélisée par une suite de variables aléatoires ϵ_t *où* $t = 1, ..., T$, centrées, homoscédastiques et indépendants; on parle de bruit blanc où,

$$E(\epsilon_t) = 0$$
$$Var(\epsilon_t) = \sigma^2$$
$$Cov(\epsilon_t, \epsilon_s) = 0 \;\; pour \;\; t \neq s.$$

- ## Définitions Moyenne Mobile :

* On appelle moyenne mobile centrée de longueur impaire $l_i = 2k + 1$ à l'instant t la valeur moyenne (arithmétique) mm_t des observations Y_{t-k}, Y_{t-k+1}, ..., Y_t, Y_{t+1}, .., Y_{t+k} définie par

$$mm_t = \frac{Y_{t-k} ,+ Y_{t-k+1} + ... + Y_{t-1} + Y_t + Y_{t+1} + .. + Y_{t+k}}{l_i}$$

* On appelle moyenne mobile centrée de longueur paire $l_p = 2k$ à l'instant t la valeur moyenne mm_t des observations Y_{t-k}, Y_{t-k+1}, ..., Y_t, Y_{t+1}, .., Y_{t+k}, la première et la dernière étant pondérée par 0.5:

$$mm_t = \frac{0.5Y_{t-k} + Y_{t-k+1} + ... + Y_{t-1} + Y_t + Y_{t+1} + .. + Y_{t+k-1} + 0.5\, Y_{t+k}}{l_p}$$

Dans la première formule, le nombre de termes de la somme est égal à $2k+1$: il s'agit bien d'une moyenne. Dans la seconde, la somme des coefficients est égale à $2k$, puisque le premier et le dernier sont égaux à 0.5 : il s'agit d'une moyenne pondérée. Dans les deux cas, le nombre d'observations prises en compte avant l'instant t est égal au nombre d'observations prises en compte après l'instant t: on qualifie, ainsi, les moyennes de centrées.

La première valeur d'une moyenne mobile de longueur $4(= 2x2)$ *ou* $5(= 2x2 + 1)$ que l'on peut calculer, **est à l'instant** $t = 3$, puisque la première observation connue est Y_1, où:

$$pour \;\; l = 4, \; on \; a \;\; mm_3 = \frac{0.5Y_1 + Y_2 + Y_3 + Y_4 + 0.5\, Y_5}{4}$$
$$pour \;\; l = 5, \; on \; a \;\; mm_3 = \frac{Y_1 + Y_2 + Y_3 + Y_4 + Y_5}{5}$$

De façon générale, ne peut pas calculer de moyennes mobiles en $t = 1, t = 2, \ldots, t = k$ puisque les formules ne peuvent être appliquées que si l'on connaît Y_{t-k}. De même, si T est le nombre total d'observations, on ne peut pas calculer mm_T ,..., mm_{T-k+1} puisqu'il faut connaître Y_{t+k}.

- L'avantage de la technique des moyennes mobiles est d'atténuer, d'éliminer ou d'amortir les mouvements cycliques, saisonniers et accidentels tout en conservant les tendances linéaires : la série est dite « lissée », et est d'autant plus lissée que la longueur de la moyenne mobile est élevée ((comme on peut le constater sur la figure 3.8 sur laquelle nous avons représenté les moyennes mobiles de longueur 14.))

Son principe est de construire une nouvelle série obtenue en calculant des moyennes arithmétiques successives de longueur p fixe à partir des données originales. Chacune de ces moyennes obtenues correspondra au "milieu" de la période pour laquelle la moyenne arithmétique vient d'être calculée.

t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9
4	6	5	3	7	5	4	3	6

- Exemple: On dispose du tableau ci-dessous contenant des mesures d'un phénomène relevées à 9 instants différents.

t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9
4	6	5	3	7	5	4	3	6
	5.00	4.67	5.00	5.00	5.33	4.00	4.33	

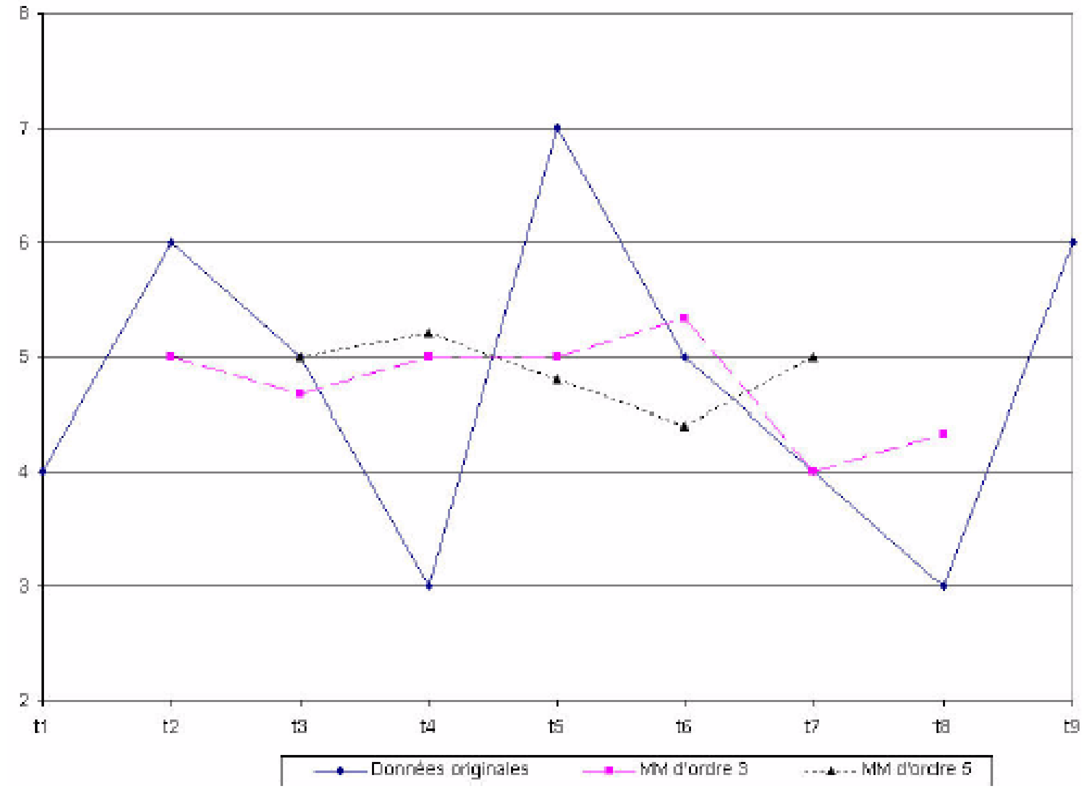
– Le calcul des moyennes mobiles d'ordre 3 donne les valeurs suivantes:

