SERIES NON STATIONNAIRES

1 Quelques définitions

1.1 Définition d'un processus stationnaire

Un processus X_t est stationnaire au second degré si

- $E(X_t) = m \;\; \forall t$ constante pour tout t
- $E(X_t^2) < \infty$; la variance de X_t est alors finie
- $Cov(X_{t,}X_{t-h}) = \Gamma(h)$ indépendant de t

Le processus stationnaire Y_t tel que $E(Y_t) = 0$, $V(Y_t) = \sigma^2$, $Cov(Y_t, Y_{t-h}) = 0$ est un bruit blanc noté (BB)

1.2 Quelques exemples de processus stationnaires ou non

1. PROCESSUS STATIONNAIRE AUTOUR D'UNE TENDANCE TS

   le processus X_t tel que $X_t = at + b + u_t$ où u_t est un BB a pour espérance $at + b$ et pour variance la constante $V(u_t) = \sigma^2$. Son espérance étant fonction du temps, il n'est pas stationnaire; mais le processus $(X_t - at\ )$ est stationnaire. Le processus X_t est dit stationnaire autour de sa tendance et noté **TS**

2. PROCESSUS ASYMPTOTIQUEMENT STATIONNAIRE

   le processus $Z_t = \rho Z_{t-1} + a + u_t$ avec $|\rho| < 1$ où u_t est un BB peut s'écrire en fonction de Z_0 en t=0

$$\begin{aligned} Z_t &= Z_0\rho^t + a\sum_{i=0}^{t-1}\rho^i + \sum_{i=0}^{t-1}\rho^i u_{t-i} \\ E(Z_t) &= Z_0\rho^t + a\sum_{i=0}^{t-1}\rho^i = Z_0\rho^t + a\frac{1-\rho^t}{1-\rho} \end{aligned}$$

   cette espérance dépend de t mais si $t \to \infty \;\; E(Z_t) \to \frac{a}{1-\rho}$

$$V(Z_t) = E\left(\sum_{i=0}^{t-1}\rho^i u_{t-i}\right)^2 = \sigma^2\sum_{i=0}^{t-1}\rho^{2i} = \sigma^2\frac{1-\rho^{2t}}{1-\rho^2} \qquad \to \quad \sigma^2\frac{1}{1-\rho^2}$$

$$\begin{aligned} Cov(Z_t, Z_{t-h}) &= \mathbb{E}\left[\left(\sum_{i=0}^{t-1}\rho^i u_{t-i}\right)\left(\sum_{i=0}^{t-1}\rho^i u_{t-i}\right)\right] = \left(\sum_{j=0}^{t-h}\rho^{h+j}\rho^j\sigma^2\right) \\ &= \sigma^2\rho^h\sum_{j=0}^{t-1}\rho^{2j} = \sigma^2\rho^h\frac{1-\rho^{2t}}{1-\rho^2} \quad \rightarrow \quad \sigma^2\frac{\rho^h}{1-\rho^2} \end{aligned}$$

le processus Z_t est donc asymptotiquement stationnaire.

3. PROCESSUS STATIONNAIRE EN DIFFERENCE DS

   le processus $Y_t = Y_{t-1} + a + u_t$ où u_t est un BB (cas précédent avec $\rho = 1$) s'écrit en fonction de Y_0 en $t = 0$

$$Y_t = Y_0 + at + \sum_{i=0}^{t-1} u_{t-i}$$

   $E(Y_t) = Y_0 + at$ cette espérance tend vers ∞ si $t \to \infty$

$$V(Y_t) = \mathbb{E}\left(\sum_{i=0}^{t-1} u_{t-i}\right)^2 = t\sigma^2$$

$$Cov(Y_t, Y_{t-h}) = \mathbb{E}\left(\sum_{i=0}^{t-1} u_{t-i}\sum_{j=0}^{t-h-1} u_{t-h-j}\right) = \mathbb{E}\left(\sum_{J=0}^{t-h-1} u_{t-h-j}^2\right) = (t-h)\sigma^2$$

   l'espérance et la variance tendent vers ∞ si $t \to \infty$ la série Y_t n'est donc pas stationnaire. Par contre la série $(Y_t - Y_{t-1}) = a + u_t$ est stationnaire. Le processus Y_t est dit stationnaire en différence et noté **DS**

4. MARCHE ALEATOIRE

   Si dans le processus précendent on a $a = 0$, $Y_t = Y_{t-1} + u_t$ est une marche aléatoire.

1.3 Un processus non stationnaire : la tendance

Etude du modèle $Y_t = at + b + \varepsilon_t$ en fonction des propriétés de ϵ_t

1. si l'erreur ε_t est un BB, ce processus est identique au processus étudié en 1-2-1 il est non stationnaire et TS

2. si $\varepsilon_t = \rho\varepsilon_{t-1} + u_t$ où u_t est un BB et $|\rho| < 1$, alors $E(Y_t) = at + b$ et $V(Y_t) = V(\varepsilon_t)$ or ε_t a les mêmes propriétés que le processus Z_t avec $a = 0$ dans la section 1-2-2.

   On en déduit donc que $E(\varepsilon_t) \to 0$, $V(\varepsilon_t) \to \sigma^2\frac{1}{1-\rho^2}$ et $Cov(\varepsilon_t, \varepsilon_{t-h}) \to \sigma^2\frac{\rho^h}{1-\rho^2}$

   le processus ε_t est donc asymptotiquement stationnaire

   $E(Y_t) \to at + b$ qui tend donc vers ∞ et comme Y_t et ε_t ont même variance et covariance seule la première propriété de staionnarité asymptotique n'est pas vérifiée. On en déduit que le processus $(Y_t - at)$ ayant les mêmes propriétés que ε_t est asymptotiquement stationnaire, le processus Y_t est asymptotiquement **TS**

3. si $\varepsilon_t = \varepsilon_{t-1} + u_t$ où u_t est un BB, le processus ε_t est identique au processus étudié en 1-2-3 avec $a = 0$

   On en déduit que $E(\varepsilon_t) = 0$, $V(\varepsilon_t) = t\sigma^2$ et $Cov(\varepsilon_t, \varepsilon_{t-h}) = (t-h)\sigma^2$

   le processus ε_t est non stationnaire car la variance$\rightarrow \infty$, mais il est stationnaire en différence **DS**

   Le processus Y_t est tel que $E(Y_t) = at + b$ qui tend donc vers ∞ et comme Y_t et ε_t ont même variance et covariance, aucunes des propriétés de la stationnarité ne sont vérifiées ici.

   On remarque que si $\varepsilon_t = \varepsilon_{t-1} + u_t$ alors $Y_t = Y_{t-1} + a + u_t$, c'est donc un processus non stationnaire mais stationnaire en différence (**DS)** car le processus $(Y_t - Y_{t-1}) = a + u_t$ est stationnaire.

1.4 Problème posé par les séries non stationnaires dans un modèle linéaire

Soit Y_t et X_t sont deux séries non stationnaires à variance dépendant de t

- En général toute combinaison linéaire de ces séries $Y_t - aX_t - b$ est aussi non stationnaire; ainsi l'erreur ε_t du modèle $Y_t = aX_t + b + \varepsilon_t$ n'est en général pas stationnaire , la méthode des MCO n'est donc pas applicable. Il s'agit de régression fallacieuse car on a des variances d'erreurs tendant vers l'infini.

- Dans certains cas cependant, cette combinaison sera stationnaire on parlera de variables cointégrées et les MCO seront utilisables. Voilà pourquoi il va être important de savoir si les erreurs sont bien stationnaires et pour faire les tests savoir la forme de non stationnarité des séries utilisées.

- Par exemple si les séries Y_t et X_t sont TS elles s'écrivent

$$\begin{aligned}
Y_t &= c_1 t + c_2 + u_t \\
X_t &= d_1 t + d_2 + v_t \\
\text{Le modèle } Y_t &= aX_t + b + \epsilon_t \text{ devient} \\
Y_t &= a(d_1 t + d_2 + v_t) + b + \varepsilon_t = c_1 t + c_2 + u_t
\end{aligned}$$

avec u et v BB ; on trouve a et b en fonction des c et d et l'erreur $\varepsilon_t = u_t - av_t$ combinaison de BB donc BB.
Par contre si les séries sont DS le problème est différent. Il nécessite la notion de racine unitaire.

