Chapitre 3

Processus non stationnaires

Révision février 2010.
On examine ici la non stationnarité associée à la présence de racines unité dans le polynôme d'autorégression.

3.1 Le problème de la racine unité

Considérons un processus autorégressif d'ordre 1.

$$Y_t = \phi_1 Y_{t-1} + Z_t$$

où donc Z_t est un bruit blanc $(0, \sigma_Z^2))$. On a vu que si $-1 < \phi_1 < 1$ alors Y_t est stationnaire.
Pour tester : $\phi_1 = 0$ on peut estimer le coefficient par MCO (Moindres Carrés Ordinaires) et faire un test de Student classique.
Supposons maintenant que $\phi_1 = 1$. On obtient :

$$Y_t = y_0 + Z_1 + Z_2 + \cdots + Z_t \qquad t \geq 0 \tag{3.1}$$

Le processus Y_t est alors appelée une marche aléatoire, voir la figure (3.1). (Il faut nécessairement un point de départ.)
L'examen de certaines séries temporelles suggèrent des extensions à ce comportement. Ainsi, les rendements composés des actions, définis par $R_t = Y_t - Y_{t-1}$ où Y_t est le log du prix de l'action, ressemblent souvent à un bruit blanc avec une moyenne légèrement positive. Ce qui suggère qu'un modèle pour Y_t est :

$$Y_t = \delta + Y_{t-1} + Z_t.$$

On dit que Y_t est une marche aléatoire avec dérive (ou drift) δ. En exprimant Y_{t-1} en fonction de Y_{t-2},... on obtient :

$$Y_t = \delta t + y_0 + Z_1 + Z_2 + \cdots + Z_t.$$

Le graphique de Y_t en fonction du temps est donc celui d'une droite à laquelle est superposée une marche aléatoire.

Propriétés élémentaires

Propriétés élémentaires d'une marche aléatoire.

$\mathsf{E}(Y_t) = y_0$, $\mathsf{var}(Y_t) = t\sigma_Z^2$, $\quad \mathsf{var}(Y_{t+1}) = (t+1)\sigma_Z^2$, $\quad \mathsf{cov}(Y_t, Y_{t+1}) = t\sigma_Z^2$,
$\lim_{t\to\infty} \mathsf{var}(Y_t) = \infty$
et donc $\rho(Y_t, Y_{t+1}) = \sqrt{\frac{t}{t+1}} < 1.$

Propriétés élémentaires d'une marche aléatoire avec dérive.

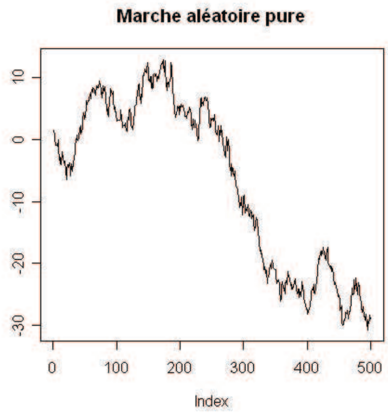

FIGURE 3.1 – Marche aléatoire

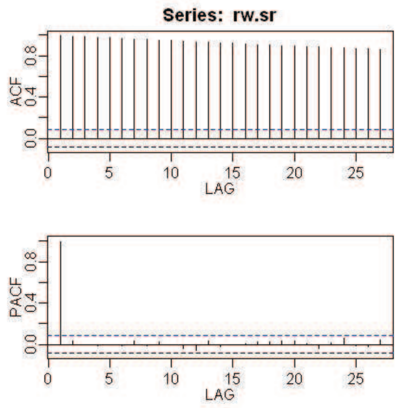

FIGURE 3.2 – ACF et PACF de la marche aléatoire ci-contre

$\mathsf{E}(Y_t) = y_0 + \delta t$, et les autres propriétés de la marche aléatoire sans dérive restent valables.

Notes.

– Une marche aléatoire (dont la variance au moins n'est pas constante) est évidemment une série non stationnaire. Avec notre vocabulaire sur les ARIMA, "marche aléatoire" est équivalent à "ARIMA(0,1,0)".
– Une série non stationnaire dont une différence première, deuxième, $\cdots$ est stationnaire est dite avoir un *trend stochastique*. Les ARIMA(p,d,q) avec $d > 0$ sont des séries ayant un trend stochastique, on dit que la série est intégrée d'ordre d, $(I(d))$, et la notation indique que la différence d[e]de la série est un ARMAp, q). Les SARIMA avec $D_s > 0$ sont des séries à saisonnalité stochastique.
  Si dans une série mensuel on est amené à modéliser la saisonnalité par des indicatrices de mois (indic. de janvier = 1 pour tous les mois de janvier, = 0 sinon, indic. de février = 1 pour tous les mois de février, = 0 sinon $\cdots$) on choisit de modéliser par une saisonnalité déterministe.

3.1.1 Illustration

Pour illustrer ces modèles, on a simulé une marche aléatoire (Fig. 3.1) avec $Y_0 = 0$, et dessiné ses ACF et PACF (Fig. 3.2)

$$Y_t = Y_{t-1} + Z_t$$

c'est-à-dire

$$Y_t = Z_1 + Z_2 + \cdots + Z_t \tag{3.2}$$

puis une marche aléatoire avec dérive (Fig. 3.3) et dessiné ses ACF et PACF (Fig. 3.4)

$$Y_t = \alpha + Y_{t-1} + Z_t$$

c'est-à-dire

$$Y_t = \alpha t + Z_1 + Z_2 + \cdots + Z_t \tag{3.3}$$

On voit sur ces graphiques que : l'ACF ne décroit jamais vers 0, une marche aléatoire sans dérive peut prendre n'importe quelle valeur et donc ne peut pas être stationnaire, une marche aléatoire avec dérive a une tendance.

La question qui nous occupe maintenant est :

comment tester dans un autorégressif d'ordre 1, que $\phi_1 = 1$ ou, comment tester que la racine du polynôme d'autorégression $1 - \phi_1 B = 0$ est 1 ?

Plus généralement, comment tester que le polynôme d'autorégression d'un autorégressif d'ordre p a une racine unité ? (Une autorégression d'ordre p n'est appelée AR(p) que si elle est stationnaire.)