La moyenne mobile d'ordre 3 à l'instant $t = 6$ est $mm_6 = \frac{7+5+4}{3} \approx 5.33$. On constate que les deux valeurs extrêmes pour $t = 1$ et $t = 9$ ont disparu, i.e. une de chaque côté.

t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9
4	6	5	3	7	5	4	3	6
		5	5.2	4.8	4.4	5		

– Le calcul des moyennes mobiles d'ordre 5, donne le résultat suivant:

Exemple, la valeur pour $t = 4$ est $mm_4 = \frac{6+5+3+7+5}{5} = 5.2$. On remarque que les deux valeurs extrêmes de la série sont perdues. On constate ainsi que si la longueur de la moyenne mobile p ou $l_p = 2k + 1$, impaire à l'instant t, alors à chaque extrémité, on aura k valeurs qui seront perdues.



Par représentation graphiquement des résultats, on observe la tendance au "lissage" de la représentation orginale avec l'utilisation de la technique des moyennes mobiles:

Modélisation et Désaisonnalisation:

• Estimation / analyse des mouvements saisonniers

Pour réaliser cette estimation/analyse des mouvements saisonniers, on doit savoir d'abord si la série brute étudiée décrit un modèle additif ou un modèle multiplicatif.

** Donc, comment déterminer la nature du modèle (*additive / multiplicative)?*

- Si pour une observation donnée, la variation saisonnière S_t s'ajoute simplement aux autres composantes; on a à faire au modèle additif;

- Si pour une observation donnée, la variation saisonnière S_t est proportionnelle aux autres composantes, c'est le modèle multiplicatif.

La distinction entre ces deux décomposition peut se faire soit par une méthode graphique/du profil ou soit par une méthode analytique.

– Par méthode graphique, on superpose les saisons représentées par des droites de profil sur un même graphique. Si ces droites sont parallèles, le modèle est additif, autrement le modèle est multiplicatif.

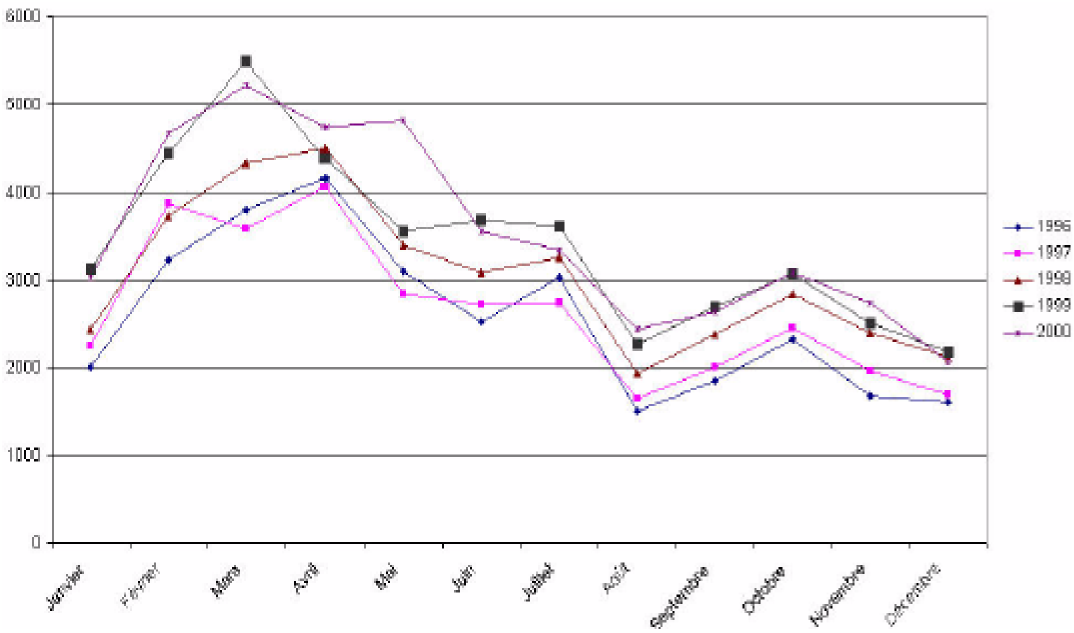
Année	Janvier	Février	Mars	Avril	Mai	Juin	Juillet	Août	Septembre	Octobre	Novembre	Décembre
1996	2006	3224	3789	4153	3100	2527	3015	1504	1847	2314	1673	1602
1997	2247	3862	3586	4047	2838	2727	2730	1648	2007	2450	1966	1695
1998	2433	3723	4325	4493	3399	3083	3247	1928	2377	2831	2388	2126
1999	3127	4437	5478	4384	3552	3678	3611	2260	2699	3071	2510	2182
2000	3016	4671	5218	4746	4814	3545	3341	2439	2637	3085	2737	2055