2 Notion de racine unitaire et degré d'intégration

Dans la suite du cours on notera les écarts d'une série $\Delta Y_t = DY_t = Y_t - Y_{t-1}$
Soit la série sous forme plus général autorégressif avec trend et constante,

$$Y_t = \rho_1 Y_{t-1} + \rho_2 Y_{t-2} + ... + \rho_p Y_{t-p} + at + b + u_t$$

où u_t est toujours un BB
en utilisant l'opérateur linéaire L cette équation s'écrit

$$(1-\rho_1 L-\rho_2 L^2...-\rho_p L^p)Y_t=\Phi(L)Y_t=at+b+u_t$$

On note z_i les racines du polynôme Φ qui s'écrit alors

$$(1-\frac{1}{z_1}L)(1-\frac{1}{z_2}L).....(1-\frac{1}{z_p}L)Y_t=at+b+u_t$$

si une racine et une seule égale à 1 on dira que la série Y_t a une racine unitaire elle est dite intégrée d'ordre 1 noté I(1)
si deux racines sont égales à 1 on dira que la série a deux racines unitaires elle est intégrée d'ordre 2 noté I(2)
Les séries économiques ne sont pas intégrées d'ordre supérieur à 2

1. dans le cas $Y_t=\rho Y_{t-1}+at+b+u_t$ où u_t est un BB, le modèle s'écrit $(1-\rho L)Y_t=at+b+u_t$

   (a) si la racine $z_1>1$ soit $|\rho|<1$ la série Y_t n'a pas de racine unitaire , elle est I(0) mais non stationnaire car son espérance est $at+b$ qui tend vers ∞ avec t . Elle est comme en l'a vu plus haut asymptotiquement stationnaire autour de sa tendance **TS**

   (b) si la racine $z_1=1/\rho=1$ alors $\rho=1$ et la série a une racine unitaire elle est I(1). On remarque que l'équation s'écrit

   $$Y_t-Y_{t-1}=at+b+u_t$$

   cette transformation n'est toujours pas stationnaire , la différence est bien I(0) mais il faut enlever le trend pour avoir une série stationnaire. La série

   $$Z_t=Y_t-Y_{t-1}-at=b+u_t \text{ est stationnaire}$$

2. dans le cas $Y_t=\rho_1 Y_{t-1}+\rho_2 Y_{t-2}+at+b+u_t$ on a $(1-\rho_1 L-\rho_2 L^2)Y_t=at+b+u_t$ soit

   $$(1-\frac{1}{z_1}L)(1-\frac{1}{z_2}L)Y_t=at+b+u_t$$

   (a) si les racines z_1 et z_2 sont >1 la série Y_t n'a pas de racine unitaire et comme dans la section 1- 2-1 elle est I(0) mais non stationnaire, Y_t est asymptotiquement TS

   (b) si une seule racine z_1 est égale à 1 on peut calculer qu'alors $\rho_1+\rho_2=1$ la série Y_t est non stationnaire et I(1)

   (c) si les deux racines sont égales à 1 on a en plus $\rho_2=-1$, Y_t alors non stationnaire et I(2)
   Cela signifie que pour obtenir une série stationnaire il faut les différencier deux fois Y.
   En effet si Y_t a 2 racines unitaires , elle est I(2) ; sa différence ΔY_t a une racine unitaire elle est I(1) et $(\Delta Y_t-\Delta Y_{t-1})=\Delta^2 Y_t$ n'a plus de racine unitaire elle est I(0)

3. plus généralement dans le cas $Y_t = \rho_1 Y_{t-1} + \rho_2 Y_{t-2} + ... + \rho_p Y_{t-p} + at + b + u_t$

   (a) si aucune racine n'est égale à 1 la série Y_t est I(0) et asymptotiquement TS si $a \neq 0$ et elle est I(0) asymptotiquement stationnaire si $a = 0$

   (b) si 1 seule racine est égale à 1 la série Y_t est non stationnaire I(1) . On peut voir qu'alors

$$\sum_{i=1}^{p} \rho^i = 1$$

   (c) si 2 racines sont égales à 1 la série est I(2) et comme précédemment il faudra différencier 2 fois pour avoir une série stationnaire. En économie les séries ne sont pas d'ordre supérieur à 2

   (d) Rappel: une série qui n'a pas de racine unitaire donc I(0) peut être stationnaire (si elle n'a pas de trend) ou non (si elle a un trend).

3 Le test de DICKEY-FULLER

Le test de D.F. teste la présence d'une seule racine unitaire. Pour tester si on a deux racines unitairesdans le cas où on a déjà trouvé une, il faudra comme on le verra recommencer la recherche sur les écarts de cette série. c'est pour cette raison que dans l'hypothèse H0 on parle d'au moins une RU.

$$\begin{array}{lcl} H0 & : & \text{le modèle a au moins une racine unitaire} \\ H1 & : & \text{le modèle n'a pas de racine unitaire} \end{array}$$

3.1 Test de base de D. F.

Il correspond à des séries que l'on peut présenter sous la forme suivante :

$$Y_t = \rho Y_{t-1} + at + b + u_t$$

avec u_t BB
en retranchant Y_{t-1}

$$Y_t - Y_{t-1} = \rho Y_{t-1} - Y_{t-1} + at + b + u_t$$

$$\begin{aligned} \Delta Y_t &= (\rho - 1) Y_{t-1} + at + b + u_t \\ \Delta Y_t &= \gamma Y_{t-1} + at + b + u_t \\ \gamma &= \rho - 1 \end{aligned}$$

le test s'écrit

$$\begin{array}{lcll} H0 & : & \rho = 1 \Longleftrightarrow \gamma = 0 & \text{au moins 1 RU} \\ H1 & : & \rho < 1 \Longleftrightarrow \gamma < 0 & \text{pas de RU} \end{array}$$

si H1 est vrai $(\rho - 1)$ suit asymptotiquement une loi normale
si H0 est vrai $(\rho - 1)$ ne suit plus une loi normale mais une loi issue des lois de Weiner. De plus cette loi change suivant que le modèle possède ou non un trend ou une constante.
Mais dans tous les cas, l'erreur du modèle doit être un BB. Dans les modèle ayant une erreur u_t présentant de l'autocorrélation cela conduit aux tests dits A.D.F.

3.2 Test de Dickey-Fuller avancé

Le plus souvent dans la pratique expliquer Y_t par un seul décalage en $t-1$ et un trend ne suffit pas à obtenir comme erreur un bruit blanc.
Supposons que r retards soient nécessaires pour que l'erreur soit un BB. Le modèle s'écrit alors

$$Y_t = \rho_1 Y_{t-1} + \rho_2 Y_{t-2} + ... + \rho_r Y_{t-r} + at + b + u_t$$

avec l'erreur u_t bruit blanc. En retranchant Y_{t-1} dans les deux membres de l'équation (vous pouvez faire la démonstration pour r=2 puis 3 et généraliser, je n'ai pas le courage de l'écrire)

$$\begin{aligned} Y_t - Y_{t-1} &= \rho_1 Y_{t-1} - Y_{t-1} + \rho_2 Y_{t-2} + ... + \rho_r Y_{t-r} + at + b + u_t \qquad (1) \\ \Delta Y_t &= \left(\sum_{i=1}^{r} \rho_i - 1 \right) Y_{t-1} + \sum_{j=1}^{r-1} \alpha_j \Delta Y_{t-j} + at + b + u_t \\ \Delta Y_t &= \gamma Y_{t-1} + \sum_{j=1}^{r-1} \alpha_j \Delta Y_{t-j} + at + b + u_t \\ \gamma &= \sum_{i=1}^{r} \rho_i - 1 \end{aligned}$$

Si l'on a une RU alors

$$\sum_{i=1}^{p} \rho^i = 1$$

les hypothèses du modèle s'écriront

$$\begin{aligned} H0 &: \sum_{i=1}^{r} \rho_i = 1 \Longleftrightarrow \gamma = 0 \quad \text{au moins 1 RU} \\ H1 &: \sum_{i=1}^{r} \rho_i < 1 \Longleftrightarrow \gamma < 0 \quad \text{pas de RU} \end{aligned}$$

et le test sera identique au test de base de D.F.

3.3 construction du test

Dans sa forme générale le modèle ADF lorsque les erreurs u_t sont bien des B.B., le test de D. F. peut donc s'appliquer.
Pour exposer ce test, un peu compliqué, on va partager sa présentation en deux. Tout d'abord le schéma théorique de recherche parmi les trois possibilités, puis dns un exemple pratique on ajoutera le calcul de décision entre H0 et H1.
D. F. ont montré que sous l'hypothèse H0 la loi suivie par $\widehat{\gamma}$ dépend de la présence ou non de trend et constante dans le modèle

$$Y_t = \rho_1 Y_{t-1} + \rho_2 Y_{t-2} + ... + \rho_r Y_{t-r} + at + b + u_t$$

ou

$$\Delta Y_t = \left(\sum_{i=1}^{r} \rho_i - 1 \right) Y_{t-1} + \sum_{j=1}^{r-1} \alpha_j \Delta Y_{t-j} + at + b + u_t$$

On procède de la façon suivante:

1. on teste

$$\begin{aligned} H0 &: \sum_{i=1}^{r} \rho_i = 1 \Longleftrightarrow \gamma = 0 \quad \text{au moins 1 RU} \\ H1 &: \sum_{i=1}^{r} \rho_i < 1 \Longleftrightarrow \gamma < 0 \quad \text{pas de RU} \\ avec\ a &\neq 0\ et\ b \neq 0 \end{aligned}$$

   soit la forme la plus complète du modèle.

   la table de D.F. donne la borne (pour avoir une borne plus précise que celles de D. F. suivant les valeurs de n voir le début des tables de MacKinnon).

