

TRAITEMENT DE L'EXEMPLE

Estimation de la tendance

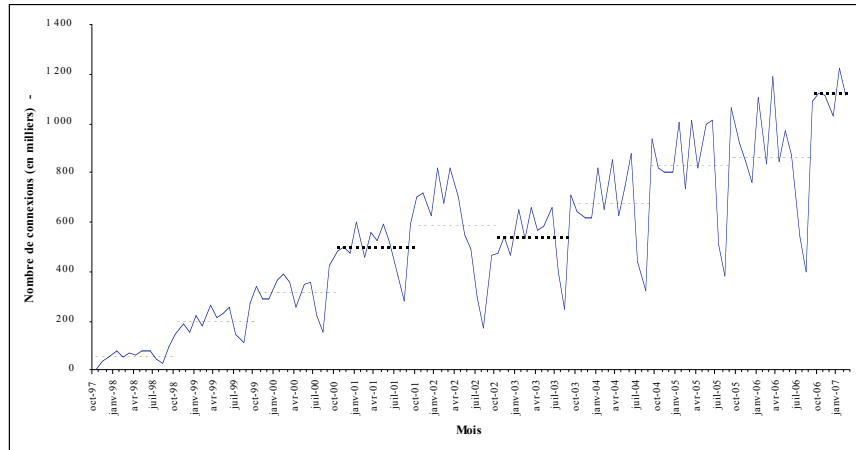
On détermine les valeurs de la tendance T par ajustement.

Plusieurs types d'ajustement sont possibles :

- **Ajustement par les moyennes annuelles**

On ajuste chaque année par la moyenne annuelle des observations correspondantes.

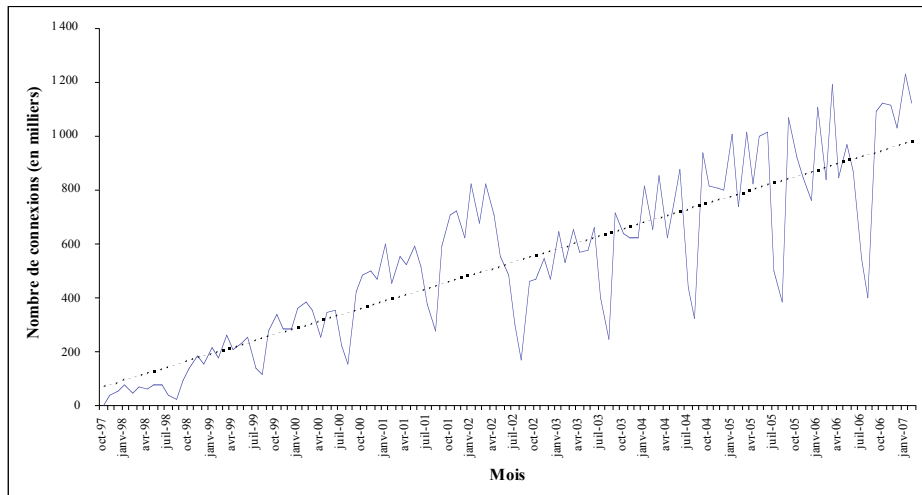
1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
55 531	198 233	314 536	495 655	586 667	540 530	679 153	831 231	865 984	1 122 004



- **Ajustement par la méthode des moindres carrés**

On ajuste les points de la série par la droite des moindres carrés.

Une équation de cette droite est $Y = 8\,378,4t + 57\,122,9$ où t, le rang du mois, varie de 1 à 113.