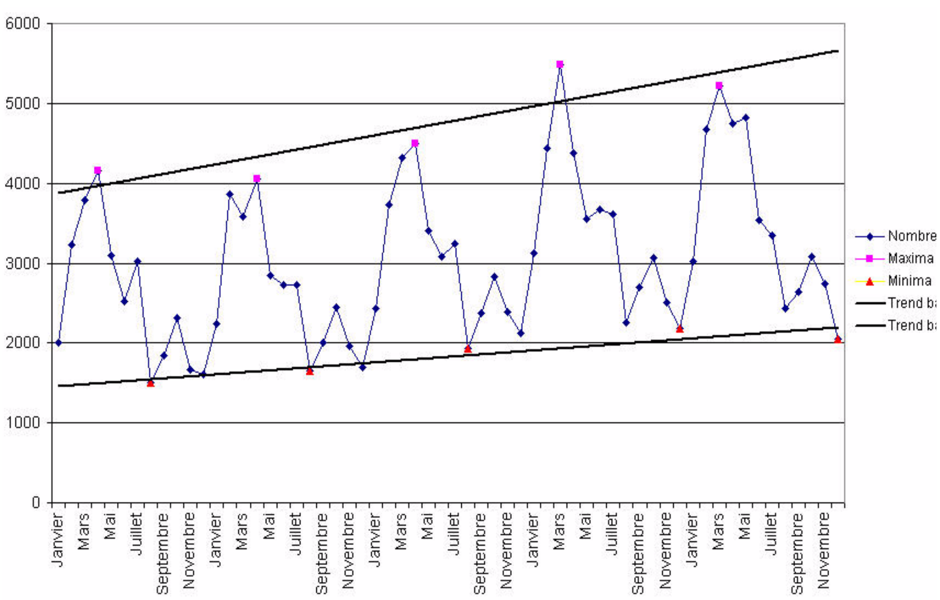
Exemple, considérons la chronique suivante:

Graphiquement, on observe que les droites de profil ne sont pas parallèles pour toutes les saisons; ainsi, le modèle de la chronique est multiplicatif.

Aussi, on peut confirmer ce résultat par la méthode graphique de la bande: A partir d'un graphique représentant la série chronologique, on trace une droite passant respectivement par les minima et par les maxima de chaque saison. Si ces deux droites sont parallèles, nous sommes en présence d'un modèle additif. Dans le cas contraire, c'est un modèle multiplicatif. Au niveau de l'exemple, on a

Par la méthode de la bande, on constate que les deux droites ne sont pas parallèles, ce qui signifie qu'on est en présence d'un modèle multiplicatif.


6000
5000
4000
3000
2000
1000
0
Janvier
Février
Mars
Avril
Mai
Juin
Juillet
Août
Septembre
Octobre
Novembre
Décembre
1996
1997
1998
1999
2000


6000
5000
4000
3000
2000
1000
0
Janvier
Mars
Mai
Juillet
Septembre
Novembre
Nombre
Maxima
Minima
Trend ba
Trend ba

	Moyenne	Écart-type
1996	2562.8	850.7
1997	2650.3	782.3
1998	3029.4	803.6
1999	3415.8	946.6
2000	3525.3	1023.4

–**Par la méthode analytique (test de Buys-Ballot):** On calcule les moyennes et écarts-types pour chacune des périodes considérées; puis, par les moindres carrées, on regresse l'écart-type sur la moyenne et une dérive, $\sigma = a\ \overline{x}\ + b$. Si le coefficient a est nul, on a un modèle additif, sinon, le modèle est multiplicatif.

Au niveau de l'exemple, on a:

Par la régression des moindres carrés, on a $\widehat{a} = 0.195$ et $\widehat{b} = 289.037$, ce qui confirme encore une fois que nous sommes bien en présence d'un modèle multiplicatif.

- Correction des variations saisonnières: Modélisation et désaisonnalisation.

Dans beaucoup de situations, il est préférable de travailler sur des données qui ne sont pas affectées par un mouvement saisonnier. Ainsi, on essaye toujours de transformer la série chronologique initiale en données désaisonnalisées ou corrigées des variations saisonnières.

— Pour un modèle additif, où la série chronologique se décompose en une tendance, c_t, des variations saisonnières s_t de période p (égales à $s_1, s_2, ..., s_p$) et d'une composante accidentelle, e_t; et, la variation saisonnière s'ajoute à la tendance:

$$y_{ij} = c_{ij} + s_j + e_{ij} \text{ pour } i = 1, ..., n \text{ et } j = 1, ..., p.$$
$$\textit{où } s_j \textit{ définit la variation saisonnière à l'instant } j \textit{ de chaque période } i,$$
$$\textit{i.e. du trimestre } j \textit{ pour des données trimestrielles } (p = 4)$$
$$\textit{et du mois } j \textit{ pour des données mensuelles } (p = 12);$$
$$\textit{Et, on note } mm_{ij} \textit{ les moyennes mobiles.}$$

Notons que:

** Pour un même trimestre, la différence entre l'observation et la tendance est à peu près constant et égale à S_j (on suppose que la composante accidentelle est relativement faible):

$y_{ij} - c_{ij} = s_j + e_{ij}$ *pour* $i = 1, ..., n$ *et* $j = 1, ..., p$.

** Aussi, la différence entre une observation y_{ij} et la moyenne mobile mm_{ij} correspondante est à peu près constante pour j fixé (une saison) :

$y_{ij} - mm_{ij} \approx s_j$ *pour* $i = 1, ..., n$ *et* $j = 1, ..., p \Rrightarrow$ c'est une approximation du coefficient S_j; propriété bien recherchée sur la représentation graphique de la série Yt pour déterminer si cette série suit un modèle additif ou non.

** En général, la différence $Yt - mm_t$ définit les écarts saisonniers / différences saisonnières entre la série initiale Yt et la série moyenne mobile, $Yt - mm_t$. C'est la moyenne des résidus obtenue pour une saison qui permet d'estimer l'effet de cette saison $estim\,(S_t)$. En effet, La moyenne des S_j, pour chaque colonne j, donne une première estimation $S_j^{'}$ définie par: $s_j^{'} = \frac{1}{n} \sum_{i=1}^{n} (y_{ij} - mm_{ij})$

** Les estimations définitives S_j sont obtenues en centrant les termes $s_j^{'}$:

$$s_j = s_j^{'} - m_{s_j^{'}} \quad \textit{pour} \quad j = 1, ..., p$$
$$\textit{où} \quad m_{s_j^{'}} = \frac{1}{p}\left(s_1^{'} + ... + s_p^{'}\right), \textit{ moyenne des } s_j^{'}$$

Remarque: la période p des variations saisonnières est la longueur exprimée en unités de temps séparant deux variations saisonnières dues à un même phénomène.