   (a) si la décision est H1 la série Y n'a pas de racine unitaire alors

   $$\sum_{i=1}^{p} \rho^i < 1$$

   Pour tester si le modèle Y a un trend at , il suffit de tester si $a = 0$ à l'aide d'un test de Student classique car nous ne sommes pas dans l'hypothèse H0, les coefficients suivent donc asymptotiquement une loi Normale; on teste de la même façon la présence de terme constant . En théorie le test est terminé.

   (b) si la décision est H0 sous l'hypothèse $a \neq 0$, $\sum_{i=1}^{p} \rho^i = 1$ et la série Y a au moins une racine unitaire. Comme les tests sous H0 sont différents en présence ou non de trend , il faut tester la présence du trend sachant que l'on a une RU.

   Ce test s'effectue comme un test de Fisher de contraintes linéaires Φ_3 mais comme nous sommes dans le cadre d'une racine unitaire la loi suivie n'est pas un Fisher mais elle est tabulée par D. et F.

$$\begin{aligned} H01 &: \sum_{i=1}^{r} \rho_i = 1 \Longleftrightarrow \gamma = 0 \quad \text{et} \quad a = 0 \\ H11 &: \text{comme} \sum_{i=1}^{r} \rho_i = 1 \text{ alors } a \neq 0 \end{aligned}$$

- si la décision est H11 alors le modèle a un trend et le résultat du test précédent est bon, le modèle a une racine unitaire au moins et un trend. Le test est terminé.
- si la décision est H01 le modèle n'a pas de trend et on doit reprendre le test dans un modèle sans trend. On passe donc à la partie suivante.

2. on teste toujours

$$\begin{aligned} H0 &: \sum_{i=1}^{r} \rho_i = 1 \Longleftrightarrow \gamma = 0 \quad \text{au moins 1 RU} \\ H1 &: \sum_{i=1}^{r} \rho_i < 1 \Longleftrightarrow \gamma < 0 \quad \text{pas de RU} \\ \textit{avec maintenant } a &= 0 \ \textit{et} \ b \neq 0 \end{aligned}$$

avec maintenant $a = 0$ et $b \neq 0$ le modèle s'écrit

$$\begin{aligned} Y_t &= \rho_1 Y_{t-1} + \rho_2 Y_{t-2} + ... + \rho_r Y_{t-r} + b + u_t \\ \Delta Y_t &= \left(\sum_{i=1}^{r} \rho_i - 1 \right) Y_{t-1} + \sum_{j=1}^{r-1} \alpha_j \Delta Y_{t-j} + b + u_t \end{aligned}$$

la borne est lue sur les tables de D. F. pour un modèle sans trend mais avec constante

(a) si la décision est H1 la série Y n'a pas de racine unitaire alors

$$\sum_{i=1}^{p} \rho^i < 1$$

Pour tester si le modèle Y a un terme constant b , il suffit de tester si $b = 0$ à l'aide d'un test de Student classique car nous ne sommes pas dans l'hypothèse H0. En théorie le test est terminé.

(b) si la décision est H0 sous l'hypothèse $b \neq 0$, $\sum_{i=1}^{p} \rho^i = 1$ et la série Y a au moins une racine unitaire. Comme les tests sous H0 sont différents en présence ou non de terme constant , il faut tester la présence du terme constant sachant que l'on a une RU.

Ce test s'effectue comme un test de Fisher de contraintes linéaires Φ_1 mais comme nous sommes dans le cadre d'une racine unitaire la loi suivie n'est pas un Fisher mais elle est tabulée par D. et F.

$$\begin{aligned} H01 &: \sum_{i=1}^{r} \rho_i = 1 \Longleftrightarrow \gamma = 0 \quad \text{et} \quad b = 0 \\ H11 &: \quad \text{comme} \sum_{i=1}^{r} \rho_i = 1 \ \text{ alors } b \neq 0 \end{aligned}$$

- si la décision est H11 alors le modèle a un terme constant et le résultat du test précédent est bon , le modèle a une racine unitaire au moins et un terme constant. Le test est terminé.
- si la décision est H01 le modèle n'a pas de terme constant et on doit reprendre le test dans un modèle sans constante.

3. on teste toujours

$$\begin{aligned} H0 &: \sum_{i=1}^{r} \rho_i = 1 \Longleftrightarrow \gamma = 0 \quad \text{au moins 1 RU} \\ H1 &: \sum_{i=1}^{r} \rho_i < 1 \Longleftrightarrow \gamma < 0 \quad \text{pas de RU} \\ avec\ a &= 0\ et\ b = 0 \end{aligned}$$

avec maintenant $a = 0$ et $b = 0$ le modèle s'écrit

$$\begin{aligned} Y_t &= \rho_1 Y_{t-1} + \rho_2 Y_{t-2} + ... + \rho_r Y_{t-r} + u_t \\ \Delta Y_t &= \left(\sum_{i=1}^{r} \rho_i - 1 \right) Y_{t-1} + \sum_{j=1}^{r-1} \alpha_j \Delta Y_{t-j} + u_t \end{aligned}$$

(a) si la décision est H1 la série Y n'a pas de racine unitaire et le test est terminé

$$\sum_{i=1}^{p} \rho^i < 1$$

(b) si la décision est H0 sous l'hypothèse $\sum_{i=1}^{p} \rho^i = 1$ et la série Y a au moins une racine unitaire dans un modèle sans trend ni constante.

4. CONCLUSION :

Si la conclusion finale est H1 Y n'a donc pas de racine unitaire , la série est donc I(0) elle sera stationnaire si elle ne présente pas de trend , et sera stationnaire en tendance (T. S.) si elle a un trend.

Si la conclusion finale est H0 la série a une racine unitaire au moins. Pour savoir si elle en a plusieurs, il faut refaire l'étude à l'identique sur la série DY des écarts de Y .

Si de nouveau la série DY a une racine unitaire, la série Y aura ainsi au moins deux racines unitaires et on étudiera la série DDY des écarts d'écart de Y , elle ne doit pas avoir de racine unitaire car les séries économiques sont au maximum I(2). La série Y sera donc I(2) il faudra la différencier deux fois pour la rendre stationnaire.

Si la série DY n'a pas de racine unitaire, la série Y n'aura qu'une racine unitaire donc I(1) .

5. REMARQUES IMPORTANTES:

   On remarque parfois que la série Y possédant ou non un coefficient du trend significatif , la série DY possède également un trend qui aurait du disparaître en passant en écart. Cela provient du fait que la série de base Y est en t^2 ou t^3,il suffit alors de complexifier le modèle de base en ajoutant un terme en t^2 ou t^3 .

   Un problème beaucoup plus important se pose lorsque la série Y présente des points de rupture (la série fonction de a_1t+b_1 sur une période et de a_2t+b_2 ensuite) . Le test de D. F. n'est plus fiable dans ce cas. C'est pour cela parfois, que vous trouvez une série Y stationnaire et son écart DY non stationnaire (ce qui est bien sur impossible en théorie). Voila pourquoi il est souvent bon de bien vérifier que si vous trouvez une série sans racine unitaire son écart est également sans racine unitaire.

   Il existe des tests spécifiques qui testent à la fois la présence de rupture (break) et de racine unitaire. Il existe plusieurs tests avec recherche de Break, voyez celui proposé par votre logiciel, ils ont tous la même présentation.

6. COMMENT CHOISIR EN PRATIQUE LA VALEUR DE r ?

   La méthode classique pour choisir r de telle sorte que l'erreur du modèle soit un BB est de minimiser l'AIC . On ajoute des retards sur Y jusqu'à obtenir le minimum du critère d'Akaiké. Pour déterminer le nombre maximum de retards pris pour la comparaison, on utilise le critère de SWERTZ qui propose de prendre $r\max = 12(n/100)^{1/4}$ ou $4(n/100)^{1/4}$ (programme RATS DFAUTOAIC.src). Il est indispensable si on utilise cette méthode de vérifier que les erreurs ne sont plus autocorrélées.

   Une autre méthode consiste simplement à ajouter des retards DY_{t-i} jusqu'à ce que le modèle ne présente plus d'autocorrélation. (programme RATS DFUNIT5.SRC non automatisé mais qui permet de suivre l'évolution des statistiques Q et H en fonction de r).

4 Exemples de recherche de RU

La procédureRats que l'on va utiliser ici est DFAUTOAIC qui donne automatiquement un modèle où les erreurs ne sont pas autocorrélées, c'est-à-dire qu'il ajoute si besoin des retards sur les écarts de la variable étudiée jusqu'à ce que les erreurs ne soient plus autocorrélées (voir 3.3.3). Ensuite elle construit le modèle et les tests de "Fisher" pour les modèles avec trend et constante puis sans trend et enfin sans trend ni constante, à l'utilisateur de vérifier dans quel cas il se trouve en utilisant ces tests de "Fisher".