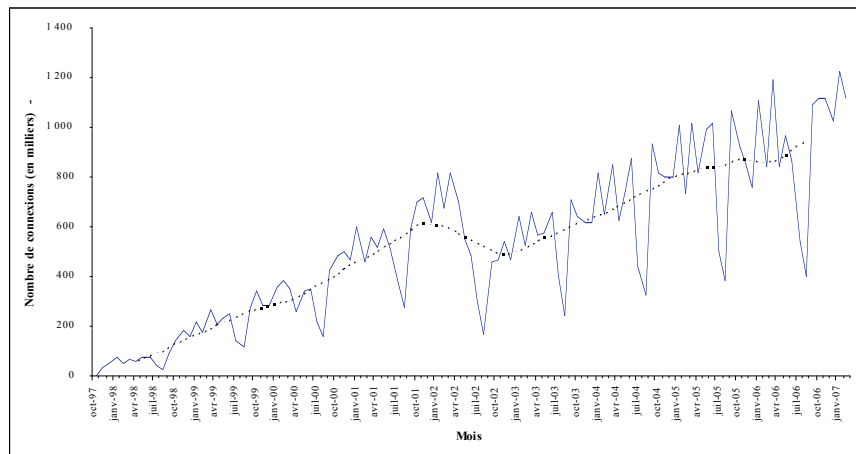


• **Ajustement par la méthode des moyennes mobiles**

On utilise la méthode des moyennes mobiles (éventuellement centrées) en affectant, quand c'est possible, à chaque période la moyenne mobile correspondante.

Détermination des moyennes mobiles centrées d'ordre 12 :

	1997-1998	1998-1999	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004	2004-2005	2005-2006
Octobre		133 158,71	272 611,83	413 154,38	611 330,96	494 027,63	616 114,42	769 335,33	873 323,96
Novembre		145 948,67	279 224,83	434 356,50	617 372,71	489 575,13	625 323,58	788 048,13	872 997,67
Décembre		159 805,21	287 955,46	451 473,13	614 471,29	497 999,21	641 266,71	804 341,92	865 657,42
Janvier		171 222,50	295 356,79	464 892,50	609 982,71	509 234,96	652 056,21	812 833,79	861 090,13
Février		179 204,42	300 345,63	476 506,21	602 110,92	516 517,42	657 182,42	817 897,21	863 373,71
Mars		190 643,17	308 300,04	488 600,33	592 080,08	530 135,67	669 852,42	825 764,88	865 004,83
Avril	61 288,92	206 546,25	320 433,96	504 901,08	576 962,58	547 517,25	686 549,46	835 728,50	874 139,25
Mai	73 294,17	218 924,25	335 282,46	523 304,29	559 997,08	557 626,50	701 642,88	842 078,42	893 386,88
Juin	83 833,58	228 473,67	351 826,79	538 786,71	546 373,63	567 006,08	716 948,75	842 299,50	915 498,38
Juillet	94 126,96	239 854,04	369 487,13	554 275,13	532 718,25	580 365,17	732 550,83	844 746,21	931 523,38
Août	105 540,63	254 483,58	382 419,75	572 558,63	519 390,38	592 549,71	743 970,83	853 228,21	948 083,88
Septembre	118 946,63	266 908,25	393 754,46	592 651,00	506 541,25	605 724,08	754 241,92	865 070,38	



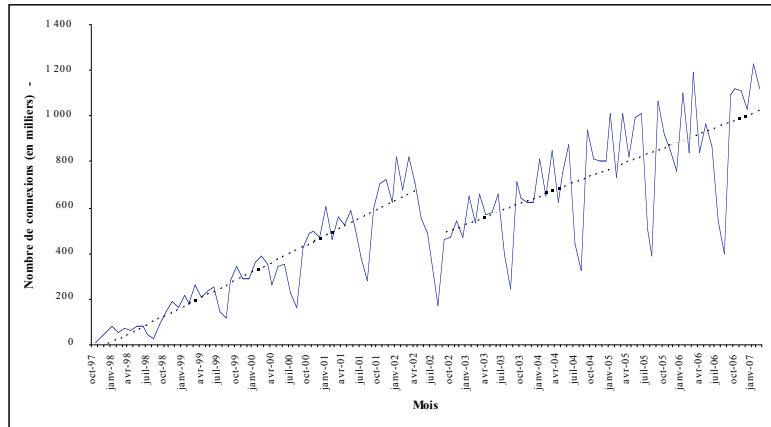
Si cette méthode permet un lissage des valeurs de la chronique, elle présente cependant plusieurs inconvénients :

- on perd les données de début et de fin de série, on réserve la méthode aux longues chroniques ;
- les moyennes mobiles peuvent engendrer des cycles et des mouvements qui n'étaient pas présents dans les données d'origine ;
- les moyennes mobiles sont fortement affectées par les valeurs "aberrantes" accidentelles.