D'où, on définit les estimations des coefficients saisonniers d'un modèle additif par:

· Le calcul des différences entre les observations et les moyennes mobiles ;

· Le calcul de la moyenne ou la médiane $s_j^{'}$ des différences de chaque colonne du tableau ;

· Le calcul de la moyenne $m_{s_j^{'}}$ de ces valeurs $s_j^{'}$;

· Les estimations s_j sont ainsi obtenues en centrant les valeurs $s_j^{'}$ où $s_j = s_j^{'} - m_{s_j^{'}} \quad pour \quad j = 1, ..., p.$

	1^{er} trimestre	2^{e} trimestre	3^{e} trimestre	4^{e} trimestre
Année 1	89.658	97.593	108.906	114.157
Année 2	96.205	99.399	112.763	119.185
Année3	99.602	105.192	116.556	121.911
Année 4	103.272	109.644	121.208	126.508
Année 5	105.637	113.428	125.641	131.147
Année 6	111.118	117.215	129.776	133.000

Tableau 3.8 : série chronologique 1
(période p = 4)

** Exemple: Soit la série chronologique suivante observée trimestriellement pendant 6 ans:

L'observation de chaque trimestre est soumis à un effet particulier qui revient tous les ans; il y a donc 4 variations saisonnières correspondant chacune à un trimestre (i.e. la périodicité est de 4 trimestre). Donc, on calcule la moyenne mobile d'ordre 4 suivante:

On remarque que les différences apparaissant dans une même colonne sont proches les uns des autres et caractérisent ainsi, l'aspect du modèle additif.

- Aussi, on en déduit les moyennes (des résidus) suivantes:
$s_1^{'} = -10.2897$; $s_2^{'} = -5.4735$; $s_3^{'} = 5.5979$ $s_4^{'} = 10.1371$
- La moyennes des $s_j^{'}$ calculée est $m_{s_j^{'}} = -0.007039938$.
- En posant $s_j = s_j^{'} - m_{s_j^{'}}$ *pour* $j = 1, ..., p$, on obtient les estimations définitives S_j :
$s_1 = -10.2827$; $s_2 = -5.4664$; $s_3 = 5.6049$ $s_4 = 10.1442$

Remarque pour l'exemple: On observe que contrairement aux moyennes mobiles de longueur 4, les moyennes mmobiles de longueur 5 n'éliminent pas les variations saisonnières - voir graphique.

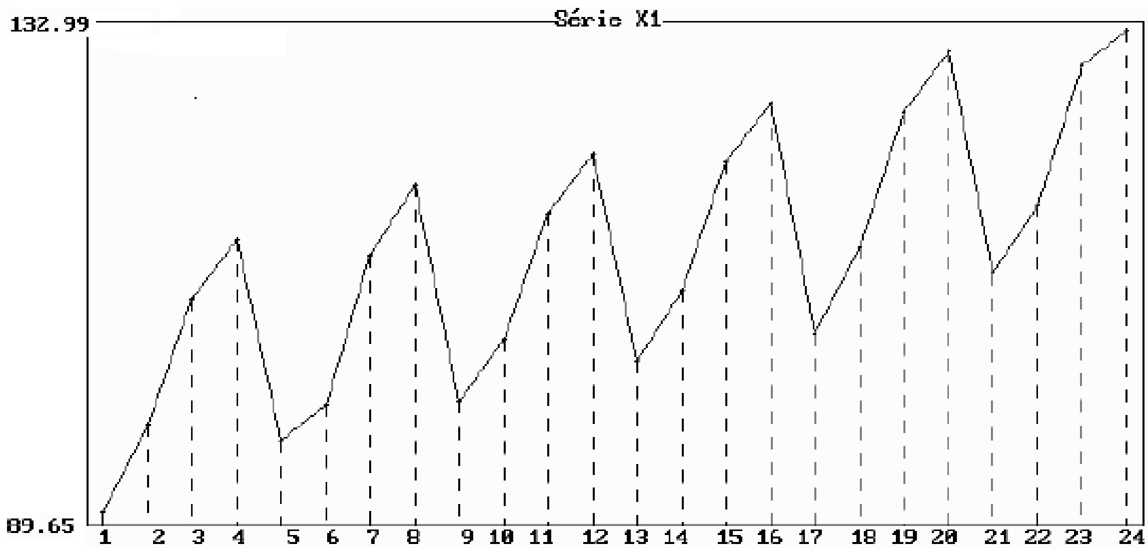


Figure 4.8 : représentation graphique de la série 1 (données observées trimestriellement pendant 6 ans)

— Pour un modèle Multiplicatif, on a:

$$Yt \ = c_t \, (1 + S_t) + \epsilon_t \ , \quad \forall \ t = 1, .., T$$

$$ou \ bien \ en \ fonction \ des \ données, \ on \ a \ Yij \ = c_{ij} \, (1 + S_{ij}) + \epsilon_{ij} \quad \forall i = 1, .., n \ \ et \ \forall j = 1, .., p$$

** Les différences entre l'observation et la tendance sont définies par

$$y_{ij} - c_{ij} = c_{ij} \ \ s_j + e_{ij} \ \ pour \ i = 1, ..., n \ \ et \ \ j = 1, ..., p.$$

Ces différences $y_{ij} - c_{ij}$, pour le $jième$ trimestre observé en l'année i, sont proportionnelles à la tendance c_{ij} : ce qui signifie que lorsque la tendance est croissante, les différences augmentent, lorsqu'elle est décroissante, ces dernières diminuent. Les différences permettent ainsi de déterminer si la série chronologique étudiée suit un modèle multiplicatif.

** On quantifie les variations saisonnières par les rapports:

$$\frac{y_{ij}}{c_{ij}} = 1 + s_j + \frac{e_{ij}}{c_{ij}} \quad pour \ \ i = 1, ..., n \quad et \quad j = 1, ..., p.$$

En admettant que les variations accidentelles e_{ij} sont faibles par rapport à la tendance c_{ij} et en utilisant l'approximation de la tendance par les moyennes mobiles, on constate donc

que les rapports $\frac{y_{ij}}{mm_{ij}}$ $j = 1, ..., p$ sont à peu près constants et donnent une approximation de $1 + s_j = S_j$, $i.e.$:

	1^{er} trimestre	2^{e} trimestre	3^{e} trimestre	4^{e} trimestre
Année 1			*103.39678*	*104.44080*
Année 2	*105.14860*	*106.25917*	*107.31233*	*108.46116*
Année3	*109.65950*	*110.47448*	*111.27404*	*112.28937*
Année 4	*113.42748*	*114.58360*	*115.45379*	*116.22236*
Année 5	*117.24943*	*118.38337*	*119.64839*	*120.80691*
Année 6	*121.79719*	*122.54573*		

