

Séries chronologiques

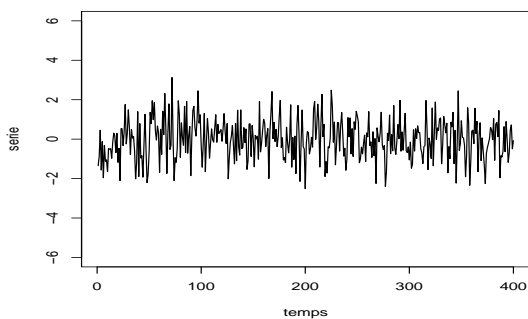
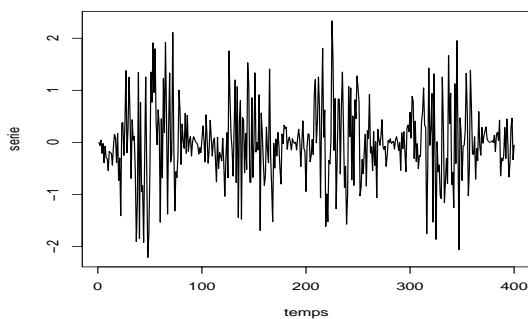
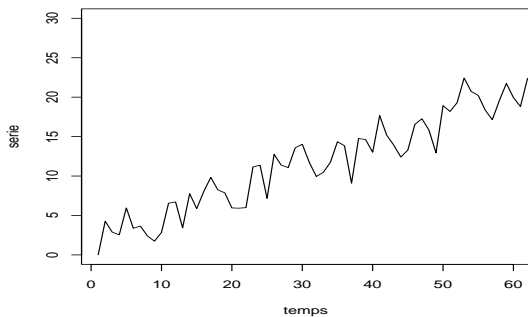
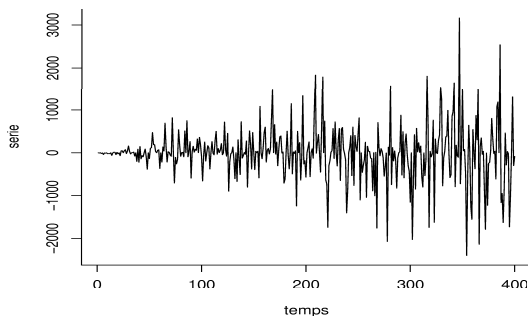
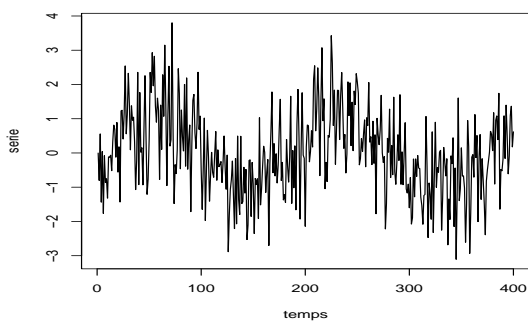
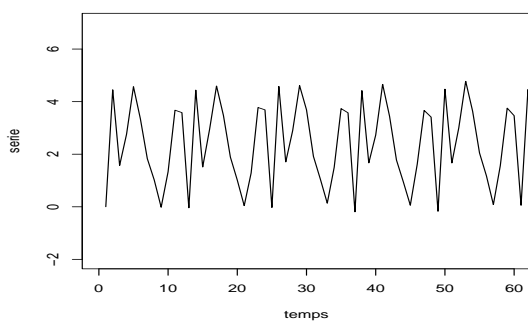
10 janvier 2007

1 Exercice 1

Nous avons simulé les séries suivantes, où ε_t est un bruit aléatoire, s_t une série d'effets saisonniers, f_t une tendance sinusoidale et a une constante positive :

$$y_t = \varepsilon_t, \quad y_t = s_t + \varepsilon_t, \quad y_t = at + s_t + \varepsilon_t, \\ y_t = at \times s_t \times \varepsilon_t, \quad y_t = f_t \times \varepsilon_t, \quad y_t = f_t + \varepsilon_t,$$

Associez à chaque série son graphique, en justifiant.



2 Exercice 2 : Utilisation des moyennes mobiles centrées sur une tendance linéaire

Soit deux séries chronologiques X et Y , comprenant une tendance linéaire, définies de la manière suivante :

$$\begin{aligned}x_t &= 2 \times t + 100 \\y_t &= 3 \times t + 200 \quad t = 1 \dots, 16\end{aligned}$$

1. Calculer les moyennes mobiles de longueur 3 et de longueur 4 sur les séries X et Y , notées $mm_3(X)$, $mm_4(X)$, $mm_3(Y)$ et $mm_4(Y)$.
2. Comparer la série X aux deux moyennes mobiles $mm_3(X)$ et $mm_4(X)$. Commenter les résultats.
3. Soit la série Z la somme des deux séries chronologiques X et Y .
 - 3.1. Calculer les moyennes mobiles de longueur 3 et de longueur 4 sur la série Z , notées $mm_3(Z)$ et $mm_4(Z)$.
 - 3.2. Calculer les sommes $mm_3(X) + mm_3(Y)$ et $mm_4(X) + mm_4(Y)$. Comparer ces résultats à ceux de la question précédente. Que peut-on constater ?