Une première série d'exemples simples va être étudiée pour illustrer tout ce qui précède. Nous utiliserons principalement la procédure dfautoaic.src. Cette première série d'exemples est issue de variables simulées pour retrouver les exemples vus dans la partie théorique.
all 1000
open data exemples_integration.rat
data(for=rats) / y y1 x x1 z w x2 x3
**** définition de la variable tendance
set tendance = t
**** appel des deux procédures que l'on va utiliser
source dfautoaic.src
source dfunit5.src

4.1 Etude de la variable x1

ATTENTION : le début de l'échantillon doit être la première variable utilisable. Par exemple si x1 va de 1 à 1000, la variable en différence première Dx1 va de 2 à 1000
@dfautoaic x1 1 1000

4.1.1 Résultat de la procédure

Rats utilise la forme

$$\Delta Y_t = \left(\sum_{i=1}^{r}\rho_i - 1\right)Y_{t-1} + \sum_{j=1}^{r-1}\alpha_j \Delta Y_{t-j} + at + b + u_t$$

La procédure donne la variable en écart DX1 en fonction de $X1_{t-1}$ d'éventuels retards sur DX1 (pour avoir l'erreur non autocorrélée) du trend et de la constante. Puis sans le trend et enfin sans la constante.

```
                   TEST UTILISANT LA PROCEDURE DFAUTOAIC.SRC

***************************************************
ETUDE DE L INTEGRATION DE LA SERIE  X1
***************************************************
********* avec tendance et constante

Linear Regression - Estimation by Least Squares
Dependent Variable dX1
Usable Observations    999      Degrees of Freedom   996
Centered R**2     0.481307      R Bar **2   0.480266
Uncentered R**2   0.482074      T x R**2     481.591
Mean of Dependent Variable      0.1092138874
Std Error of Dependent Variable 2.8410251013
Standard Error of Estimate      2.0481686212
Sum of Squared Residuals        4178.2147222
Regression F(2,996)                 462.1062
Significance Level of F           0.00000000
Log Likelihood                   -2132.24643
Durbin-Watson Statistic             1.980173

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  X1{1}                    -0.950900141  0.031280176    -30.39945  0.00000000
2.  Constant                  9.429488913  0.331976027     28.40413  0.00000000
3.  TENDANCE                  0.095224915  0.003142664     30.30070  0.00000000
```

```
 statistique Q( 63 )=       63.50848   niveau de signific.   0.4584
   calcul de phi3 avec H0 (a,0,1) :     462.10619

*****modele sans le tendance avec la constante

Linear Regression - Estimation by Least Squares
Dependent Variable dX1
Usable Observations    999      Degrees of Freedom   997
Centered R**2     0.003166      R Bar **2   0.002167
Uncentered R**2   0.004639      T x R**2       4.634
Mean of Dependent Variable      0.1092138874
Std Error of Dependent Variable 2.8410251013
Standard Error of Estimate      2.8379457980
Sum of Squared Residuals        8029.7745436
Regression F(1,997)                   3.1669
Significance Level of F           0.07544822
Log Likelihood                   -2458.55599
Durbin-Watson Statistic             2.878074

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  X1{1}                  -0.005514892  0.003098974     -1.77959  0.07544822
2.  Constant                0.439956310  0.206406234      2.13151  0.03329143

 statistique Q( 63 )=       63.50848   niveau de signific.   0.4584

 calcul de phi1 avec H0 (0,0,1) :       2.32321

 ******** sans tendance ni constante

Linear Regression - Estimation by Least Squares
Dependent Variable dX1
Usable Observations    999      Degrees of Freedom   998
Centered R**2    -0.001376      R Bar **2  -0.001376
Uncentered R**2   0.000103      T x R**2       0.103
Mean of Dependent Variable      0.1092138874
Std Error of Dependent Variable 2.8410251013
Standard Error of Estimate      2.8429792875
Sum of Squared Residuals        8066.3661665
Log Likelihood                   -2460.82704
Durbin-Watson Statistic             2.882242

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  X1{1}                  0.0004328576 0.0013504735      0.32052  0.74863913

 statistique Q( 63 )=       65.66159   niveau de signific.   0.3847
```

4.1.2 Etude du test

On constate ici qu'il n'y a pas de retards sur l'endogène DX1, la procédure a donc décidé que le modèle n'avait pas besoin de retards sur DX1 pour que les erreurs soient non autocorrélées. Le modèle n'est donc pas autorégressif donc on peut utiliser DW.

- On commence par l'étude de l'autocorrélation du modèle. Comme on l'a vu en théorie il faut toujours que les erreurs du modèle ne soient pas autocorrélées pour construire D. F. Ici on constate qu'il n'y a pas de retards sur DX1, le modèle n'est donc pas autorégressif, on regarde donc la statistique de DW=1.98 , il n'y a pas d'autocorrélation d'ordre 1. La statistique de Ljung-Box Q(63) a un niveau de significativité de .458 on décide donc la non autocorrélation d'ordre supérieur à 1. IL faut toujours bien vérifier qu'il n'y a pas d'autocorrélation des erreurs. On peut donc faire le test qui ici est le test simple de D. F. car comme on l'a vu la procédure n'a pas eu besoin de mettre des retards sur DX1 pour obtenir des erreurs non autocorrélées (voir partie 3.1). Donc dans cet exemple on trouve que le modèle s'écrit

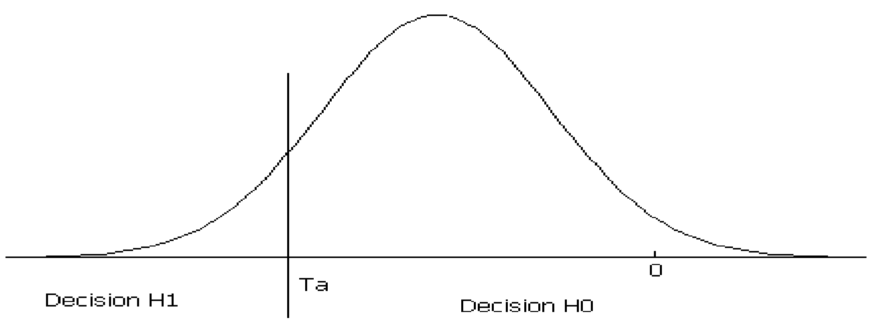
$$X1_t = \rho X1_{t-1} + at + b + u_t$$

identique à

$$\Delta X1_t = (\rho - 1) X1_{t-1} + at + b + u_t$$

Le travail est donc ici de tester si $(\rho - 1) = 0$ soit si $\rho = 1$ donc si on a une racine unitaire.

- On commence par le modèle avec tendance et constante et on regarde tout d'abord si sous ces deux hypothèses de tendance+constante il y a ou non une racine unitaire. Pour cela on fait le test de "Student", $(\widehat{\rho} - 1) = \widehat{\gamma}$ a pour t=-30.399, la table de D.F. pour n=1000 donne la borne ta=-3.41 pour le modèle avec tendance et constante, On décide donc très nettement H1 pas de R. U.



Le test s'arrête ici car on décide H1 pas de racine unitaire donc nous sommes dans le cas classique des tests de Student. On en déduit que $(\rho - 1) \neq 0$. De plus la tendance ayant un t=30.30 son coefficient est très significatif. Il en est de même pour la constante. Le modèle est donc

$$X1_t = \rho X1_{t-1} + at + b + u_t$$

On va vérifier ces résultats en construisant ce modèle

lin x1
 # constant x1{1} tendance

```
Linear Regression - Estimation by Least Squares
Dependent Variable X1
Usable Observations    999      Degrees of Freedom   996
Centered R**2     0.995011      R Bar **2   0.995001
Uncentered R**2   0.999060      T x R**2     998.061
Mean of Dependent Variable      60.081817754
Std Error of Dependent Variable 28.967522781
```

```
Standard Error of Estimate       2.048168621
Sum of Squared Residuals        4178.2147222
Regression F(2,996)               99316.0881
Significance Level of F           0.00000000
Log Likelihood                   -2132.24643
Durbin-Watson Statistic             1.980173

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  Constant                 9.4294889135 0.3319760269     28.40413  0.00000000
2.  X1{1}                    0.0490998588 0.0312801759      1.56968  0.11680725
3.  TENDANCE                 0.0952249151 0.0031426642     30.30070  0.00000000
```

On constate bien que ρ est très différent de 1 ($\widehat{\rho} = 0.049$) , de plus il n'est pas significatif, la tendance et la constante sont très significatives. En fait ce modèle est donc de la forme

$$X1_t = at + b + u_t$$

C'est une série non stationnaire et stationnaire en tendance T.S. (voir 1.2.1)

4.2 Etude de la série X2

On utilise DFAUTOAIC.SRC pour étudier cette nouvelle série.