Tableau 4.8 : moyennes mobiles de longueur 4 de la série 1

	1^{er} trimestre	2^{e} trimestre	3^{e} trimestre	4^{e} trimestre
Année 1			*5.50932*	*9.71580*
Année 2	*-8.94393*	*-6.86050*	*5.45047*	*10.72334*
Année3	*-10.05746*	*-5.28258*	*5.28226*	*9.62153*
Année 4	*-10.15538*	*-4.93910*	*5.75471*	*10.28534*
Année 5	*-11.61263*	*-4.95497*	*5.99271*	*10.33979*
Année 6	*-10.67929*	*-5.33023*		

Tableau 5.8 : différences entre les observations et les moyennes mobiles de la série 1

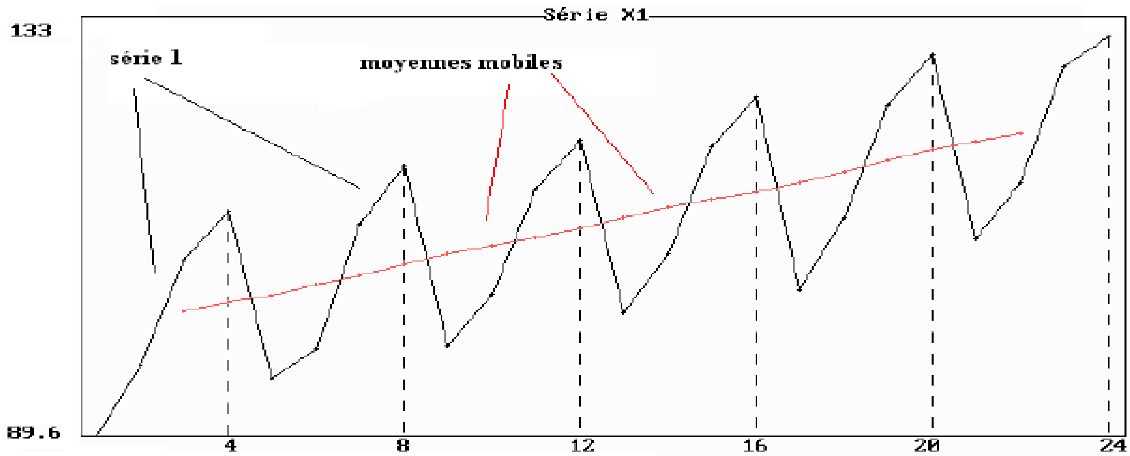


Figure 5.8 : représentation simultanée de la série 1 et des moyennes mobiles de longueur 4

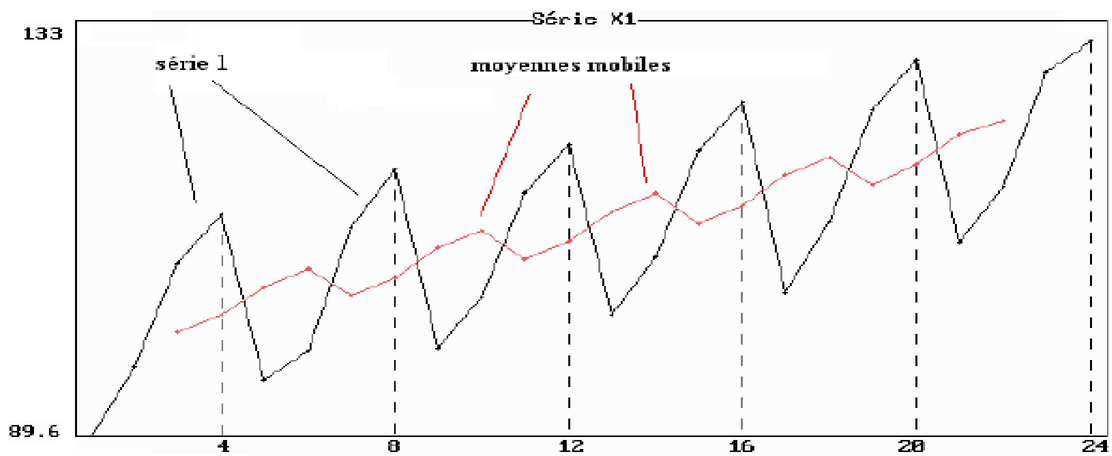


Figure 6.8 : représentation simultanée de la série 1 et des moyennes mobiles de longueur 4

$$pour\ tout \quad j=1,...,p,\ on \quad a \quad \frac{y_{ij}}{mm_{ij}}=1+s_j=S_j$$

** La moyenne des S_j, pour chaque colonne j, donne une première estimation S_j' définie par: $s_j'=\frac{1}{n}\sum_{i=1}^{n} \frac{y_{ij}}{mm_{ij}}$.

** Par analogie avec les coefficients saisonniers s_j du modèle additif, dont la moyenne est égale à 0 (i.e.$m_{s_j'}=1$), on cherche des estimations définitives S_j de moyenne 1 (i.e.$m_{s_j'}=1$) où:

$$on\ calcule\ la\ moyenne\ \ des\ s_j':$$
$$m_{s'} \ = \frac{1}{p}\left(s_1'+...+s_p'\right) \ \ pour\ \ j=1,...,p$$
$$En\ posant\ que\ \ S_j=\frac{s_j'}{m_{s'}} \ \ pour\ \ j=1,...,p\ ,$$
$$les\ coefficients\ saisonniers\ estimés\ S_j\ sont\ ainsi\ de\ somme\ p:$$
$$S_1+...+\ S_p=\frac{1}{m_{s'}}\left(s_1'+...+s_p'\right)=p$$
$$ce\ qui\ équivaut\ à\ une\ moyenne\ des\ s_j\ \ égale\ à\ 0$$
$$puisqu'on\ a\ les\ coefficients\ saisonniers\ du\ \ \text{mod}\,èle\ multiplicatif\ \ :\ S_j=1+s_j$$

D'où, on définit les estimations des coefficients saisonniers d'un modèle multiplicatif par:

- Le calcul des rapports des observations aux moyennes mobiles ;
- Le calcul de la moyenne ou la médiane des rapports s_j' de chaque colonne du tableau ;
- Le calcul de la moyenne $m_{s_j'}$ de ces valeurs s_j' ;
- Les estimations S_j sont ainsi obtenues en posant $S_j=\frac{s_j'}{m_{s'}}$.