• **Ajustement "raisonné"**

On peut constater sur le diagramme de l'ajustement par les moyennes mobiles centrées d'ordre 12 que les données sont classées en deux groupes dans lesquels la tendance est linéaire : jusqu'à avril 2002, puis à partir de septembre 2002. On peut ajuster les deux séries par la méthode des moindres carrés. Une équation de la courbe

d'ajustement obtenue est :
$$\begin{cases} Y = 13\,016,50 t + 40\,128,4 & \text{pour } t \text{ variant de } 5 \text{ à } 55 \\ Y = 8\,746,0 t - 25\,727,4 & \text{pour } t \text{ variant de } 60 \text{ à } 113 \end{cases}$$



Estimation des variations saisonnières, indice saisonnier

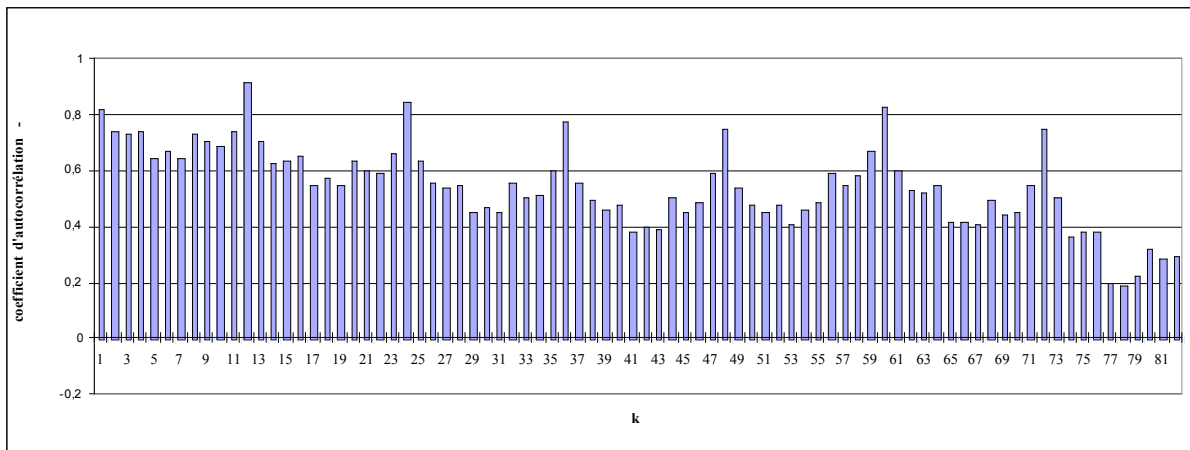
Corrélogramme

Lorsque la saisonnalité n'apparaît pas nettement sur le graphique cartésien, on utilise un autre graphique appelé **corrélogramme**. C'est le diagramme en bâton représentant la **fonction d'autocorrélation** qui, à $k \in \mathbb{N}$, associe

$\rho(k)$, le **coefficient d'autocorrélation d'ordre k**, défini par $\rho(k) = \frac{\text{Cov}(Y_i, Y_{i+k})}{\sigma(Y_i) \sigma(Y_{i+k})}$ où Y_{i+k} est la chronique

obtenue en décalant la série initiale de k unités de temps.

$\rho(k)$ est le coefficient de corrélation entre la chronique Y_i et la chronique Y_{i+k} ; lorsque $\rho(k)$ est voisin de 1, Y_i et Y_{i+k} sont voisines et la chronique présente une saisonnalité de période k.



Corrélogramme de la série de l'exemple.

Détermination du modèle

Dans un modèle multiplicatif, l'effet saisonnier et l'aléatoire sont amplifiés (ou diminués) proportionnellement à la croissance (ou décroissance) de la tendance. Ceci peut se voir sur le graphique cartésien de la chronique en observant les fluctuations autour de la tendance.

On utilise aussi un autre graphique appelé **diagramme mu-sigma** : si la saisonnalité est d'ordre p, on calcule les moyennes μ et les écarts types σ de tous les ensembles consécutifs de p valeurs de la chronique ; on porte alors sur un diagramme cartésien les points d'abscisses μ et d'ordonnées σ .