4.2.1 Résultat de la procédure

```
                  TEST UTILISANT LA PROCEDURE DFAUTOAIC.SRC

***************************************************
ETUDE DE L INTEGRATION DE LA SERIE  X2
***************************************************
********* avec tendance et constante

Linear Regression - Estimation by Least Squares
Dependent Variable dX2
Usable Observations    999      Degrees of Freedom   996
Centered R**2     0.255486      R Bar **2   0.253991
Uncentered R**2   0.255487      T x R**2     255.231
Mean of Dependent Variable      0.0032879662
Std Error of Dependent Variable 4.6136789895
Standard Error of Estimate      3.9849173558
Sum of Squared Residuals        15816.048067
Regression F(2,996)                 170.8930
Significance Level of F           0.00000000
Log Likelihood                   -2797.15139
Durbin-Watson Statistic             2.037535

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  X2{1}                    -0.510921781  0.027636135    -18.48745  0.00000000
2.  Constant                  1.142116995  0.260100915      4.39105  0.00001249
3.  TENDANCE                 -0.000324957  0.000437528     -0.74271  0.45783184

 statistique Q( 63 )=       58.84965   niveau de signific.   0.6249
   calcul de phi3 avec H0 (a,0,1) :     170.89301

*****modele sans le tendance avec la constante

Linear Regression - Estimation by Least Squares
Dependent Variable dX2
Usable Observations    999      Degrees of Freedom   997
Centered R**2     0.255074      R Bar **2   0.254327
Uncentered R**2   0.255074      T x R**2     254.819
```

```
Mean of Dependent Variable      0.0032879662
Std Error of Dependent Variable 4.6136789895
Standard Error of Estimate      3.9840211872
Sum of Squared Residuals        15824.807546
Regression F(1,997)                 341.3879
Significance Level of F           0.00000000
Log Likelihood                   -2797.42796
Durbin-Watson Statistic             2.038103

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  X2{1}                    -0.510105415  0.027608058    -18.47669  0.00000000
2.  Constant                  0.977754157  0.136637673      7.15582  0.00000000

 statistique Q( 63 )=       58.84965   niveau de signific.   0.6249

 calcul de phi1 avec H0 (0,0,1) :     170.69430

 ******** sans tendance ni constante

Linear Regression - Estimation by Least Squares
Dependent Variable dX2
Usable Observations    999      Degrees of Freedom   998
Centered R**2     0.216815      R Bar **2   0.216815
Uncentered R**2   0.216815      T x R**2     216.598
Mean of Dependent Variable      0.0032879662
Std Error of Dependent Variable 4.6136789895
Standard Error of Estimate      4.0830023526
Sum of Squared Residuals        16637.566395
Log Likelihood                   -2822.44510
Durbin-Watson Statistic             2.096691

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  X2{1}                    -0.433850593  0.026101309    -16.62179  0.00000000

 statistique Q( 63 )=       59.01049   niveau de signific.   0.6192
```

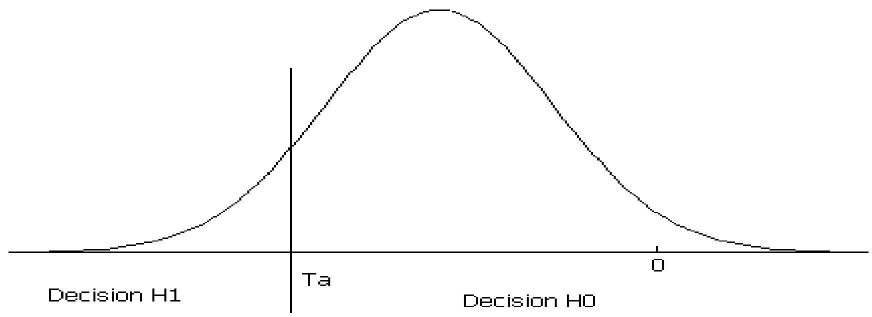
4.2.2 Etude du test

On constate ici qu'il n'y a pas de retards sur l'endogène DX2 le modèle n'est donc pas autorégressif donc on peut utiliser DW.

- On commence par l'étude de l'autocorrélation du modèle. La statistique de DW=2.037 , il n'y a pas d'autocorrélation d'ordre 1. La statistique de Ljung-Box Q(63) a un niveau de significativité de .625 on décide donc la non autocorrélation d'ordre supérieur à 1. On peut donc faire le test qui ici est le test simple de D. F. car comme on l'a vu la procédure n'a pas eu besoin de mettre des retards sur DX2 pour obtenir des erreurs non autocorrélées (voir partie 3.1). Donc dans cet exemple on trouve que le modèle s'écrit
$$X2_t = \rho X2_{t-1} + at + b + u_t$$
identique à
$$\Delta X2_t = (\rho - 1)X2_{t-1} + at + b + u_t$$
Le travail est donc ici de tester si $(\rho - 1) = 0$ soit si $\rho = 1$ donc si on a une racine unitaire.

- On commence par le modèle avec tendance et constante et on regarde tout d'abord si sous ces deux hypothèses de tendance+constante il y a ou non une racine unitaire.

Pour cela on fait le test de "Student", $(\widehat{\rho}-1)$ a pour t=-18.48, la table de D.F. pour n=1000 donne la borne ta=-3.41 pour le modèle avec tendance et constante, On décide donc très nettement H1 pas de R. U.



Le test s'arrête ici car on décide H1 pas de racine unitaire donc nous sommes dans le cas classique des tests de Student. On en déduit que $(\rho-1) \neq 0$. De plus la tendance ayant un t=-0.74 son coefficient n'est pas significatif. La constante est significative. Le modèle est donc

$$X1_t = \rho X1_{t-1} + b + u_t \qquad \text{avec } \rho \neq 1$$

On va vérifier ces résultats en construisant ce modèle

lin x2
 # constant x2{1} tendance

```
Linear Regression - Estimation by Least Squares
Dependent Variable X2
Usable Observations    999      Degrees of Freedom   996
Centered R**2     0.240438      R Bar **2   0.238913
Uncentered R**2   0.353942      T x R**2     353.588
Mean of Dependent Variable      1.9136111318
Std Error of Dependent Variable 4.5677470239
Standard Error of Estimate      3.9849173558
Sum of Squared Residuals        15816.048067
Regression F(2,996)                 157.6408
Significance Level of F           0.00000000
Log Likelihood                   -2797.15139
Durbin-Watson Statistic             2.037535

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  Constant                 1.142116995  0.260100915      4.39105  0.00001249
2.  X2{1}                    0.489078219  0.027636135     17.69706  0.00000000
3.  TENDANCE                -0.000324957  0.000437528     -0.74271  0.45783184
```

La tendance n'est pas significative par contre la constante est significative de même que le coefficient de $X2_{t-1}$.Cette série est donc asymptotiquement stationnaire (voir 1.2.2)

4.3 Etude de la série W

On utilise DFAUTOAIC.SRC pour étudier cette nouvelle série.

4.3.1 Résultat de la procédure

```
                  TEST UTILISANT LA PROCEDURE DFAUTOAIC.SRC

***************************************************
ETUDE DE L INTEGRATION DE LA SERIE  W
***************************************************
********* avec tendance et constante

Linear Regression - Estimation by Least Squares
Dependent Variable dW
Usable Observations    998      Degrees of Freedom   994
Centered R**2     0.278778      R Bar **2   0.276601
Uncentered R**2   0.278787      T x R**2     278.229
Mean of Dependent Variable      0.0042237625
Std Error of Dependent Variable 1.1899255648
Standard Error of Estimate      1.0120657243
Sum of Squared Residuals        1018.1313682
Regression F(3,994)                 128.0720
Significance Level of F           0.00000000
Log Likelihood                   -1426.06616
Durbin-Watson Statistic             1.973514

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  W{1}                     -0.003423807  0.002187099     -1.56546  0.11779441
2.  Constant                 -0.006679723  0.071622882     -0.09326  0.92571389
3.  TENDANCE                  0.000117860  0.000168215      0.70065  0.48368242
4.  dW{1}                     0.528471610  0.027035653     19.54721  0.00000000

valeur de la statistique de Durbin   h=       0.80435

 statistique Q( 63 )=       36.77833   niveau de signific.   0.9966
 stat. modifiee  Q(63-1)=   36.77833   niveau de signific.   0.9955

   calcul de phi3 avec H0 (a,0,1) :       1.48126

*****modele sans le tendance avec la constante

Linear Regression - Estimation by Least Squares
Dependent Variable dW
Usable Observations    998      Degrees of Freedom   995
Centered R**2     0.278422      R Bar **2   0.276971
Uncentered R**2   0.278431      T x R**2     277.874
Mean of Dependent Variable      0.0042237625
Std Error of Dependent Variable 1.1899255648
Standard Error of Estimate      1.0118067840
Sum of Squared Residuals        1018.6342034
Regression F(2,995)                 191.9608
Significance Level of F           0.00000000
Log Likelihood                   -1426.31255
Durbin-Watson Statistic             1.971666

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  W{1}                     -0.002274639  0.001446476     -1.57254  0.11614368
2.  Constant                  0.035710449  0.038324142      0.93180  0.35166575
3.  dW{1}                     0.526845033  0.026928905     19.56429  0.00000000


 statistique Q( 63 )=       37.16320   niveau de signific.   0.9961
 stat. modifiee Q( 63 - 1 ) 37.16320   niveau de signific.   0.9948

 calcul de phi1 avec H0 (0,0,1) :       1.23977

 ******** sans tendance ni constante

Linear Regression - Estimation by Least Squares
```

```
Dependent Variable dW
Usable Observations    998      Degrees of Freedom   996
Centered R**2     0.277792      R Bar **2   0.277067
Uncentered R**2   0.277801      T x R**2     277.245
Mean of Dependent Variable      0.0042237625
Std Error of Dependent Variable 1.1899255648
Standard Error of Estimate      1.0117398623
Sum of Squared Residuals        1019.5230788
Log Likelihood                   -1426.74779
Durbin-Watson Statistic             1.970851

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  W{1}                     -0.001534474  0.001208770     -1.26945  0.20457703
2.  dW{1}                     0.526549700  0.026925259     19.55598  0.00000000


 statistique Q( 63 )=       37.69381   niveau de signific.   0.9953
 stat. modifiee Q(63-1)=    37.69381   niveau de signific.   0.9937
```