	1^{er} trimestre	2^e trimestre	3^e trimestre	4^e trimestre
Année 1	224.3705	253.2811	201.2421	248.9411
Année 2	274.3802	300.1641	248.9038	298.4386
Année 3	331.9657	371.4032	303.4313	365.9029
Année 4	406.6326	437.9967	361.5774	444.8447
Année 5	488.4166	536.5268	435.5698	549.3614
Année 6	598.0016	659.2896	533.2156	669.2675

** Exemple: Soit la série chronologique suivante observée trimestriellement sur 6 ans qui définit un modèle multiplicatif, de période 4 (i.e. de longueur 4):

La série est soumise à des variations saisonnières de période 4. La tendance, caractérisée par les moyennes mobiles de longueur 4, est croissante et la différence entre une observation Yt et la moyenne mobile mm_t a tendance à augmenter pour une même variation saisonnière.

** Pour quantifier les variations saisonnières, on considère les rapports qui sont des approximations des termes $1 + s_j$ que l'on appelle les coefficients saisonniers dans le cas du modèle multiplicatif:

$$\frac{y_{ij}}{c_{ij}} = 1 + s_j + \frac{e_{ij}}{c_{ij}} \quad pour \;\; i = 1, ..., n \quad et \quad j = 1, ..., p.$$

où en admettant que les variations accidentelles e_{ij} sont faibles par rapport à la tendance c_{ij} et en utilisant l'approximation de la tendance par les moyennes mobiles- voir tableaux:

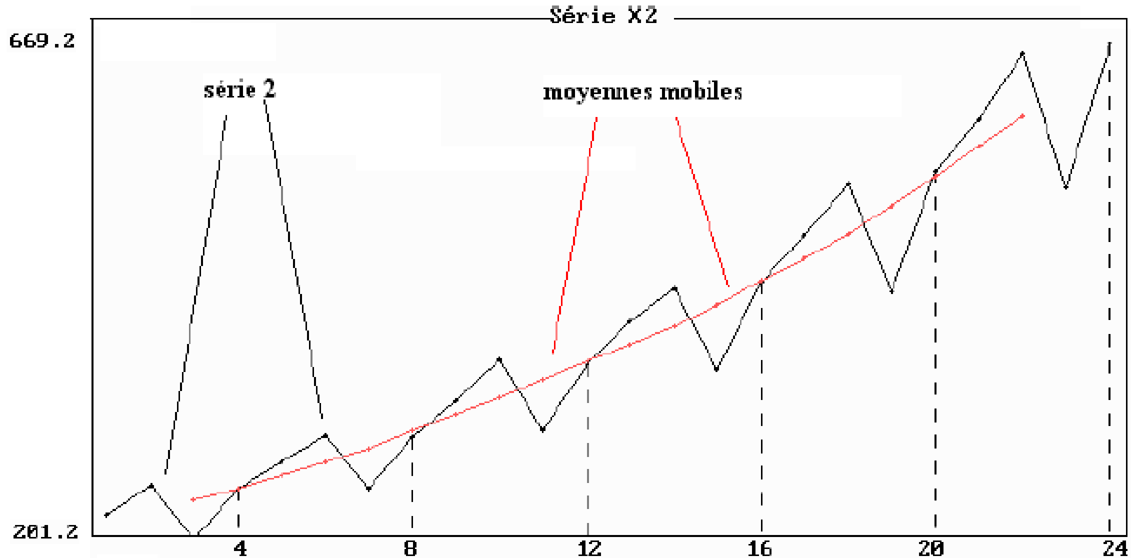
** On en déduit les moyennes suivantes:

$s_1^{'}$= 1.045913 ; $s_2^{'}$= 1.097236 ; $s_3^{'}$ = 0.8539006 et $s_4^{'}$= 0.9942986

** La moyenne $s_j^{'}$ est $m_{s_j^{'}}$ = .9978371

** Et, les valeurs définitives sont obtenues de façon que les $s_j^{'}$soient de moyenne 1 :

S_1 = 1.04818 ; S_2 = 1.099614 ; S_3= 0.8557515 et S_4= 0.9964539.



	1^{er} trimestre	2^{e} trimestre	3^{e} trimestre	4^{e} trimestre
Année 1			238.210	250.322
Année 2	262.140	274.284	287.670	303.773
Année3	319.494	334.743	352.509	370.167
Année 4	385.759	402.895	422.986	445.525
Année 5	467.090	489.404	516.167	545.210
Année 6	572.761	599.955		

Tableau 7.8 : moyennes mobiles de longueur 4 de la série 2

	1^{er} trimestre	2^{e} trimestre	3^{e} trimestre	4^{e} trimestre
Année 1			0.84481	0.99449
Année 2	1.04670	1.09435	0.86524	0.98244
Année3	1.03904	1.10952	0.86078	0.98848
Année 4	1.05411	1.08712	0.85482	0.99847
Année 5	1.04566	1.09629	0.84385	1.00761
Année 6	1.04407	1.09890		

ableau 8.8 : rapports des observations aux moyennes mobiles de la série .

$\Rrightarrow$ 2)$-$Méthode de lissage regroupe l'ensemble des techniques empiriques qui ont pour caractéristiques communes d'accorder un poids (une influence) plus important aux valeurs récentes de la chronique. Ces méthodes portent aussi le nom de filtrage, où il s'agit d'une opération mathématique transformant un entrant x_t en une nouvelle chronique sortante y_t .

L'idée générale est telle qu'**: on dispose d'une série chronologique** $x = (x_1, x_2, \cdots, x_n)$ **de longueur** n **enregistrée aux dates** $1, ..., n$**. On se situe à la date** n **et on souhaite prévoir la valeur** $x_n + h$ **non encore observée à l'horizon** h**. On note cette prévision** $\widehat{x}_{n\ ,\ h}$**. L'entier** n **est parfois appelé base de la prévision. Dans les techniques du lissage exponentiel, il s'agit d'ajuster à la chronique, localement, une fonction simple telle que:**

— une constante dans le lissage exponentiel simple,

— une droite dans le lissage exponentiel double,

— des fonctions polynomiales ou périodiques dans les lissages plus généraux.