Séries	X	Y	Z
1	102	203	305
2	104	206	310
3	106	209	315
4	108	212	320
5	110	215	325
6	112	218	330
7	114	221	335
8	116	224	340
9	118	227	345
10	120	230	350
11	122	233	355
12	124	236	360
13	126	239	365
14	128	242	370
15	130	245	375
16	132	248	380

TAB. 1 – Valeurs des séries chronologiques X , Y et Z

3 Exercice 3 : Lissage par moyenne mobile simple et désaisonnalisation

Nous donnons au tableau 2 la série y_t des indices trimestriels de la production industrielle (base 100 en 1970) de 1963 à 1982 :

Année	Trimestre 1	Trimestre 2	Trimestre 3	Trimestre 4
1970	100	104	87	110
1971	107	108	92	117
1972	114	115	96	123
1973	122	122	103	130
1974	129	129	108	126
1975	117	117	97	125
1976	127	127	108	134
1977	133	130	107	132
1978	133	134	110	140
1979	138	136	118	146
1980	145	138	115	141

TAB. 2 – indice trimestriel de la production industrielle

1. A quoi cela sert-il de désaisonnaliser la chronique étudiée ? A quel type de questions cherche-t-on à répondre lorsque l'on décompose une chronique ?

2. Au vu du graphique, quel modèle de décomposition proposez-vous pour cette série ?

3. Nous allons adopter un schéma de décomposition additif pour cette série :

$$y_t = f_t + s_t + \varepsilon_t, \tag{1}$$

où f_t est la tendance de la série, s_t le coefficient saisonnier et ε_t la perturbation erratique. Rappelez la signification et les hypothèses habituelles faites pour ces différentes composantes. En particulier, rappelez le principe de conservation des aires pour les coefficients saisonniers. Pourquoi cette contrainte de renormalisation est-elle nécessaire ?

4. Estimation de la tendance par une moyenne mobile simple centrée sur 4 points

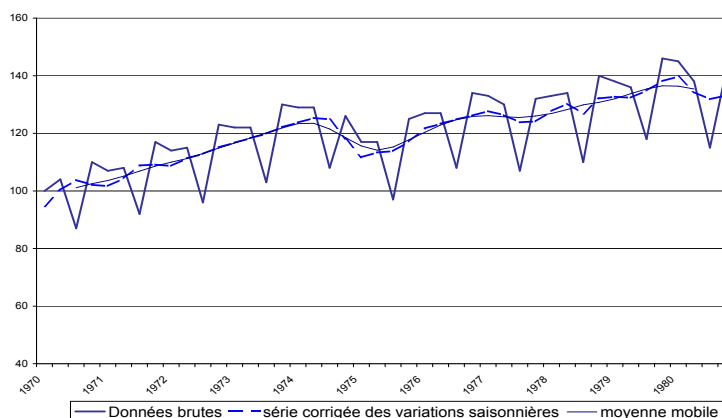


FIG. 1 – Indice trimestriel de la production industrielle

4.1. Rappelez le principe de la moyenne mobile simple centrée sur 4 points.

4.2. Calculez les valeurs manquantes du tableau 3 (année 1975).

4.3. Calculez les valeurs manquantes du tableau 4 donnant des estimations provisoires des coefficients saisonniers (année 1979).

4.4. Complétez le tableau 5 donnant les estimations provisoires et finale des coefficients saisonniers et commentez les valeurs trouvées :

Année	Trimestre 1	Trimestre 2	Trimestre 3	Trimestre 4
1970			101,125	102,500
1971	103,625	105,125	106,875	108,625
1972	110,000	111,250	113,000	114,875
1973	116,625	118,375	120,125	121,875
1974	123,375	123,500	121,500	118,500
1975				
1976	120,375	122,875	124,750	125,875
1977	126,125	125,750	125,500	126,000
1978	126,875	128,250	129,875	130,750
1979	132,000	133,750	135,375	136,500
1980	136,375	135,375		