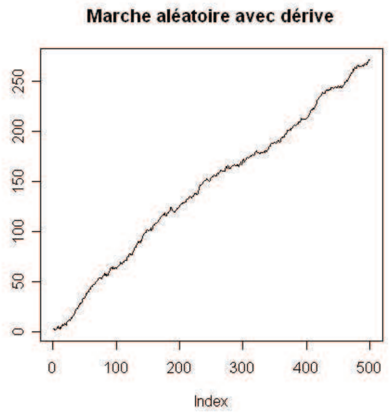

FIGURE 3.3 – Marche aléatoire avec dérive

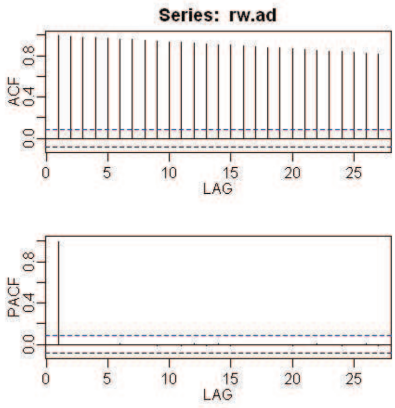

FIGURE 3.4 – ACF et PACF de la marche aléatoire ci-contre

3.1.2 Trend stochastique, trend déterministe

Supposons un processus AR(1) avec un trend déterministe :

$$Y_t - a - bt = \phi_1(Y_{t-1} - a - b(t-1)) + Z_t \qquad |\phi_1| < 1. \tag{3.4}$$

Posant $\mu_t = a + bt$, on voit qu'on peut l'écrire :

$$Y_t = \mu_t + \frac{Z_t}{1 - \phi_1 B} \tag{3.5}$$

Y_t est donc bien la somme d'une tendance déterministe μ_t et d'un AR(1) centré [1]. Que se passe-t-il quand $\phi_1 = 1$?
Développons d'abord (3.4). On obtient :

$$Y_t = \beta_1 + \beta_2 t + \phi_1 Y_{t-1} + Z_t \tag{3.6}$$

où : $\beta_1 = a(1 - \phi_1) + b\phi_1, \;\; \beta_2 = b(1 - \phi_1)$.

Donc si $\phi_1 = 1$ alors $\beta_2 = 0$, $\beta_1 = b$ et

$$Y_t = b + Y_{t-1} + Z_t,$$

est une marche aléatoire avec dérive. En résumé, si Y_t a un trend déterministe μ_t, alors si $\phi_1 = 1$ nécessairement, $\beta_2 = 0$ (donc on ne peut avoir à la fois un trend stochastique et un trend déterministe dans le modèle (3.6)).

La démarche suivante, qui arrive au même résultat, est peut-être plus claire.
Enlevant Y_{t-1} des deux côtés de (3.4) et posant $\pi = \phi_1 - 1$, on obtient :

$$\Delta Y_t = b + \pi(Y_{t-1} - b(t-1) - a) + Z_t$$

de la forme, si $\pi < 0$:

$$\Delta Y_t = \beta_0 + \beta_1 t + \pi Y_{t-1} + Z_t \tag{3.7}$$

ou, si $\pi = 0$:

$$\Delta Y_t = b + Z_t$$

1. Ce pourrait être le modèle du Lac Huron.

et on voit directement ici que quand $\pi = 0$, Y_t a un trend stochastique avec dérive b, et quand $\pi < 0$, Y_t oscille autour de $a + bt$.

Cas d'une autorégression d'ordre supérieur à 1. Il est peu probable qu'un modèle autorégressif d'ordre 1 suffise à décrire la dynamique. Considérons donc, en étendant ainsi (3.6), un autorégressif d'ordre p, pas nécessairement stationnaire, avec tendance :

$$Y_t = \beta_1 + \beta_2\, t + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + Z_t. \tag{3.8}$$

Cette série n'est pas stationnaire si 1 est racine de $1 - \phi_1 \mathrm{B} - \cdots - \phi_p \mathrm{B}^p = 0$ c'est-à-dire si $\rho = \phi_1 + \cdots + \phi_p$ prend la valeur 1. Il serait utile de voir apparaître le coefficient ρ dans (3.8). C'est ce que permet l'écriture ECM (*Error Correction Model*) de (3.8) :

$$Y_t = \beta_1 + \beta_2\, t + \rho Y_{t-1} + \zeta_1 \triangle Y_{t-1} + \cdots + \zeta_{p-1} \triangle Y_{t-p+1} + Z_t, \tag{3.9}$$

ou, retranchant des deux côtés Y_{t-1}, et posant $\pi = \phi_1 + \cdots + \phi_p - 1 = \rho - 1$:

$$\triangle Y_t = \beta_1 + \beta_2\, t + \pi Y_{t-1} + \zeta_1 \triangle Y_{t-1} + \cdots + \zeta_{p-1} \triangle Y_{t-p+1} + Z_t \tag{3.10}$$

où les ζ s'expriment en fonction des ϕ. Cette équation est la version pour un ordre p quelconque, de (3.7) pour $p = 1$.
(On obtient la représentation 3.10 en écrivant $Y_{t-p} = Y_{t-p+1} - \triangle Y_{t-p+1}, Y_{t-p-1} = Y_{t-p+2} - \triangle Y_{t-p}, \cdots,$ et en reportant dans (3.8).)
Examinons (3.10). Le côté gauche est la variation de la série de $t-1$ à t, l'erreur en t ; elle est exprimée du côté droit comme une combinaison linéaire d'un trend, de la valeur de la série à la date précédente, des variations antérieures et de l'innovation Z_t ; on comprend la terminologie "modèle de correction d'erreur". Si la série a une racine unité, alors le côté gauche de (3.10) qui est la série différenciée, est stationnaire, donc le côté droit doit l'être aussi, au terme déterministe $\beta_1 + \beta_2\ t$ près, ce qui exige $\pi = 0$ pour annuler le terme non stationnaire. Si la série n'a pas de racine unité alors les deux côtés sont stationnaires et la série peut avoir un trend linéaire.

3.2 Test de non stationnarité due à la présence d'une racine unité. Introduction et pratique

La théorie et la pratique des tests de racine unité sont difficiles. Ces tests considèrent l'hypothèse nulle : "1 est racine du polynôme d'autorégression". (Si 1 est racine double, on doit différencier deux fois la série pour se ramener à la stationnarité). Les tests de racine unité étant délicats à mettre en œuvre, il est indispensable de *conduire ces tests parallèlement à un examen du chronogramme*. Ici nous présentons seulement la technique du test ADF (*Augmented Dickey-Fuller*) et la marche à suivre pour mettre en œuvre un tel test.
Principe.
Cas I : le chronogramme de la série ne montre pas de tendance donc elle ne peut pas être une marche aléatoire avec dérive (Fig. 3.3), ni une série avec un trend déterministe qui montrerait une évolution autour d'une droite. La régression à considérer est donc

$$Y_t = \beta_1 + \phi Y_{t-1} + Z_t,$$

où, si l'on pense qu'il faut régresser à un ordre supérieur à 1 :

$$Y_t = \beta_1 + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + Z_t. \tag{3.11}$$