En terme de modélisation, les méthodes de lissage exponentiel non saisonnier supposent que la chronique de départ est structurée comme suit:

$$
\begin{gathered}
x_t = f_t^{(k)} + \epsilon_t \\
\textit{avec} \\
E(\epsilon_t) = 0, \forall\, t \in \mathbb{Z} \\
E\left(\epsilon_t^2\right) = \sigma_\epsilon^2, \forall\, t \in \mathbb{Z} \\
E(\epsilon_t\ \epsilon_{t-h}) = 0, \forall\, h \neq 0
\end{gathered}
$$

$f_t^{(k)}$ *fonction polynômiale de degré k dont les paramètres dépendent du temps.*

Les méthodes de lissage se différencient entre elles selon le degré de $f_t^{(k)}$; les degrés les plus utilisés sont $k = 0$ et $k = 1$.

- Si $k = 0$, la série s'écrit:

$$
\begin{gathered}
x_t = a_t + \epsilon_t, \quad \forall\, t \\
\textit{avec } a_t = a \ \textit{ et } \ \textit{si } t \in \{n_1; n_2\} \subset \{\ 1;\ n\}
\end{gathered}
$$

$\{n_1; n_2\}$ *intervalle qui désigne le pas de la chronique ou la période de prévision.*

Et à ce type de modèle, on adopte la méthode de lissage exponentiel simple, *LES*.

- Si $k = 1$, la série s'écrit:

$$
\begin{gathered}
x_t = a_t + b_t\, t + \epsilon_t, \quad \forall\, t \\
\textit{avec } a_t = a\ ,\ b_t = b \textit{ et si } t \in \{n_1; n_2\} \subset \{\ 1;\ n\}
\end{gathered}
$$

La méthode employée pour ce type de modèle est le lissage exponentiel double, LED.

Lorsque k est quelconque, la méthode de lissage porte le nom de lissage exponentiel général, LEG.

1°) Principe du Lissage Exponentiel Simple, LES:

Dans cette partie, la prévision est très simple, ne dépend pas de l'horizon h de prévision.

1–1) Définition:

La prévision $\widehat{x}_{n\ ,\ h}$ est construite en prenant en compte toute l'histoire de la chronique, de sorte que, plus on s'éloigne de la base n de la prévision, moins l'influence des observations correspondantes est importante. Cette décroissance de l'influence est de type exponentiel. De là vient le nom de la technique.

Alors, $x_{\ t}$ peut être considéré comme le résultat d'une combinaison linéaire infinie de ses valeurs passées, où le poids / l'influence du passé sur le présent étant décroissant avec son ancienneté.

Et, la prévision $\widehat{x}_{n\ ,\ h}$ (notée aussi, $\widehat{x}_n\ (h)$) fournie par la méthode LES, avec une constante de lissage β est définie par:

$$\widehat{x}_{n\ ,\ h} = (1-\beta)\ \sum_{i\ =\ 0}^{n\ -\ 1} \beta^i\ x_{n\ -\ i} \qquad (1)$$
$$et\ 0 \ \langle\ \beta\ \langle\ 1$$

On donne un poids d'autant moins important que les observations sont loins (dans la passé), avec une décroissance exponentielle:

* Pour β proche de 1: **C'est la prise en compte de tout le passé** . Car, on a l'influence des observations éloignées dans le temps de la base n est plus grande ç-à-d que la mémoire du phénomène étudié est forte et la prévision est peu réactive aux dernières observations.

* Pour β proche de 0: **C'est la prise en compte d'avantage des valeurs récentes** (plus sensible aux fluctuations). C'est à dire que les observations les plus récentes ont un poids prépondérant (dominant) sur les termes anciens; la mémoire du phénomène est faible et le lissage est très réactif aux dernières observations.

Remarquons que si β ne dépend pas de h, $\widehat{x}_{n\ ,\ h}$ ne dépend pas de h et $\widehat{x}_{n\ ,\ h} = \widehat{x}_n$. Cette valeur $\widehat{x}_n$ est la prévision faite en n de la valeur en $n\ +\ 1$. On appelle cette série $\widehat{x}_n$ la valeur lissée de la chronique $x_{\ t}$ calculée en $t-1$

. Par hypothèse et lorsque les paramètres sont estimés, cette valeur lissée est considérée comme la valeur prévue ou prédite de x_t calculée en $t-1$ pour t.

1–2) Formules récursives de mise à jour (ou méthode adaptive):

La définition vérifie les formules de récurrences suivantes:

$$\widehat{x}_{n\ ,\ h} = (1-\beta)\ x_n\ + \beta\ \widehat{x}_{n\ -\ 1,\ h} \tag{2}$$

ou encore :

$$\widehat{x}_{n\ ,\ h} = \widehat{x}_{n\ -\ 1,\ h} + (1-\beta)\ \ (\ x_n\ -\ \widehat{x}_{n\ -\ 1,\ h}) \tag{3}$$

Par conséquent, l'équation (2) implique que la connaissance de la prévision à l'horizon sur la base $n-1$, soit $\widehat{x}_{n\ -\ 1,\ h}$ et l'observation x_n suffisent pour calculer immédiatement la prévision suivante $\widehat{x}_{n\ ,\ h}$. Alors sous la forme (2), le lissage apparaît comme étant une moyenne pondérée de la dernière réalisation et de la dernière valeur lissée.

Remarquons aussi que la formule (2) permet d'interpréter $\widehat{x}_{n\ ,\ h}$ comme le barycentre de x_n et $\widehat{x}_{n\ -\ 1,\ h}$ affectés respectivement des masses $(1-\beta)$ et β . Et, l'influence de la dernière observation x_n est d'autant plus grande que β est proche de 0.

Et, sous la forme (3), le lissage apparaît comme le résultat de la dernière valeur lissée corrigé par une pondération de l'écart entre la réalisation et la prévision.

Cependant, notons que l'utilisation des formules récursives nécessite d'initialiser la récurrence en prenant par exemple $\widehat{x}_{1\ ,\ h} = x_1$; mais, on remarque que pour n assez grand la valeur initiale a peu d'influence.