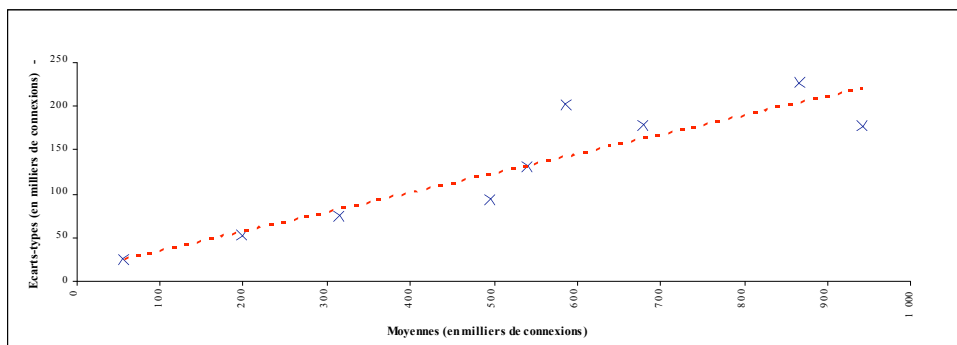


Diagramme mu-sigma de la série de l'exemple.

Dans le cas du modèle additif, les points auront tendance à être alignés horizontalement (les fluctuations restant constantes), tandis que dans le modèle multiplicatif, il s'aligneront à peu près sur une droite oblique.

On peut adopter un modèle multiplicatif pour l'exemple.

Estimation des variations saisonnières

On suppose que la saisonnalité est de période p et que l'étude de la série chronologique se fait sur n périodes.

- **Cas du modèle additif :**

$Y = T + C + S + I$ et $Z = Y - T = C + S + I$ donne une bonne approximation de S mais n'est pas périodique en général. On calcule alors les p coefficients $X(i)$ moyenne ou médiane de $Z(i), Z(i+p), \dots, Z(i+(n-1)p)$ pour i entier compris entre 1 et p .

Si la somme des $X(t)$ n'est pas 0, on procède à un ajustement : on obtient $S(i) = X(i) - \frac{X(1) + X(2) + \dots + X(p)}{p}$.

- **Cas du modèle multiplicatif :**

$Y = T \times C \times S \times I$ et $Z = \frac{Y}{T} = C \times S \times I$ donne une approximation de S mais n'est pas périodique en général.

On calcule alors les p coefficients $X(i)$ moyenne ou médiane de $Z(i), Z(i+p), \dots, Z(i+(n-1)p)$ pour i entier compris entre 1 et p .

Si la somme des $X(t)$ n'est pas p , on procède à un ajustement : on obtient $S(i) = \frac{X(i) \times p}{X(1) + X(2) + \dots + X(p)}$.

Les coefficients $S(i)$ sont donnés sous forme de pourcentage et portent alors le nom d'**indices saisonniers**.

Il y a autant de méthodes de calcul des indices saisonniers que de méthodes de détermination de la tendance, mais tous reposent sur le même principe, exposé ici dans le cas d'un modèle multiplicatif qui nous intéresse :

1. On détermine la tendance des observations.
2. On exprime chaque donnée comme pourcentage de la valeur de tendance annuelle correspondante.
3. On calcule une valeur centrale (moyenne ou médiane) des séries obtenues sur un même mois.
4. On ajuste pour que la moyenne des indices obtenus soit 100, en multipliant par un facteur convenable.

• Avec l'ajustement "raisonné"

Les valeurs de l'ajustement sont données par : $\begin{cases} Y = 13\,016,50t + 40\,128,4 & \text{pour } t \text{ variant de } 5 \text{ à } 55 \\ Y = 8\,746,0t - 25\,727,4 & \text{pour } t \text{ variant de } 60 \text{ à } 113 \end{cases}$