TAB. 3 – Moyenne mobile simple centrée sur 4 points

Année	Trimestre 1	Trimestre 2	Trimestre 3	Trimestre 4
1970			-14,125	7,500
1971	3,375	2,875	-14,875	8,375
1972	4,000	3,750	-17,000	8,125
1973	5,375	3,625	-17,125	8,125
1974	5,625	5,500	-13,500	7,500
1975	1,375	2,875	-18,250	7,250
1976	6,625	4,125	-16,750	8,125
1977	6,875	4,250	-18,500	6,000
1978	6,125	5,750	-19,875	9,250
1979				
1980	8,625	2,625		

TAB. 4 – Première estimation des coefficients saisonniers

Estimation	1er trim	2eme trim	3eme trim	4eme trim
provisoire			-16,738	
finale		3,663		7,875

TAB. 5 – Estimation finale des coefficients saisonniers

4.5. Calculez les valeurs manquantes (année 1974) de la série corrigée des variations saisonnières (tableau 6).

Année	Trimestre 1	Trimestre 2	Trimestre 3	Trimestre 4
1970	94,700	100,338	103,838	102,125
1971	101,700	104,338	108,838	109,125
1972	108,700	111,338	112,838	115,125
1973	116,700	118,338	119,838	122,125
1974				
1975	111,700	113,338	113,838	117,125
1976	121,700	123,338	124,838	126,125
1977	127,700	126,338	123,838	124,125
1978	127,700	130,338	126,838	132,125
1979	132,700	132,338	134,838	138,125
1980	139,700	134,338	131,838	133,125

TAB. 6 – Série corrigée des variations saisonnières

4 Exercice 4 : Régression et désaisonnalisation pour une tendance linéaire

On reprend les données de l'exercice 3 à partir du premier trimestre de l'année 1976. Notre objectif est maintenant d'obtenir la tendance par une régression linéaire, puis d'en déduire à partir de la série corrigée les coefficients saisonniers.

1. Pourquoi a-t-on choisi la période premier trimestre 1976- dernier trimestre 1982 pour faire la régression linéaire ?

2. On modélise l'indice trimestriel y_t de la production industrielle par : $y_t = at + b$ pour $t = 1, \dots, 28$ correspondant à chacun des trimestres entre 1976 et 1982. On donne :

$$\sum_{t=1}^{28} t = 406, \quad \sum_{t=1}^{28} y_t = 3647, \quad \sum_{t=1}^{28} t^2 = 7714, \quad \sum_{t=1}^{28} y_t^2 = 478765, \quad \sum_{t=1}^{28} ty_t = 53500.$$

Calculer $\text{cov}(Y, T)$, $V(T)$, $V(Y)$, \bar{Y} , \bar{T} . Calculer les estimations des moindres carrés \hat{a} et \hat{b} de a et b . Calculez le coefficient de détermination R^2 de la régression.

3. Comment calculer les coefficients saisonniers ? Pourquoi le principe de conservation des aires est-il automatiquement satisfait ?

4. Compléter les tableaux suivants :

T	Y	$\hat{Y} = \hat{a}T + \hat{b}$	$\hat{\varepsilon} = Y - \hat{Y}$	Trimestre
1	127	125,680	1,320	1
2	127	126,018	0,982	2
3	108	126,357	-18,357	3
4	134	126,695	7,305	4
5	133		5,966	1
6	130		2,628	2
7	107		-20,711	3
8	132		3,950	4
9	133		4,612	1
10	134	128,727	5,273	2
11	110	129,065		3
12	140	129,404		4
13	138	129,742		1
14	136	130,081		2
15	118	130,419		3
16	146	130,758	15,242	4
17	145	131,096	13,904	1
18	138	131,435	6,565	2
19	115	131,773	-16,773	3
20	141	132,112	8,888	4
21	137	132,450	4,550	1
22	136	132,789	3,211	2
23	115	133,128	-18,128	3
24	143	133,466	9,534	4
25	137	133,805	3,195	1
26	136	134,143	1,857	2
27	111	134,482	-23,482	3
28	140	134,820	5,180	4

TAB. 7 – Valeurs expliquées par la régression et estimation provisoire des coefficients saisonniers

Trimestres	1	2	3	4
Coefficients saisonniers		3,776	-18,419	

5. Comparer avec les valeurs trouvées à l'exercice 3. Interprétez.

6. Comment utiliser les résultats précédents pour faire des prévisions ? Prévoir les valeurs de l'indice pour l'année 1987.



FIG. 2 – Séries Y (Données) \hat{Y} (Régression) et \hat{Y} redressée par les coefficients saisonniers.

5 Exercice 5 : Un modèle multiplicatif sur une série mensuelle

Une compagnie aérienne régionale désire connaître la structure du trafic aérien d'une de ses lignes. Pour cela, elle fournit la série mensuelle du nombre de passagers entre 2000 et 2004.