1–3) Choix et détermination de la constante de lissage, β:

La méthode n'est pas bonne en présence de tendance, de composante saisonnière ou de fluctuations de hautes fréquences. En pratique, le lissage exponentiel simple est assez peu utilisé.

* En effet, si on est tenté de choisir des valeurs de β comprises entre 0.7 et 0.99,ç-à-d β proche de 1 qui tiennent compte des prévisions du passé lointain et immédiat et augmentent la réactivité (ç-à-d la réapparition du phénomène qui ne se manifestait plus); alors, ces valeurs de β seront rigides et peu sensibles aux variations conjoncturelles où le rôle de filtrage (ç-à-d le contrôle des aléas que doit assurer le lissage) sera inopérant et inefficace. En effet, à partir de la formule du lissage, on vérifie ce résultat:

Lorsque β proche de 1 $\Longrightarrow$ $\widehat{x}_t = x^p_{t\ -\ 1,\ t} = x_{t\ -1}$ *et* $var\,(x_t) = var\,(\widehat{x}_t) = \sigma^2_\epsilon$; dans ce cas, on parle de prévision naïve.

$*$ Il n'en n'est plus de même si β est compris entre 0.01 et 0.3, ç-à-d β faible; car on assure un rôle de filtrage efficace où la prévision sera souple c'est à dire influencée par les observations récentes et ainsi, on perd en réactivité.

$*$ Pour déterminer la constante de lissage β, on définit deux approches de calculs permettant en partie de répondre au dilemme précédent:

(i°)- La valeur du coefficient de lissage est celle qui minimise la somme des carrés des erreurs de prévision passée (qui est la somme des carrés des écarts prévisionnels), c'est à dire la valeur qui en moyenne réalise la meilleure prévision à l'horizon h sur les bases 1, 2, ..., $n - h$ en utilisant le critère des moindres carrés:

$$Min \sum_{t\ =\ 1}^{n} e^2_t = Min \sum_{t\ =\ 1}^{n} (x_t\ - \widehat{x}_t)^2$$

C'est la technique de calcul la plus couramment employée.

(ii°)- Procédure de régulation d'un coefficient de lissage: A partir d'un certain nombre d'indicateurs (amplitude de l'erreur de prévision, prévision biaisée, ...) cette procédure de régulation détermine la valeur idéale du coefficient de lissage à un instant donné. Contrairement à la méthode précédente, dans laquelle un coefficient optimal sur la totalité de la période historique est déterminé, cette approche permet d'estimer la valeur jugée la plus efficace à l'instant où la prévision est calculée:

1–4) Initialisation des calculs: Dans la formule (2) du LES:

$$\widehat{x}_{n\ ,\ h} = (1 - \beta)\ x_n\ + \beta\ \widehat{x}_{n\ -\ 1,\ h}$$

- Pour $t\ = 1$, $\widehat{x}_1 = (1 - \beta)\ x_0\ + \beta\ \widehat{x}_0$, on a *la valeur brute* x_0 et la valeur lissée $\widehat{x}_0 = x^p_{-\ 1,\ 0}$ sont inconnus.
- Pour $t\ = 2$, $\widehat{x}_2 = (1 - \beta)\ x_1\ + \beta\ \widehat{x}_1$, où seul $\widehat{x}_1$est inconnu.

Plusieurs solutions existent pour démarrer les calculs à la période $t\ = 2$. Et, on retient la méthode la plus simple et la plus utilisée est celle de poser $\widehat{x}_1 =\ x_1$.

1–5) Intervalle de confiance de la prévision: on a d'après la formule (1) :

$$var\,(\widehat{x}_t) = var\left(\lambda \sum_{i\,=\,0}^{\infty} (1-\lambda)^i\; x_{t\;-1-\;i}\right) \qquad avec\ \lambda = 1-\beta$$
$$= \lambda^2\ var\left(\sum_{i\,=\,0}^{\infty} (1-\lambda)^i\; x_{t\;-1-\;i}\right)$$
$$or,\ x_t\ = a_t\ +\ \epsilon_t,\ \ \forall\ t$$
$$alors\ sur\ un\ intervalle\ local, la\ var\,(x_t\ ) = \sigma_\epsilon^2$$
$$et$$
$$\forall\ t\ et\ t',\ \ t \neq t',\ les\ variables\ x_t\ \ et\ x_{t'} sont\ indépendantes.$$
$$D'où :$$
$$var\,(\widehat{x}_t) = \lambda^2\ var\,(x_t\ ) \sum_{i\,=\,0}^{\infty} (1-\lambda)^{2\ i} = \lambda^2\ var\,(x_t\ )\ \left[1+(1-\lambda)^2+(1-\lambda)^4+\cdots\right]$$
$$= \lambda^2\ var\,(x_t\ )\,\tfrac{1}{1-(1-\lambda)^2}$$
$$= var\,(x_t\ )\,\tfrac{\lambda}{2-\lambda}$$
$$Donc,\ on\ a :$$
$$\tfrac{var(\widehat{x}_t)}{var(x_t\ )} = \tfrac{\lambda}{2-\lambda}\ ;\ et\ dans\ le\ cas\ où\ \lambda = 1, \widehat{x}_t\ a\ la\ même\ variance\ que\ \ x_t.$$

Alors, le résidu de la prévision effectuée en $t-1$ pour t est défini par:

$$e_t\ = x_t\ -\ \widehat{x}_t$$
$$\Longrightarrow\ \ var\,(e_t\ ) = var\,(x_t\ ) + var\,(\ \widehat{x}_t)\ \ \ car\ x_t\ \ et\ \ \widehat{x}_t\ sont\ indépendantes.$$
$$\Longleftrightarrow\ var\,(e_t\ ) = \tfrac{2\ \sigma_x^2}{2\ -\ \lambda}$$

L'intervalle de confiance à un niveau $1-\alpha$ de la valeur prévue de x_t en $t-1$ (ou de $t = n+1$ à $n+h$) est défini par:

$$IC^{1-\alpha}\,(.) = \pm\ t^{\alpha/2}\ \sigma_x\ \sqrt{\tfrac{2}{2\ -\ \lambda}}$$
$$t^{\alpha/2}\ valeur\ lue\ sur\ la\ table\ de\ Laplace\ Gauss.$$