Y _a	Oct.	Nov.	Déc.	Janv.	Fév.	Mars	Avr.	Mai	Juin	Juil.	Août	Sept.
2001-2002												493 146,99
2002-2003	503 242,22	513 337,44	523 432,66	533 527,88	543 623,10	553 718,32	563 813,55	573 908,77	584 003,99	594 099,21	604 194,43	614 289,66
2003-2004	624 384,88	634 480,10	644 575,32	654 670,54	664 765,76	674 860,99	684 956,21	695 051,43	705 146,65	715 241,87	725 337,09	735 432,32
2004-2005	745 527,54	755 622,76	765 717,98	775 813,20	785 908,42	796 003,65	806 098,87	816 194,09	826 289,31	836 384,53	846 479,76	856 574,98
2005-2006	866 670,20	876 765,42	886 860,64	896 955,86	907 051,09	917 146,31	927 241,53	937 336,75	947 431,97	957 527,19	967 622,42	977 717,64
2006-2007	987 812,86	997 908,08	1 008 003,30	1 018 098,52	1 028 193,75							

Indices saisonniers en pourcentage : exemple de calcul pour la première valeur $463\,407 = 493\,146,99 \times 94\%$.

Z	Oct.	Nov.	Déc.	Janv.	Fév.	Mars	Avr.	Mai	Juin	Juil.	Août	Sept.
2001-2002												94,0
2002-2003	93,8	106,2	89,7	121,3	97,7	118,5	101,0	101,1	112,9	67,0	40,5	116,0
2003-2004	102,5	97,7	96,1	124,9	98,2	126,0	91,2	107,4	124,3	61,6	44,9	127,3
2004-2005	109,6	106,5	104,8	130,1	93,6	127,5	102,0	122,2	123,0	60,3	45,2	124,6
2005-2006	106,8	96,9	86,1	123,5	92,7	130,1	90,8	103,5	91,6	56,7	41,2	111,6
2006-2007	113,5	111,8	102,0	120,6	108,8							
X moyenne	105,2	103,8	95,7	124,1	98,2	125,5	96,3	108,5	112,9	61,4	42,9	114,7
S moyenne ajustée	106,2	104,8	96,6	125,2	99,1	126,7	97,1	109,5	114,0	106,2	104,8	96,6

On travaillera désormais avec l'estimation de la tendance par l'ajustement "raisonné" et on utilisera les indices saisonniers correspondants calculés avec la moyenne.

Désaisonnalisation des données

Quand, selon le modèle, on retranche S à Y ou on divise Y par S, on dit qu'on ajuste les données suivant les variations saisonnières. Les nouvelles données prennent en compte la tendance, les mouvements cyclique et irrégulier. On divise Y par S pour obtenir la tendance et les variations cycliques et irrégulières.

	Oct.	Nov.	Déc.	Janv.	Fév.	Mars	Avr.	Mai	Juin	Juil.	Août	Sept.
2001-2002												400 463,3
2002-2003	444 648,0	520 479,0	486 164,2	517 053,1	536 027,2	518 168,4	586 277,0	529 969,8	578 810,1	642 702,8	564 321,3	616 048,2
2003-2004	602 601,5	592 010,7	641 649,3	653 198,0	659 165,0	671 495,4	643 088,6	681 414,8	769 042,3	710 756,0	750 957,5	808 943,5
2004-2005	769 811,6	768 350,3	830 719,2	806 431,9	742 184,8	801 186,8	846 959,7	910 730,0	891 827,7	813 863,0	883 991,1	922 308,2
2005-2006	871 489,3	810 785,9	790 187,2	884 607,3	848 880,9	942 129,8	867 115,1	885 706,2	761 286,8	877 048,2	920 131,6	942 605,8
2006-2007	1 055 864,5	1 064 904,1	1 064 007,9	980 563,5	1 128 819,9							