4.3.2 Etude du test

On constate ici qu'il y a 1 retard sur l'endogène ΔW le modèle est donc autorégressif donc on remplace le test de Durbin et Watson par le test H de Durbin.

- On commence par l'étude de l'autocorrélation du modèle. La statistique de H=0.804 , il n'y a pas d'autocorrélation d'ordre 1. La statistique de Ljung-Box Q(63-1) a un niveau de significativité de .9955 on décide donc la non autocorrélation d'ordre supérieur à 1. On peut donc faire le test qui ici est le test A. D. F. car comme on l'a vu la procédure a mis un retard sur DW pour obtenir des erreurs non autocorrélées (voir partie 3.2). Donc dans cet exemple on trouve que le modèle s'écrit avec r=2

$$W_t = \rho_1 W_{t-1} + \rho_2 W_{t-2} + at + b + u_t$$

identique à

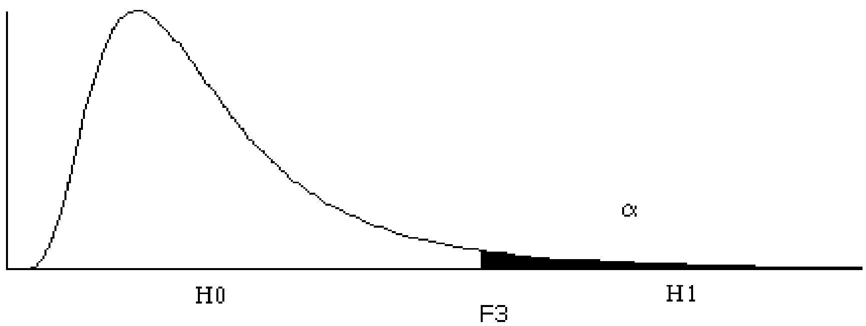
$$\Delta W_t = \left(\sum_{i=1}^{2} \rho_i - 1 \right) W_{t-1} + \alpha_1 \Delta W_{t-1} + at + b + u_t$$

Le travail est donc ici de tester si $\sum_{i=1}^{2} \rho_i - 1 = 0$ soit si $\sum_{i=1}^{2} \rho_i = 1$ donc si on a une racine unitaire.

- On commence par le modèle avec tendance et constante et on regarde tout d'abord si sous ces deux hypothèses de tendance+constante il y a ou non une racine unitaire. Pour cela on fait le test de "Student", $\sum_{i=1}^{2} \rho_i - 1$ a pour t=- 1.565 la table de D.F. pour n=1000 donne la borne ta=-3.41 pour le modèle avec tendance et constante, On décide donc très nettement H0 on a au moins une RU.

- Mais nous ne savons pas si nous avons bien fait de travailler avec la tendance, on va donc construire un test qui teste à la fois il y a une racine unitaire ET la tendance est nulle (remarquons que dans l'exemple précédent il ne faut surtout pas faire ce test car l'une des hypothèse est fausse dans la mesure où nous n'avons pas trouvé de RU.

). C'est le test Φ_3 qui se construit comme un test de Fisher (les résultats sont donnés directement) le logiciel donne $\Phi_3 = 1.48$. La table de Φ_3 donne la borne F3 =6.25, comme la valeur trouvée est très inférieure à cette borne on déduit l'hypothèse H01: la tendance est nulle.



- Comme la tendance est nulle il faut recommencer le test sans la tendance avec la constante. On trouve t=-1.572 la table de D. F. pour le modèle sans tendance donne la borne ta=-2.86.on décide donc très nettement encore H0 il y a une RU. au moins.
- Nous testons si nous avons bien travaillé en gardant la constante. Pour cela on fait le test qui teste la présence d'une RU. et de la constante c'est le teste Φ_1, le logiciel donne $\Phi_1 = 1.239$ et la table de Φ_1 donne la borne F1= 4.59, comme la valeur trouvée est très nettement inférieure à la borne on décide H01: la constante est nulle.
- Comme la constante est nulle il faut donc recommencer le test sans la constante. On trouve t=-1.269 la table de D. F. pour un modèle sans tendance ni constante donne la borne ta=-1.95 on décide donc H0 il y a une racine unitaire au moins. OUF !
- Mais ce n'est pas terminé car on a trouvé une racine mais il peut y en avoir plusieurs. Pour la savoir nous avons vu que l'étude doit être reprise sur les écarts de W soit ΔW. Si W a une seule R.U. alors ΔW n'a pas de R.U. et si W a deux R.U. alors ΔW en a une (voir la théorie ci-dessus).

4.3.3 La série ΔW a-t-elle une R.U. ?

set dw = w-w{1}
ATTENTION dw va de 2 à 1000
@dfautoaic dw 2 1000

```
                  TEST UTILISANT LA PROCEDURE DFAUTOAIC.SRC

****************************************************
ETUDE DE L INTEGRATION DE LA SERIE  DW
****************************************************
********* avec tendance et constante

Linear Regression - Estimation by Least Squares
Dependent Variable dDW
Usable Observations    998      Degrees of Freedom   995
Centered R**2     0.237680      R Bar **2   0.236147
```

```
Uncentered R**2   0.237680      T x R**2     237.205
Mean of Dependent Variable      0.0011639854
Std Error of Dependent Variable 1.1588315049
Standard Error of Estimate      1.0128032216
Sum of Squared Residuals        1020.6415138
Regression F(2,995)                 155.1128
Significance Level of F           0.00000000
Log Likelihood                   -1427.29491
Durbin-Watson Statistic             1.968823

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  DW{1}                   -0.474960741  0.026966236    -17.61316  0.00000000
2.  Constant                 0.042544691  0.064398295      0.66065  0.50899014
3.  TENDANCE                -0.000079616  0.000111361     -0.71494  0.47481682

 statistique Q( 63 )=       38.41261   niveau de significativ.   0.9939
   calcul de phi3 avec H0 (a,0,1) :     155.11282

*****modele sans le tendance avec la constante

Linear Regression - Estimation by Least Squares
Dependent Variable dDW
Usable Observations    998      Degrees of Freedom   996
Centered R**2     0.237288      R Bar **2   0.236522
Uncentered R**2   0.237289      T x R**2     236.814
Mean of Dependent Variable      0.0011639854
Std Error of Dependent Variable 1.1588315049
Standard Error of Estimate      1.0125546334
Sum of Squared Residuals        1021.1658181
Regression F(1,996)                 309.8666
Significance Level of F           0.00000000
Log Likelihood                   -1427.55118
Durbin-Watson Statistic             1.969234

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  DW{1}                   -0.474228254  0.026940152    -17.60303  0.00000000
2.  Constant                 0.002615018  0.032051963      0.08159  0.93499166

 statistique Q( 63 )=       38.41261   niveau de significativ.   0.9939

 calcul de phi1 avec H0 (0,0,1) :     154.93396

 ******** sans tendance ni constante

Linear Regression - Estimation by Least Squares
Dependent Variable dDW
Usable Observations    998      Degrees of Freedom   997
Centered R**2     0.237283      R Bar **2   0.237283
Uncentered R**2   0.237284      T x R**2     236.809
Mean of Dependent Variable      0.0011639854
Std Error of Dependent Variable 1.1588315049
Standard Error of Estimate      1.0120500871
Sum of Squared Residuals        1021.1726427
Log Likelihood                   -1427.55451
Durbin-Watson Statistic             1.969232

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  DW{1}                   -0.474222601  0.026926639    -17.61165  0.00000000

 statistique Q( 63 )=       40.39814   niveau de significativ.   0.9881
```

On constate qu'il n'y a pas d'autocorrélation des erreurs (DW=1.968 et Q(63)= 38.41 avec un niveau de significativité de ns=0.99).
Dans le modèle avec tendance et constante le "Student" t=-17.61, la borne étant ta=-3.41

on déduit très nettement qu'il n'y a pas de R. U. donc ΔW est I(0) sans tendance ni constante qui ne sont pas significatives d'après le test de Student. Donc W n'a qu'une racine unitaire, W est donc I(1) .