La représentation ECM, parallèle à (3.10) est

$$\triangle Y_t = \beta_1 + \pi Y_{t-1} + \zeta_1 \triangle Y_{t-1} + \cdots + \zeta_{p-1} \triangle Y_{t-p+1} + Z_t \tag{3.12}$$

la constante β_1 est là pour capter une moyenne non nulle ($= \beta_1/\pi$ dans le cas où $\pi < 0$ mais si $\pi = 0$, β_1 doit être nulle puisque le graphique montre qu'il n'y a pas de dérive. Si l'on sait apriori que la série est de moyenne nulle (dans le cas où elle serait stationnaire) ou sans dérive (dans le cas où elle aurait une racine unité), ce modèle devient :

$$\triangle Y_t = \pi Y_{t-1} + \zeta_1 \triangle Y_{t-1} + \cdots + \zeta_{p-1} \triangle Y_{t-p+1} + Z_t \tag{3.13}$$

On teste $H_0 : \pi = 0$, ou $H_0 : \rho = 1$, c'est-à-dire la série est I(1) sans dérive contre $H_1 : \pi < 0$, ou $H_1 : \rho < 1$ et la série est I(0) avec une moyenne éventuellement non nulle. Cette situation est souvent celles des séries financières sans tendance, telles que taux d'intérêt, taux de change.
Mise en pratique dans SAS.
Ce test est une option de l'étape `identify` dans la `proc arima`. D'abord, comme le retard p n'est pas connu, on doit choisir une série de valeurs de p et SAS calcule les statistiques pour chacune de ces valeurs.
Deux sous-cas sont envisagés dans la mise en œuvre par SAS.
A : Si on sait apriori (examen du chronogramme,...) que l'éventuelle moyenne de la série est nulle ($\beta_1 = 0$ dans 3.11), qui donne le modèle (3.13), on est dans le cas que SAS appelle `Zero Mean` et ce sont les seules sorties de ce cas qu'il faut considérer. On veut tester $H_0 : \pi = 0$. La statistique de test est `Tau`. C'est la statistique de Student pour tester $\pi = 0$ mais qui *n'est pas* distribuée suivant un t de Student sous l'hypothèse nulle. La région critique est de la forme $-\infty, t$, SAS donne la p-value. Si l'hypothèse nulle est rejetée on conclut que la série est stationnaire de moyenne nulle.
B : Si l'on n'a pas d'information sur β_1, on est dans le cas que SAS note `Single Mean` et ce sont les seules sorties de ce cas qu'il faut alors considérer. L'hypothèse à tester est $H_0 : (\beta_1, \pi) = (0, 0)$, car s'il n'y a pas de tendance sur le chronogramme, il ne peut y avoir de dérive en cas de non stationnarité = d'acceptation de H_0.. La statistique de test est `F`. C'est la statistique de Fisher $(\beta_1, \pi) = (0, 0)$ dans (3.12) dans la méhode des MCO mais qui *ne suit pas* une loi de Fisher sous H_0. On rejette l'hypothèse nulle pour de grandes valeurs de la statistique. Si l'hypothèse nulle est rejetée, la situation réelle pourrait être : (1) $(\beta_1 \neq 0, \pi = 0)$ ce qui indique une série non stationnaire avec dérive, c'est impossible puisque le chronogramme ne montre pas de tendance, ou (2) $\pi \neq 0$ c'est-à-dire série stationnaire de moyenne constante nulle ou non seul cas possible.
(On peut également tester $(\pi = 0)$. La statistique de test est `Tau`. C'est la statistique de Student pour tester $\pi = 0$ mais qui *n'est pas* distribuée suivant un t de Student sous l'hypothèse nulle. La région critique est de la forme $-\infty, t$, SAS donne la p-value. Si ce qu'on obtient n'est pas cohérent avec ce qui précède, on peut s'interroger sur l'aptitude du test à traiter la série considérée.)
Il est donc clair que si la série ne montre pas de tendance, cas qui nous intéresse en ce moment, la sortie `Trend` de SAS ne doit pas être considérée. D'autre part, sans connaissance apriori de la série qui permettrait d'affirmer qu'elle est de moyenne nulle, le cas `Zero Mean` ne doit pas non plus être considéré.

Cas II : le chronogramme de la série montre une tendance ; celle-ci apparaît si la série est une marche aléatoire avec dérive ou si elle est stationnaire à un trend déterministe près. Le test devra décider entre ces deux hypothèses. La régression considérée est

$$Y_t = \beta_1 + \beta_2\, t + \phi Y_{t-1} + Z_t, \tag{3.14}$$

et si l'autorégression est d'ordre supérieur à 1, l'écriture ECM (3.9) :

$$Y_t = \beta_1 + \beta_2\, t + \rho Y_{t-1} + \zeta_1 \triangle Y_{t-1} + \cdots + \zeta_{p-1} \triangle Y_{t-p+1} + Z_t,$$

ou (3.10)

$$\triangle Y_t = \beta_1 + \beta_2\, t + \pi Y_{t-1} + \zeta_1 \triangle Y_{t-1} + \cdots + \zeta_{p-1} \triangle Y_{t-p+1} + Z_t$$