Mois	2000	2001	2002	2003	2004
janvier	713	800	900	1092	1070
février	630	700	820	1000	1030
mars	940	1030	1190	1380	1440
avril	1040	1280	1450	1700	1720
mai	720	840	880	960	1060
juin	1260	1520	1730	1950	2200
juillet	812	1010	1034	1203	1190
août	380	510	590	660	730
septembre	870	1060	1203	1282	1278
octobre	1120	1280	1500	1600	1760
novembre	1200	1350	1550	1700	1940
décembre	910	1000	1140	1160	1320

TAB. 8 – Trafic mensuel d'une ligne aérienne en nombre de passagers

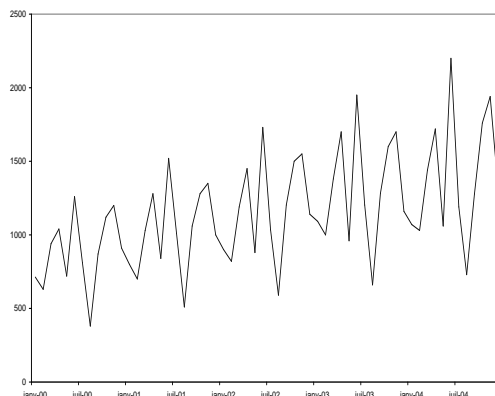


FIG. 3 – Trafic mensuel d'une ligne aérienne en nombre de passagers

1. Au vue du graphique, quel type de modèle faut-il appliquer ?
2. Après avoir justifié la longueur de la moyenne mobile à employer, calculer la série Y des moyennes mobiles.
3. Calculer les coefficients saisonniers et donner la série Z corrigée des variations saisonnières. Vérifier que ces coefficients vérifient les hypothèses de départ.

6 Exercice 6 : Préviation par régression

Nous nous intéressons à la consommation C_t de vin en France (en millions d'hectolitres) de $t = 1967$ à 1994.

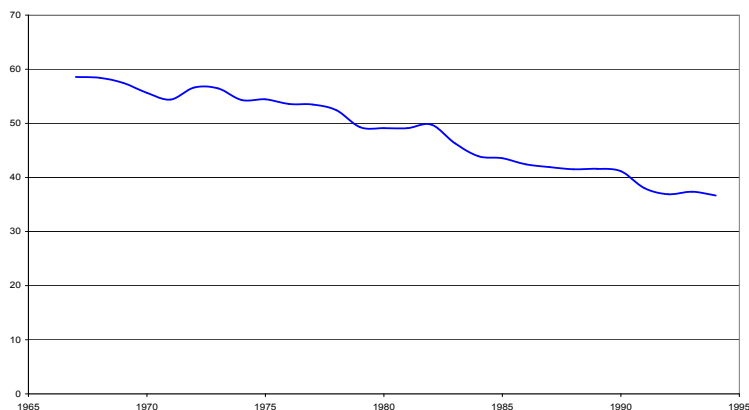


FIG. 4 – Consommation de vin en France (en million d'hectolitres)

Année	T	C	Année	T	C	Année	T	C	Année	T	C
1967	0	58,561	1974	7	54,317	1981	14	49,100	1988	21	41,500
1968	1	58,413	1975	8	54,427	1982	15	49,755	1989	22	41,600
1969	2	57,463	1976	9	53,580	1983	16	46,300	1990	23	41,157
1970	3	55,634	1977	10	53,479	1984	17	43,906	1991	24	38,019
1971	4	54,389	1978	11	52,408	1985	18	43,550	1992	25	36,903
1972	5	56,632	1979	12	49,268	1986	19	42,420	1993	26	37,354
1973	6	56,469	1980	13	49,100	1987	20	41,900	1994	27	36,663

On donne : $\bar{T} = 13,5$, $\bar{C} = 48,295$, $Cov(C, T) = -56,747$, $V(T) = 65,250$, $V(C) = 50,741$.
On choisit de modéliser la chronique par :

$$C_t = at + b, \quad (2)$$

où a et b sont des paramètres à estimer.

1. Interprétez les coefficients a et b .
2. Calculez les estimations des moindres carrés \hat{a} et \hat{b} de a et b . Donnez le coefficient de détermination de la régression et l'écart-type des résidus.
3. Afin d'éliminer la tendance linéaire, on introduit la chronique $z_t = C_t - C_{t-1}$ pour $t > 1967$.
 - 3.1. Montrez que lorsque C satisfait (2), la chronique z est telle que pour tout $t > 1967$, $z_t = a$.
 - 3.2. Donnez une relation entre \bar{z} , C_{1967} et C_{1994} .
4. On désire prévoir la consommation de vin.
 - 4.1 Comment faire des prévisions en utilisant le modèle de régression estimé précédemment ?