4.4 Etude de la série Z1

On utilise DFAUTOAIC.SRC pour étudier cette nouvelle série.

4.4.1 Résultat de la procédure

```
                  TEST UTILISANT LA PROCEDURE DFAUTOAIC.SRC
 ***************************************************
ETUDE DE L INTEGRATION DE LA SERIE  Z1
***************************************************
********* avec tendance et constante

Linear Regression - Estimation by Least Squares
Dependent Variable dZ1
Usable Observations    997      Degrees of Freedom   992
Centered R**2     0.997931      R Bar **2   0.997922
Uncentered R**2   0.998556      T x R**2     995.560
Mean of Dependent Variable      14.570982805
Std Error of Dependent Variable 22.163532871
Standard Error of Estimate       1.010253055
Sum of Squared Residuals        1012.4463455
Regression F(4,992)              119596.1884
Significance Level of F           0.00000000
Log Likelihood                   -1422.34567
Durbin-Watson Statistic             1.967881

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  Z1{1}                   -0.000023756  0.000010367     -2.29137  0.02215100
2.  Constant                -0.129847777  0.088843077     -1.46154  0.14418379
3.  TENDANCE                 0.000491710  0.000232615      2.11384  0.03477813
4.  dZ1{1}                   1.517695706  0.027123537     55.95493  0.00000000
5.  dZ1{2}                  -0.521350341  0.027203310    -19.16496  0.00000000

valeur de la statistique de Durbin   h=       0.98222

 statistique Q( 63 )=         37.7666   niveau de significativit\U{e9}   0.9951
 stat. modifiee  Q( 63 - 2 )= 37.7666   niveau de significativit\U{e9}   0.9916

     calcul de phi3 avec H0 (a,0,1) :       2.89500

*****modele sans le tendance avec la constante

Linear Regression - Estimation by Least Squares
Dependent Variable dZ1
Usable Observations    997      Degrees of Freedom   993
Centered R**2     0.997921      R Bar **2   0.997915
Uncentered R**2   0.998549      T x R**2     995.553
Mean of Dependent Variable      14.570982805
Std Error of Dependent Variable 22.163532871
Standard Error of Estimate       1.012015805
Sum of Squared Residuals        1017.0067573
Regression F(3,993)              158905.0772
Significance Level of F           0.00000000
Log Likelihood                   -1424.58605
Durbin-Watson Statistic             1.966706

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  Z1{1}                   -0.000008619  0.000007510     -1.14764  0.25139366
```

```
2.  Constant                  0.039412357  0.038557469      1.02217  0.30694856
3.  dZ1{1}                    1.521551084  0.027109365     56.12640  0.00000000
4.  dZ1{2}                   -0.522578537  0.027244560    -19.18102  0.00000000


 statistique Q( 63 )=          38.66872 niveau de signific.   0.9933
 stat. modifiee Q( 63 - 2 )=   38.66872 niveau de signific.   0.9886

     calcul de phi1 avec H0 (0,0,1) :       1.07342

 ******** sans tendance ni constante

Linear Regression - Estimation by Least Squares
Dependent Variable dZ1
Usable Observations    997      Degrees of Freedom   994
Centered R**2     0.997919      R Bar **2   0.997915
Uncentered R**2   0.998548      T x R**2     995.552
Mean of Dependent Variable      14.570982805
Std Error of Dependent Variable 22.163532871
Standard Error of Estimate       1.012038628
Sum of Squared Residuals        1018.0768523
Log Likelihood                   -1425.11030
Durbin-Watson Statistic             1.966272

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  Z1{1}                    -0.000007844  0.000007472     -1.04974  0.29409121
2.  dZ1{1}                    1.522371150  0.027098102     56.17999  0.00000000
3.  dZ1{2}                   -0.522700746  0.027244912    -19.18526  0.00000000


 statistique Q( 63 )=          39.60296 niveau de signific.   0.9908
 stat. modifiee Q( 63 - 2 )=   39.60296 niveau de signific.   0.9847
```

4.4.2 Etude du test

On constate ici qu'il y a 2 retard sur l'endogène ΔZ1 le modèle est donc autorégressif donc on remplace le test de Durbin et Watson par le test H de Durbin.

- On commence par l'étude de l'autocorrélation du modèle. La statistique de H=098 , il n'y a pas d'autocorrélation d'ordre 1. La statistique de Ljung-Box Q(63-1) a un niveau de significativité de .9916 on décide donc la non autocorrélation d'ordre supérieur à 1. On peut donc faire le test qui ici est le test A. D. F. car comme on l'a vu la procédure a mis 2 retards sur DZ1 pour obtenir des erreurs non autocorrélées (voir partie 3.2). Donc dans cet exemple on trouve que le modèle s'écrit avec r=2

  $$Z1_t = \rho_1 Z1_{t-1} + \rho_2 Z1_{t-2} + \rho_3 Z1_{t-3} + at + b + u_t$$

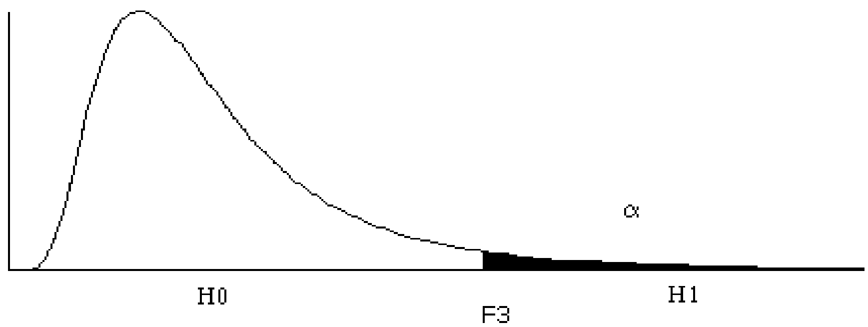
  identique à

  $$\Delta Z1_t = \left(\sum_{i=1}^{3} \rho_i - 1\right) Z1_{t-1} + \alpha_1 \Delta Z1_{t-1} + \alpha_2 \Delta Z1_{t-2} + at + b + u_t$$

  Le travail est donc ici de tester si $\sum_{i=1}^{3} \rho_i - 1 = 0$ soit si $\sum_{i=1}^{3} \rho_i = 1$ donc si on a une racine unitaire.

- On commence par le modèle avec tendance et constante et on regarde tout d'abord si sous ces deux hypothèses de tendance+constante il y a ou non une racine unitaire. Pour cela on fait le test de "Student", $\sum_{i=1}^{3} \rho_i - 1$ a pour t=- 2.29 la table de D.F. pour n=1000 donne la borne ta=-3.41 pour le modèle avec tendance et constante, On décide donc très nettement H0 on a au moins une RU.

- Mais nous ne savons pas si nous avons bien fait de travailler avec la tendance, on va donc construire un test qui teste à la fois il y a une racine unitaire ET la tendance est nulle (remarquons que dans l'exemple précédent il ne faut surtout pas faire ce test car l'une des hypothèse est fausse dans la mesure où nous n'avons pas trouvé de RU.). C'est le test Φ_3 qui se construit comme un test de Fisher (les résultats sont donnés directement) le logiciel donne $\Phi_3 = 2.89$. La table de Φ_3 donne la borne F3 =6.25, comme la valeur trouvée est très inférieure à cette borne on déduit l'hypothèse H01: la tendance est nulle.



- Comme la tendance est nulle il faut recommencer le test sans la tendance avec la constante. On trouve t=-1.147 la table de D. F. pour le modèle sans tendance donne la borne ta=-2.86.on décide donc très nettement encore H0 il y a une RU. au moins.

- Nous testons si nous avons bien travaillé en gardant la constante. Pour cela on fait le test qui teste la présence d'une RU. et de la constante c'est le teste Φ_1, le logiciel donne $\Phi_1 = 1.07$ et la table de Φ_1 donne la borne F1= 4.59, comme la valeur trouvée est très nettement inférieure à la borne on décide H01: la constante est nulle.

- Comme la constante est nulle il faut donc recommencer le test sans la constante. On trouve t=-1.049 la table de D. F. pour un modèle sans tendance ni constante donne la borne ta=-1.95 on décide donc H0 il y a une racine unitaire au moins.

- Mais ce n'est pas terminé car on a trouvé une racine mais il peut y en avoir plusieurs. Pour la savoir nous avons vu que l'étude doit être reprise sur les écarts de Z1 soit DZ1. Si Z1 a une seule R.U. alors DZ1 n'aura pas de R.U. et si Z1 a deux R.U. alors DZ1 en aura une (voir la théorie ci-dessus).