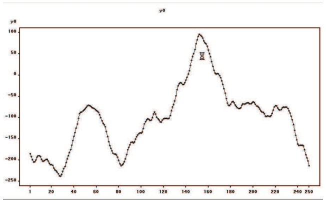

FIGURE 3.5 – y0

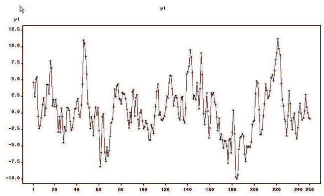

FIGURE 3.6 – y1

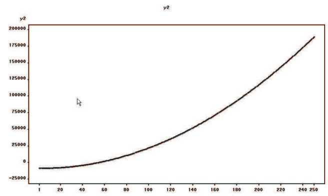

FIGURE 3.7 – y2

β_1 et β_2 sont là pour capter l'éventuelle tendance déterministe suggérée par le chronogramme, si $|\phi| < 1$. Cette approche convient aux séries ayant une tendance, comme le cours d'un titre, le niveau d'un agrégat macroéconomique.
Mise en pratique dans SAS.
Ce cas est noté `Trend` par SAS et seules les sorties correspondantes sont à considérer. D'abord, comme le retard p n'est pas connu, on doit choisir une série de valeurs de p et SAS calcule les statistiques pour chacune de ces valeurs.
L'hypothèse à tester est H_0 : $(\pi = 0,\ \beta_2 = 0)$, ou c'est-à-dire la série est I(1) avec dérive.
La statistique de test est la statistique de Fisher F, pour tester H_0, calculée comme dans la méthode MCO mais, ici encore, sous l'hypothèse nulle, cette statistique ne suit pas une loi de Fisher. Elle est notée `F` dans les sorties. On rejette l'hypothèse nulle pour de grandes valeurs de la statistique.
Si l'on rejette l'hypothèse nulle les trois situations possibles sont :
$(\pi \neq 0,\ \beta_2 = 0)$, c'est-à-dire série stationnaire sans tendance, impossible vu le chronogramme.
$(\pi = 0,\ \beta_2 \neq 0)$, c'est-à-dire série non stationnaire avec tendance, peu réaliste.
$(\pi \neq 0,\ \beta_2 \neq 0)$c'est-à-dire série stationnaire avec tendance, seul cas réaliste pour l'alternative.
(On peut également tester $(\pi = 0)$. La statistique de test est `Tau`. C'est la statistique de Student pour tester $\pi = 0$ mais qui *n'est pas* distribuée suivant un t de Student sous l'hypothèse nulle. La région critique est de la forme $-\infty, t$, SAS donne la p-value. Si l'hypothèse nulle est rejetée on conclut que la série est stationnaire. Si ce qu'on obtient n'est pas cohérent avec ce qui précède, on peut s'interroger sur l'aptitude du test à traiter la série considérée.)
Il peut évidemment arriver que 1 soit racine multiple. On détecte cette situation si après avoir différencié une première fois la série on conclut par un test de racine unité, à la non stationnarité de la série différenciée.

3.2.1 Exemples simulés

On a simulé trois séries : y0, y1 et y2. Le code figure en annexe de la section.
Effectuons le test ADF de racine unité sur ces trois séries. La syntaxe est :

```
proc arima data =a ;
i var= y0 stationarity=(adf=(1,2,3,4,5));
i var= y1 stationarity=(adf=(1,2,3,4,5));
i var= y2 stationarity=(adf=(1,2,3,4,5));
run;
quit;
```

On comprend, et l'aide en ligne de la proc arima donne les détails, que les tests ADF sont eux-mêmes un des choix de tests de racine unité (et non de stationnarité !) offerts par SAS par l'option `stationarity`. On a choisi des ordres assez élevés, jusqu'à 5, pour capter l'autorégression dont on ne connaît pas apriori l'ordre.

Examinons les résultats pour la série y0.

```
Name of Variable = y0
```

```
Mean of Working Series    -101.386
Standard Deviation        77.84165
Number of Observations         250

                              Autocorrelations

Lag  Covariance  Correlation  -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1

  0    6059.322      1.00000  |                    |********************|
  1    6001.856      0.99052  |                 .  |********************|
  2    5915.181      0.97621  |                .   |********************|
  3    5804.830      0.95800  |              .     |******************* |
  4    5676.041      0.93675  |             .      |******************* |
  5    5536.596      0.91373  |             .      |******************  |
  6    5389.025      0.88938  |            .       |******************  |
  7    5231.784      0.86343  |           .        |*****************   |
  8    5063.099      0.83559  |           .        |*****************   |
  9    4881.993      0.80570  |          .         |****************    |
 10    4688.590      0.77378  |          .         |***************     |
 11    4485.301      0.74023  |          .         |***************     |
 12    4277.746      0.70598  |         .          |**************      |
 13    4070.092      0.67171  |         .          |*************       |
 14    3864.726      0.63781  |         .          |*************       |
 15    3663.966      0.60468  |        .           |************        |
 16    3467.248      0.57222  |        .           |***********.        |
 17    3273.373      0.54022  |        .           |***********.        |
 18    3083.389      0.50887  |        .           |********** .        |
 19    2897.888      0.47825  |        .           |********** .        |
 20    2719.006      0.44873  |        .           |*********  .        |
 21    2546.459      0.42025  |        .           |********   .        |
 22    2377.180      0.39232  |       .            |********    .       |
 23    2211.725      0.36501  |       .            |*******     .       |
 24    2051.551      0.33858  |       .            |*******     .       |
```

Augmented Dickey-Fuller Unit Root Tests

Type	Retards	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	1	-4.1165	0.1625	-1.27	0.1877		
	2	-1.3159	0.4214	-0.68	0.4219		
	3	-3.4534	0.2010	-1.14	0.2324		
	4	-1.7328	0.3631	-0.72	0.4056		
	5	-3.1600	0.2214	-0.98	0.2912		
Single Mean	1	-12.2769	0.0748	-2.29	0.1774	2.62	0.4025
	2	-4.2332	0.5111	-1.31	0.6246	0.87	0.8483
	3	-10.8673	0.1061	-2.08	0.2509	2.18	0.5142
	4	-6.3124	0.3199	-1.53	0.5198	1.19	0.7663
	5	-11.6175	0.0881	-2.01	0.2814	2.06	0.5450
Trend	1	-10.4493	0.3990	-1.74	0.7296	2.71	0.6359
	2	-1.3028	0.9838	-0.38	0.9879	1.98	0.7813
	3	-8.0922	0.5712	-1.43	0.8499	2.41	0.6959
	4	-3.9374	0.8889	-0.88	0.9559	1.51	0.8765
	5	-9.9474	0.4330	-1.54	0.8141	2.07	0.7636

Lecture de ces résultats. On voit sur le graphique de y0, figure (3.5) qu'elle n'a pas de tendance, le modèle est (3.12) donc le paquet `Trend` ne la concerne pas. Sur le paquet `Single Mean` et quelque soit l'ordre de décalage, les p-values correspondant à `TAU` sont élevées donc on conclut qu'il y a une racine unité, dans (3.12) : $\pi = 0$. On examine dans le même paquet la p-value de verb"F". Elle est élevée donc on conclut sur l'expression (3.12) que $\pi = 0$ et $\beta_1 = 0$. On peut examiner le bloc Zero Mean correspondant au test de $\pi = 0$ dans le modèle (3.13). Les p-value pour `Tau` sont élevées quelque soit le décalage ; on garde l'hypothèse $\pi = 0$.