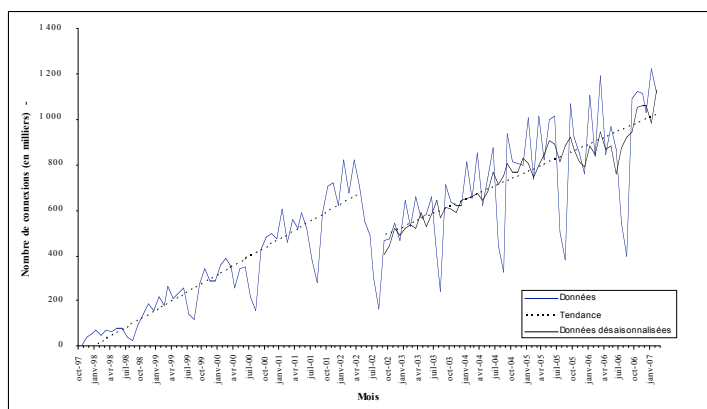


Diagramme cartésien des données désaisonnalisées

Estimation des variations cycliques

Une fois que les données ont été ajustées suivant la variation saisonnière, on peut encore les ajuster suivant la tendance. On obtient alors, selon le modèle, $C + I$ ou $C \times I$. On lisse les variations irrégulières par une moyenne mobile bien choisie. On obtient alors C . S'il y a périodicité (ou périodicité approchée) des cycles, on procède comme les variations saisonnières.

La mise en évidence des variations cycliques se fait alors comme pour les variations saisonnières. Ici, il n'y a pas lieu de regarder ce phénomène en raison de la courte durée de la série.

Estimation des variations irrégulières ou aléatoires

Les variations irrégulières ou aléatoires sont obtenues en ajustant les données suivant la tendance et les variations saisonnières et cycliques. On divise les données précédentes par T . Comme on trouve, ici, des valeurs très proches de 1, elles sont négligeables.

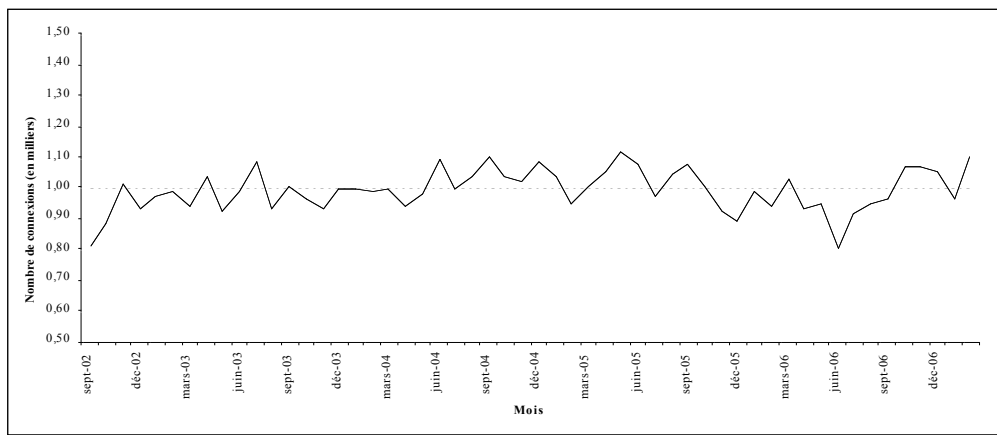


Diagramme cartésien des variations aléatoires

Prévision

Le problème de la prévision repose sur deux questions fondamentales :

- Quelle(s) est (sont) la (les) grandeur(s) à prévoir ?
- Pour quoi faire ?

Le problème se pose alors dans les termes suivants :

on connaît x_1, x_2, \dots, x_N et on cherche $x_{N+1}, x_{N+2}, \dots, x_{N+h}$, h étant l'**horizon** de la prévision.

Les idées précédentes permettent de résoudre le problème important de la prévision des séries temporelles. Cependant, il faut prendre garde que le traitement mathématique des données ne résout pas tous les problèmes. L'analyse mathématique est néanmoins appréciable si le statisticien y associe du bon sens, de l'ingéniosité et un bon jugement.

Prévision obtenue pour 1982 en utilisant les résultats précédents : $Y_p = T S = (8\,746,0 t - 25\,727,4) \times S$ où t , rang du mois, varie de 114 à 84.

Y_a	Mars	Avr.	Mai	Juin	Juil.	Août	Sept.	Oct.	Nov.	Déc.	Janv.	Fév.
2007-2008	1 230 231	951 742	1 082 861	1 136 794	623 463	439 801	1 184 709	1 096 172	1 090 811	1 014 176	1 325 499	1 057 546

Résultat à comparer aux valeurs réellement observées :

Y_a	Mars	Avr.	Mai	Juin	Juil.	Août	Sept.	Oct.	Nov.	Déc.	Janv.	Fév.
-------	------	------	-----	------	-------	------	-------	------	------	------	-------	------