Série y1, figure (3.6).

```
                               Name of Variable = y1
    Name of Variable = y1

Mean of Working Series    0.339101
Standard Deviation        3.823035
Number of Observations         250

                              Autocorrelations

Lag  Covariance  Correlation  -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1

  0   14.615594      1.00000  |                    |********************|
```

```
   1    12.002254      0.82120  |                 .  |****************    |
   2     8.831633      0.60426  |                .   |************        |
   3     7.131012      0.48790  |                .   |**********          |
   4     6.084211      0.41628  |               .    |********            |
   5     5.334948      0.36502  |               .    |*******             |
   6     4.214475      0.28835  |               .    |******              |
   7     3.737295      0.25571  |               .    |*****               |
   8     3.551456      0.24299  |               .    |*****               |
   9     3.565609      0.24396  |               .    |*****               |
  10     3.131829      0.21428  |               .    |****.               |
  11     2.225617      0.15228  |              .     |***  .              |
  12     1.772117      0.12125  |              .     |**   .              |
  13     1.473019      0.10078  |              .     |**   .              |
  14     0.599322      0.04101  |              .     |*    .              |
  15    -0.178320      -.01220  |              .     |     .              |
  16    -0.370761      -.02537  |              .    *|     .              |
  17    -0.378226      -.02588  |              .    *|     .              |
  18    -0.783470      -.05361  |              .    *|     .              |
  19    -1.342420      -.09185  |              .   **|     .              |
  20    -1.516293      -.10374  |              .   **|     .              |
  21    -1.731373      -.11846  |              .   **|     .              |
  22    -1.892747      -.12950  |              .  ***|     .              |
  23    -2.313251      -.15827  |              .  ***|     .              |
  24    -2.806362      -.19201  |              . ****|     .              |
```

La figure (3.6) ne montre pas de tendance, la série semble de moyenne nulle. Le graphe de l'ACF montre que la série est stationnaire.

```
       Name of Variable = y1


                  Augmented Dickey-Fuller Unit Root Tests

Type          Retards        Rho   Pr < Rho       Tau   Pr < Tau          F   Pr > F

Zero Mean           1   -12.5314     0.0132     -2.36     0.0181
                    2    -4.9888     0.1239     -1.51     0.1240
                    3   -10.0685     0.0268     -2.04     0.0394
                    4    -7.4122     0.0592     -1.76     0.0739
                    5    -9.8987     0.0282     -2.01     0.0430
Single Mean         1   -17.0838     0.0220     -2.80     0.0601       3.94   0.0921
                    2    -6.7914     0.2857     -1.78     0.3914       1.58   0.6675
                    3   -14.2744     0.0452     -2.47     0.1233       3.08   0.2849
                    4   -10.4651     0.1172     -2.12     0.2371       2.26   0.4955
                    5   -14.1599     0.0465     -2.40     0.1420       2.90   0.3320
Trend               1   -20.9024     0.0531     -2.86     0.1763       4.30   0.3188
                    2    -7.1881     0.6442     -1.62     0.7845       1.59   0.8610
                    3   -17.8490     0.1005     -2.54     0.3088       3.38   0.5022
                    4   -12.2468     0.2937     -2.06     0.5638       2.36   0.7070
                    5   -17.5443     0.1069     -2.40     0.3778       3.11   0.5559
```

Lecture des résultats sur le bloc `Single Mean`. Sauf pour le retard 1, la p-value pour `F` est élevée. On comprend que la série (et le code de la simulation nous le confirme) est un autorégressif d'ordre supérieur à 1. On accepte l'hypothèse $(\pi = 0, \beta_1 = 0)$, série non stationnaire, en contradiction avec l'examen de l'ACF ! Examinons alors le bloc `Zero Mean`, ce qui est cohérent avec l'allure du graphique. Les p-values pour `Tau` sont assez faibles (sauf au retard 2) : on rejette l'hypothèse $(\pi = 0)$, on conclut que la série est stationnaire ! ! ! Que se passe-t-il ? Si on admet que la série est de moyenne nulle, comme le graphique le suggère, alors on peut considérer le paquet `Zero Mean` comme le paquet valable et et on conclut à la stationnarité de la série.

Série y2.

```
     Name of Variable = y2

Mean of Working Series    56942.47
Standard Deviation        59258.78
Number of Observations         250


                              Autocorrelations

Lag   Covariance   Correlation   -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1

  0   3511603577       1.00000  |                    |********************|
  1   3467712901       0.98750  |                 .  |********************|
```

```
  2  3423823929      0.97500 |                .   |********************|
  3  3379936890      0.96251 |              .     |******************* |
  4  3336054050      0.95001 |             .      |******************* |
  5  3292178998      0.93751 |             .      |******************* |
  6  3248318260      0.92502 |            .       |******************* |
  7  3204477037      0.91254 |           .        |******************  |
  8  3160658374      0.90006 |           .        |******************  |
  9  3116866501      0.88759 |          .         |******************  |
 10  3073103599      0.87513 |          .         |******************  |
 11  3029371219      0.86267 |         .          |*****************   |
 12  2985673140      0.85023 |         .          |*****************   |
 13  2942012801      0.83780 |        .           |*****************   |
 14  2898391308      0.82538 |        .           |*****************   |
 15  2854809980      0.81296 |        .           |****************    |
 16  2811272107      0.80057 |       .            |****************    |
 17  2767784054      0.78818 |       .            |****************    |
 18  2724350428      0.77581 |       .            |****************    |
 19  2680973557      0.76346 |      .             |***************     |
 20  2637655333      0.75113 |      .             |***************     |
 21  2594397697      0.73881 |      .             |***************     |
 22  2551202334      0.72651 |      .             |***************     |
 23  2508073288      0.71422 |     .              |**************.     |
 24  2465018168      0.70196 |     .              |**************.     |
```

Name of Variable = y2

Augmented Dickey-Fuller Unit Root Tests

Type	Retards	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	1	1.7348	0.9797	-12.04	<.0001		
	2	1.6879	0.9776	-3.82	0.0002		
	3	1.6979	0.9781	-5.25	<.0001		
	4	1.6714	0.9768	-3.18	0.0016		
	5	1.6612	0.9763	-3.54	0.0005		
Single Mean	1	1.6908	0.9962	0.35	0.9803	180.74	0.0010
	2	1.8316	0.9970	0.74	0.9929	25.97	0.0010
	3	1.7117	0.9964	0.63	0.9902	55.61	0.0010
	4	1.7254	0.9964	0.79	0.9938	27.89	0.0010
	5	1.7406	0.9965	0.67	0.9913	36.22	0.0010
Trend	1	-0.0277	0.9960	-0.99	0.9419	6.94	0.0335
	2	-0.0254	0.9960	-1.74	0.7325	24.17	0.0010
	3	-0.0220	0.9960	-1.03	0.9369	9.75	0.0010
	4	-0.0214	0.9960	-1.31	0.8836	15.65	0.0010
	5	-0.0207	0.9960	-1.05	0.9345	9.63	0.0010

Le graphique (Fig. 3.7) montre qu'elle a une tendance. On est dans le cadre du modèle (3.10) et on doit donc examiner le paquet `Trend` Par la statistique F, on rejette $\pi = 0$ et $\beta_1 = 0$ dans (3.10). Par contre les p-value pour tester $\pi = 0$ sont élevées, donc on garde $\pi = 0$. Ainsi la série est non stationnaire et avec tendance.

Annexe

Code de la simulation des séries y0, y1, y2. Il est fortement conseillé de comprendre ce code pour voir la nature des séries simulées : stationnaire ? avec racine unité ? présence d'une dérive ?

```
Options PS=5000 LS=78 NoDate PageNo=1 NoCenter
FORMCHAR='|----|+|---+=|-/\<>*';

data a;
y0m2 = 0; y0m1=0; y1m2 = 0; y1m1=0; y2m2 = 0; y2m1=0;
wm1=0; um1=0;  vm1=0;
do t =-50 to 250;
u = 2*rannor(34289);  v= 2*rannor(13489);  w = 2*rannor(134);
y0 = 1.8* y0m1 - 0.8* y0m2 + u + 0.9* um1;
y1 = .5 +0.2136752* y1m1 +0.4273504* y1m2 + v + 0.9* vm1;
y2 = 1.3*t+  1.8* y2m1 - 0.8* y2m2 + w + 0.9* wm1;
if t >0 then output;
y0m2 = y0m1;  y0m1 = y0;
 y1m2 = y1m1; y1m1 = y1;
 y2m2 = y2m1; y2m1 = y2;
um1 = u;
vm1 = v;
```

```
wm1 = w;
end;
run;

proc arima data =a ;
i var= y0 stationarity=(adf=(1,2,3,4,5));
i var= y1 stationarity=(adf=(1,2,3,4,5));
i var= y2 stationarity=(adf=(1,2,3,4,5));
run;
quit;
```

3.2.2 Séries réelles

Poivre noir

On considère la série du prix spot mensuel moyen du poivre noir d'assez bonne qualité, octobre 73, juillet 94 en $ par million de tonne.

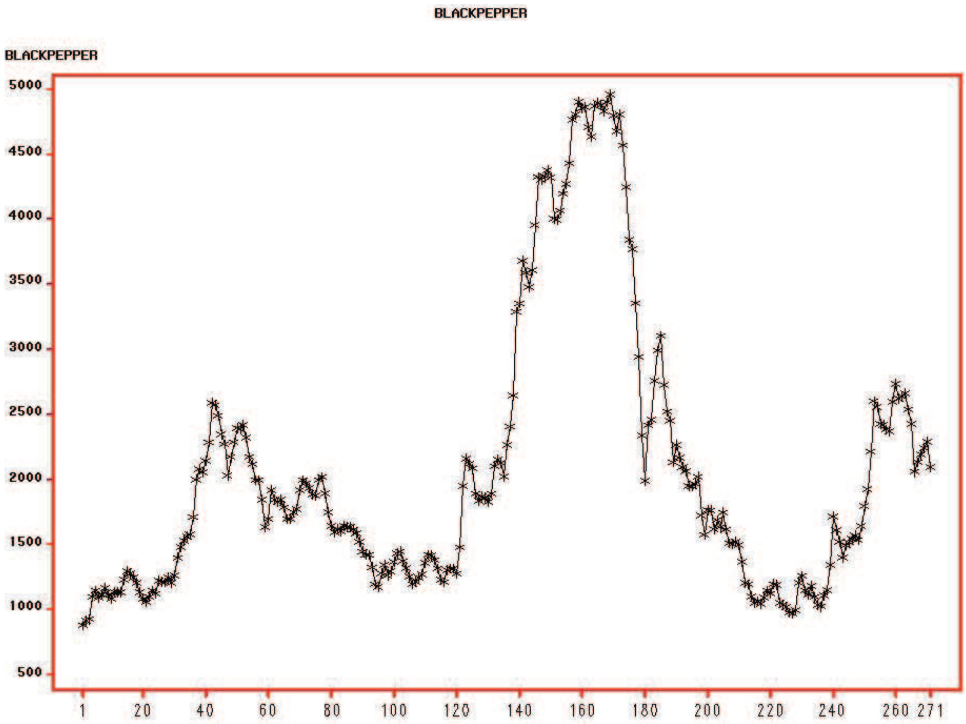

FIGURE 3.8 – Poivre noir

```
proc arima data = pepper;
i var = BLACKPEPPER(1)
stationarity=(adf=(2,3,4,5));
 run;

 The ARIMA Procedure

  Name of Variable = BLACKPEPPER

Mean of Working Series    2077.022 Standard Deviation
1050.497 Number of Observations         271


                              Autocorrelations

Lag  Covariance  Correlation  -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8
9 1

  0     1103543      1.00000  |                     |********************|
  1     1090271      0.98797  |                   . |********************|
  2     1068960      0.96866  |                 .   |******************* |
  3     1044945      0.94690  |                .    |******************* |
  4     1020629      0.92487  |               .     |******************  |
```

```
 5     996646     0.90313 |          .      |******************  |
 6     972702     0.88144 |         .       |******************  |
 7     949906     0.86078 |         .       |*****************   |
 8     925115     0.83831 |        .        |*****************   |
 9     898460     0.81416 |        .        |****************    |
10     868510     0.78702 |       .         |****************    |
11     837544     0.75896 |       .         |***************     |
12     805021     0.72949 |       .         |***************     |
13     770302     0.69803 |      .          |**************      |
14     735205     0.66622 |      .          |*************       |
15     700673     0.63493 |      .          |*************       |
16     665133     0.60272 |      .          |************        |
17     627922     0.56901 |     .           |***********.        |
18     590406     0.53501 |     .           |***********.        |
19     552108     0.50030 |     .           |********** .        |
20     511853     0.46383 |     .           |*********  .        |
21     470232     0.42611 |     .           |*********  .        |
22     430102     0.38975 |     .           |********   .        |
23     391276     0.35456 |     .           |*******    .        |
24     351535     0.31855 |     .           |******     .        |
```

Augmented Dickey-Fuller Unit Root Tests

Type	Retards	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	2	-0.8722	0.4950	-0.57	0.4694		
	3	-0.9398	0.4828	-0.61	0.4531		
	4	-0.8639	0.4965	-0.57	0.4677		
	5	-0.8428	0.5004	-0.56	0.4744		
Single Mean	2	-6.4140	0.3127	-1.86	0.3491	1.78	0.6153
	3	-6.3347	0.3186	-1.80	0.3795	1.65	0.6499
	4	-6.1355	0.3337	-1.77	0.3954	1.59	0.6635
	5	-6.3673	0.3161	-1.80	0.3780	1.66	0.6454
Trend	2	-6.2212	0.7237	-1.77	0.7156	1.76	0.8245
	3	-6.2242	0.7234	-1.74	0.7331	1.63	0.8524
	4	-5.9914	0.7422	-1.70	0.7507	1.58	0.8624
	5	-6.1733	0.7275	-1.72	0.7399	1.65	0.8475

Reprenant les étapes comme pour les séries simulées, on conclut que BLACKPEPPER a une racine unité et n'a pas de dérive.

Part de marché

On examine la part de marché d'un bien de consommation très "volatile", mesurée chaque semaine (Fig. 3.9).
La série décroit fortement, manifestement elle a une tendance (sans doute stochastique, mais n'anticipons pas !).

Augmented Dickey-Fuller Unit Root Tests

Type	Retards	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	2	-1.5059	0.3927	-1.27	0.1885		
	3	-1.4253	0.4041	-1.38	0.1545		
	4	-1.4918	0.3946	-1.43	0.1412		
	5	-1.6954	0.3669	-1.41	0.1478		
	6	-1.3896	0.4093	-1.51	0.1230		
	7	-1.4862	0.3954	-1.77	0.0726		
	8	-1.7400	0.3611	-2.52	0.0122		
Single Mean	2	-2.0190	0.7734	-0.85	0.7999	0.82	0.8609
	3	-1.4584	0.8379	-0.71	0.8390	0.95	0.8298
	4	-1.5622	0.8265	-0.75	0.8284	1.02	0.8118
	5	-2.3314	0.7351	-0.95	0.7702	1.02	0.8110
	6	-1.0750	0.8771	-0.59	0.8675	1.14	0.7798
	7	-1.0322	0.8812	-0.63	0.8587	1.60	0.6671
	8	-1.2618	0.8586	-0.94	0.7717	3.20	0.2646

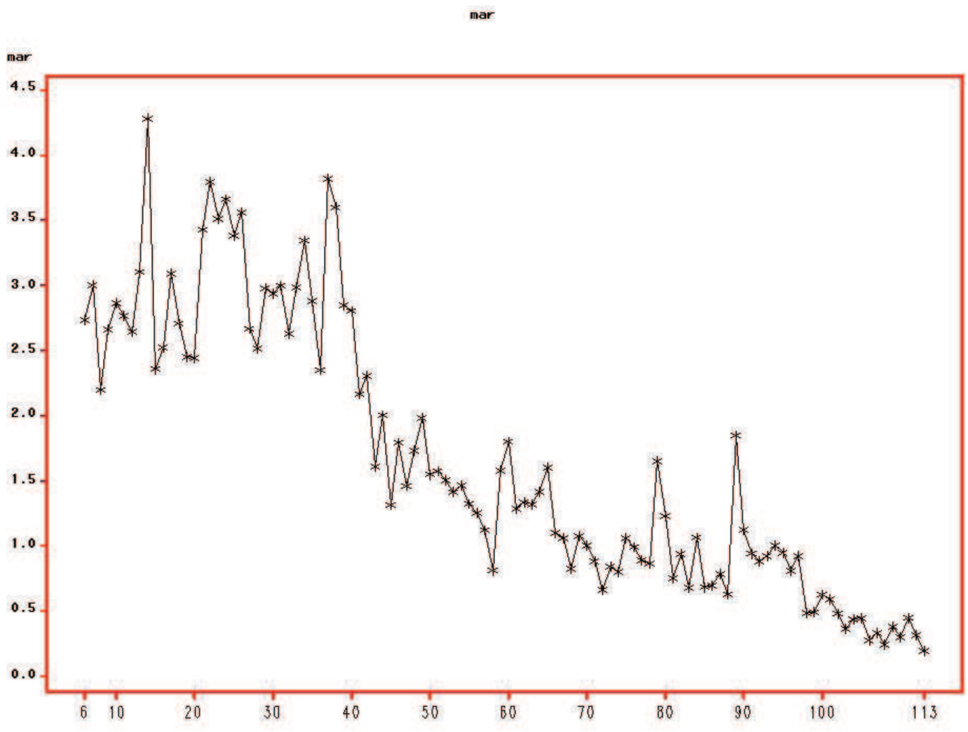

FIGURE 3.9 – Part de marché

```
Trend                2  -22.4073    0.0325    -3.28    0.0749     5.48  0.0950
                     3  -18.2600    0.0827    -2.85    0.1829     4.13  0.3511
                     4  -20.7537    0.0473    -2.85    0.1845     4.09  0.3593
                     5  -38.2874    0.0005    -3.29    0.0730     5.45  0.0970
                     6  -20.6694    0.0480    -2.64    0.2644     3.53  0.4715
                     7  -15.6699    0.1424    -2.25    0.4582     2.53  0.6746
                     8   -8.6991    0.5125    -1.70    0.7426     1.54  0.8697
```

Le test pour Trend montre une certaine instabilité, niveaux de signification empiriques faibles aux retards 2 et 5. Les p-values des deux statistiques pour $\pi = 0$ sont assez différentes. Aussi, voyons si la série n'est pas la somme d'une tendance déterministe et d'un processus stationnaire.

```
data market;
infile 'D:\Donn\ST\franses\market.txt';
input DV mar Prix Sem;
prix100 = prix/100;
t = _n_;
run;

proc autoreg data= market;
model mar = t ;
output out= resumkt r= residmar;
run;

proc arima data= resumkt;
i var= residmar stationarity=(adf=(2,3,4,5,6,7,8));
run;
quit;
```

```
                 Augmented Dickey-Fuller Unit Root Tests
```

Type	Retards	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	2	-22.2847	0.0006	-3.29	0.0012		
	3	-18.0787	0.0023	-2.84	0.0048		
	4	-20.4861	0.0011	-2.83	0.0050		
	5	-37.7709	<.0001	-3.28	0.0013		
	6	-20.0016	0.0012	-2.59	0.0099		
	7	-15.1707	0.0056	-2.22	0.0261		
	8	-8.8211	0.0373	-1.77	0.0736		
Single Mean	2	-22.2688	0.0047	-3.28	0.0186	5.45	0.0250
	3	-18.0647	0.0148	-2.84	0.0571	4.09	0.0828
	4	-20.4538	0.0076	-2.83	0.0585	4.05	0.0856
	5	-37.6213	0.0010	-3.27	0.0189	5.39	0.0270
	6	-19.9932	0.0086	-2.59	0.0982	3.42	0.1999
	7	-15.1926	0.0321	-2.22	0.2006	2.49	0.4434
	8	-8.8255	0.1686	-1.76	0.4002	1.54	0.6812
Trend	2	-22.4073	0.0325	-3.28	0.0749	5.48	0.0950
	3	-18.2600	0.0827	-2.85	0.1829	4.13	0.3511
	4	-20.7537	0.0473	-2.85	0.1845	4.09	0.3593
	5	-38.2874	0.0005	-3.29	0.0730	5.45	0.0970
	6	-20.6694	0.0480	-2.64	0.2644	3.53	0.4715
	7	-15.6699	0.1424	-2.25	0.4582	2.53	0.6746
	8	-8.6991	0.5125	-1.70	0.7426	1.54	0.8697

L'ACF du résidu de la régression de la série sur le temps décroit très rapidement et le test ci-dessus Zero Mean, conduit à considérer que les résidus sont stationnaires. Finalement la part de marché semble être une série formée d'une tendance déterministe + une erreur stationnaire.

Commerce de la Chine

On examine le revenu national réel de la Chine par secteur, base 100 en 1952.

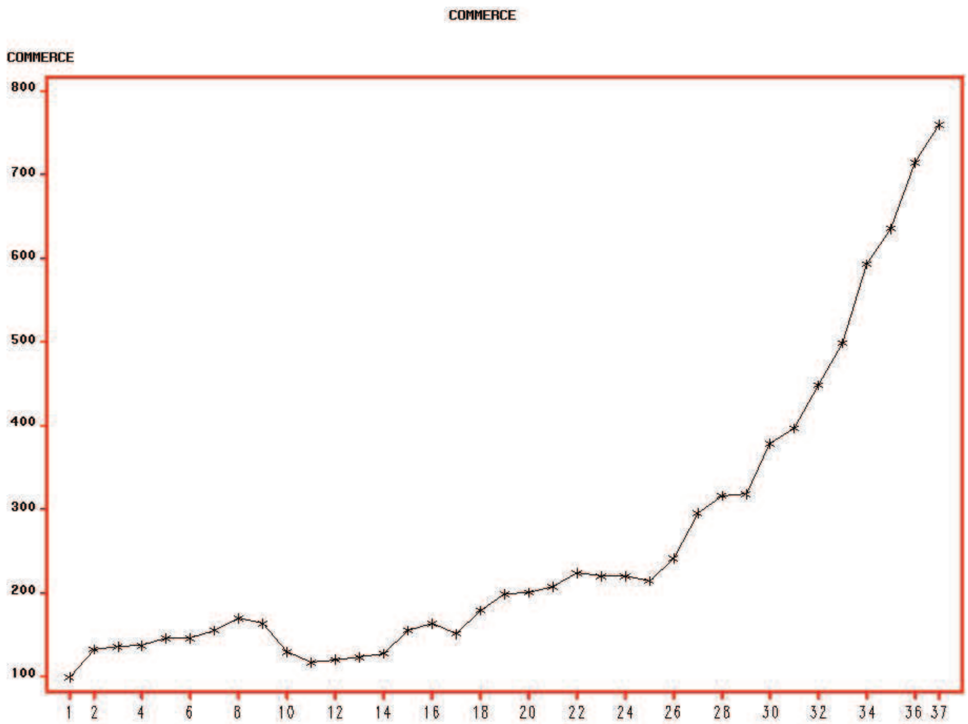

FIGURE 3.10 – Commerce de la Chine

```
proc arima data= chine;
```

```
i var= COMMERCE    stationarity=(adf=(2,3,4,5,6,7,8));
run;
quit;

Le Système SAS
1

The ARIMA Procedure

   Name of Variable = COMMERCE

Mean of Working Series    260.9865 Standard Deviation
173.8352 Number of Observations          37


                              Autocorrelations

Lag  Covariance  Correlation  -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8
9 1

  0   30218.665      1.00000  |                    |********************|
  1   25983.261      0.85984  |              .     |*****************   |
  2   21854.031      0.72320  |           .        |**************      |
  3   18219.890      0.60293  |         .          |************        |
  4   14625.703      0.48400  |       .            |**********   .      |
  5   11878.751      0.39309  |       .            |********     .      |
  6    9529.081      0.31534  |      .             |******        .     |
  7    7568.117      0.25045  |      .             |*****         .     |
  8    5614.468      0.18579  |      .             |****          .     |
  9    4255.672      0.14083  |      .             |***           .     |

"." marks two standard errors


                  Augmented Dickey-Fuller Unit Root Tests

Type         Retards       Rho  Pr < Rho      Tau  Pr < Tau       F   Pr > F

Zero Mean          2    2.9847    0.9975     2.64    0.9973
                   3    2.9742    0.9974     2.12    0.9903
                   4    3.5235    0.9990     1.53    0.9659
                   5    2.6937    0.9956     1.91    0.9844
                   6    4.3573    0.9997     1.20    0.9372
                   7    3.5772    0.9991     1.22    0.9390
                   8    6.0172    0.9999     1.06    0.9197
Single Mean        2    4.2804    0.9999     3.00    0.9999     5.56  0.0314
                   3    4.1307    0.9999     2.61    0.9999     4.21  0.0851
                   4    4.1549    0.9999     2.04    0.9998     2.48  0.4560
                   5    3.5531    0.9996     2.65    0.9999     4.29  0.0808
                   6    3.6161    0.9996     2.10    0.9998     2.43  0.4675
                   7    3.2116    0.9994     2.38    0.9999     3.10  0.3057
                   8    3.0790    0.9993     1.90    0.9997     2.05  0.5609
Trend              2    2.9106    0.9998     1.50    0.9999     5.43  0.1340
                   3    3.0551    0.9999     1.41    0.9999     4.11  0.3821
                   4    2.6393    0.9997     0.92    0.9998     3.25  0.5459
                   5    2.9678    0.9998     1.71    0.9999     4.00  0.4029
                   6    2.5565    0.9997     1.12    0.9999     3.69  0.4623
                   7    2.4848    0.9997     1.47    0.9999     4.36  0.3352
                   8    2.3521    0.9996     1.31    0.9999     3.25  0.5448
```

Exercice. Etudier cette série.

Considérations pratiques sur la non stationnarité et la différenciation

– Que se passe-t-il si, y_t obéissant par exemple à

$$y_t = \beta_0 + \beta_1 t + u_t$$

où u_t suit un modèle ARMA (stationnaire), on différencie la série ? La série différenciée est stationnaire, de moyenne β_1 mais sa structure de corrélation est celle de Δu_t et elle est généralement plus compliquée que celle de u_t. Par exemple, si u_t est un BB, alors Δu_t est un MA(1) non inversible.

- Comme il est plus expéditif de différencier une série et modéliser la série différenciée que de lui ajuster une tendance et de modéliser l'erreur, on est tenté de différencier plus souvent que nécessaire.
  ► Une structure de corrélation qui se complique après différenciation de la série est une indication qu'on a trop différencié.
- Les tests de racine unité n'ont pas une puissance élevée, on ne peut donc pas savoir avec beaucoup de certitude si on doit différencier une série ou lui ajuster une tendance linéaire.
- Sur une série courte on peut ne pas voir un trend stochastique mais en cas d'hésitation entre trend déterministe et trend stochastique, il est préférable d'ajuster un trend stochastique.
- On doit toujours examiner l'ACF et le graphique de la série quand on fait un test de racine unité.(Bien qu'on n'ait pas suivi ce chemin pour les séries y0, y1, y2.)