4.4.3 La série DZ1 a-t-elle une R.U. ?

```
                  TEST UTILISANT LA PROCEDURE DFAUTOAIC.SRC
 **************************************************
ETUDE DE L INTEGRATION DE LA SERIE  DZ1
**************************************************
********* avec tendance et constante

Linear Regression - Estimation by Least Squares
Dependent Variable dDZ1
Usable Observations    997      Degrees of Freedom   993
Centered R**2     0.279008      R Bar **2   0.276830
Uncentered R**2   0.279017      T x R**2     278.180
Mean of Dependent Variable      0.0041895622
Std Error of Dependent Variable 1.1905222763
Standard Error of Estimate      1.0124128551
Sum of Squared Residuals        1017.8049307
Regression F(3,993)                 128.0900
Significance Level of F           0.00000000
Log Likelihood                   -1424.97713
Durbin-Watson Statistic             1.972692

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  DZ1{1}                  -0.003461823  0.002188886     -1.58155  0.11407186
2.  Constant                -0.009565690  0.071829719     -0.13317  0.89408452
3.  TENDANCE                 0.000123570  0.000168576      0.73302  0.46371713
4.  dDZ1{1}                  0.528948943  0.027058150     19.54860  0.00000000

valeur de la statistique de Durbin   h=       0.82964

 statistique Q( 63 )=         36.5083   niveau de significativit\U{e9}   0.9970
 stat. modifiee  Q( 63 - 1 )= 36.5083   niveau de significativit\U{e9}   0.9959

     calcul de phi3 avec H0 (a,0,1) :       1.48544

*****modele sans le tendance avec la constante

Linear Regression - Estimation by Least Squares
Dependent Variable dDZ1
Usable Observations    997      Degrees of Freedom   994
Centered R**2     0.278618      R Bar **2   0.277167
Uncentered R**2   0.278627      T x R**2     277.791
Mean of Dependent Variable      0.0041895622
Std Error of Dependent Variable 1.1905222763
Standard Error of Estimate      1.0121772041
Sum of Squared Residuals        1018.3556763
Regression F(2,994)                 191.9557
Significance Level of F           0.00000000
Log Likelihood                   -1425.24680
Durbin-Watson Statistic             1.970650

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  DZ1{1}                  -0.002258358  0.001447342     -1.56035  0.11899594
2.  Constant                 0.034943028  0.038366414      0.91077  0.36263671
3.  dDZ1{1}                  0.527212494  0.026947981     19.56408  0.00000000


 statistique Q( 63 )=          36.92882 niveau de signific.   0.9964
 stat. modifiee Q( 63 - 1 )=   36.92882 niveau de signific.   0.9953

     calcul de phi1 avec H0 (0,0,1) :       1.21939

 ******** sans tendance ni constante

Linear Regression - Estimation by Least Squares
Dependent Variable dDZ1
Usable Observations    997      Degrees of Freedom   995
Centered R**2     0.278016      R Bar **2   0.277291
Uncentered R**2   0.278025      T x R**2     277.191
Mean of Dependent Variable      0.0041895622
```

```
Std Error of Dependent Variable 1.1905222763
Standard Error of Estimate      1.0120904808
Sum of Squared Residuals        1019.2055057
Log Likelihood                   -1425.66263
Durbin-Watson Statistic             1.969942

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  DZ1{1}                 -0.001534074  0.001209190     -1.26868  0.20485202
2.  dDZ1{1}                 0.526948557  0.026944114     19.55709  0.00000000


 statistique Q( 63 )=          37.44632 niveau de signific.   0.9957
 stat. modifiee Q( 63 - 1 )=   37.44632 niveau de signific.   0.9943
```

On constate ici qu'il y a 1 retard sur l'endogène $DDZ1$ le modèle est donc autorégressif donc on remplace le test de Durbin et Watson par le test H de Durbin.

- On commence par l'étude de l'autocorrélation du modèle. La statistique de H=0829 , il n'y a pas d'autocorrélation d'ordre 1. La statistique de Ljung-Box Q(63-1) a un niveau de significativité de .9959 on décide donc la non autocorrélation d'ordre supérieur à 1. On peut donc faire le test qui ici est le test A. D. F. car comme on l'a vu la procédure a mis 1 retards sur DDZ1 pour obtenir des erreurs non autocorrélées (voir partie 3.2). Donc dans cet exemple on trouve que le modèle s'écrit avec r=2

$$DZ1_t = \rho_1 DZ1_{t-1} + \rho_2 DZ1_{t-2} + at + b + u_t$$

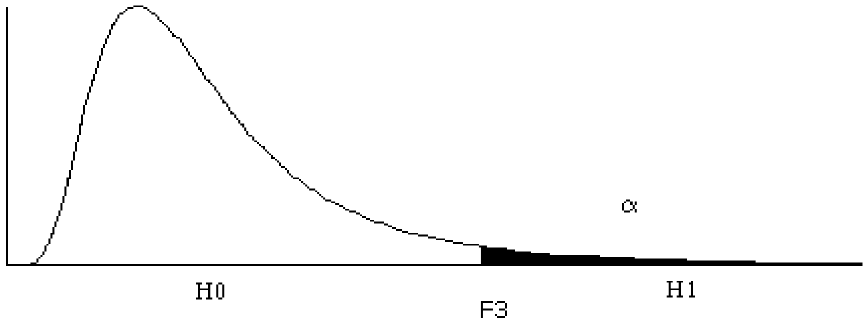
identique à

$$DDZ1_t = \left(\sum_{i=1}^{2} \rho_i - 1 \right) DZ1_{t-1} + \alpha_1 DZ1_{t-1} + at + b + u_t$$

Le travail est donc ici de tester si $\sum_{i=1}^{2} \rho_i - 1 = 0$ soit si $\sum_{i=1}^{2} \rho_i = 1$ donc si on a une racine unitaire.

- On commence par le modèle avec tendance et constante et on regarde tout d'abord si sous ces deux hypothèses de tendance+constante il y a ou non une racine unitaire. Pour cela on fait le test de "Student", $\sum_{i=1}^{3} \rho_i - 1$ a pour t=- 1.58 la table de D.F. pour n=1000 donne la borne ta=-3.41 pour le modèle avec tendance et constante, On décide donc très nettement H0 on a au moins une RU.

- Mais nous ne savons pas si nous avons bien fait de travailler avec la tendance, on va donc construire un test qui teste à la fois il y a une racine unitaire ET la tendance est nulle (remarquons que dans l'exemple précédent il ne faut surtout pas faire ce test car l'une des hypothèse est fausse dans la mesure où nous n'avons pas trouvé de RU.). C'est le test Φ_3 qui se construit comme un test de Fisher (les résultats sont donnés directement) le logiciel donne $\Phi_3 = 1.485$. La table de Φ_3 donne la borne F3 =6.25, comme la valeur trouvée est très inférieure à cette borne on déduit l'hypothèse H01: la tendance est nulle.



- Comme la tendance est nulle il faut recommencer le test sans la tendance avec la constante. On trouve t=-1.56 la table de D. F. pour le modèle sans tendance donne la borne ta=-2.86.on décide donc très nettement encore H0 il y a une RU. au moins.
- Nous testons si nous avons bien travaillé en gardant la constante. Pour cela on fait le test qui teste la présence d'une RU. et de la constante c'est le teste Φ_1, le logiciel donne $\Phi_1 = 1.219$ et la table de Φ_1 donne la borne F1= 4.59, comme la valeur trouvée est très nettement inférieure à la borne on décide H01: la constante est nulle.
- Comme la constante est nulle il faut donc recommencer le test sans la constante. On trouve t=-1.268 la table de D. F. pour un modèle sans tendance ni constante donne la borne ta=-1.95 on décide donc H0 il y a une racine unitaire au moins.
- L'étude est donc terminée, si DZ1 a une RU alors Z1 en a deux. Nous pouvons vérifier qu'il n'y en a pas trois.

4.4.4 Etude de la série DDZ1

```
 **************************************************
ETUDE DE L INTEGRATION DE LA SERIE  DDZ1
**************************************************
********* avec tendance et constante

Linear Regression - Estimation by Least Squares
Dependent Variable dDDZ1
Usable Observations    997      Degrees of Freedom   994
Centered R**2     0.237270      R Bar **2   0.235735
Uncentered R**2   0.237270      T x R**2     236.558
Mean of Dependent Variable      0.0001237071
Std Error of Dependent Variable 1.1589467811
Standard Error of Estimate      1.0131771143
Sum of Squared Residuals        1020.3686977
Regression F(2,994)                 154.6068
Significance Level of F           0.00000000
Log Likelihood                   -1426.23123
Durbin-Watson Statistic             1.967835

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  DDZ1{1}                 -0.474559400  0.026987422    -17.58447  0.00000000
2.  Constant                 0.040407047  0.064555377      0.62593  0.53150557
3.  TENDANCE                -0.000076402  0.000111577     -0.68475  0.49366118

valeur de la statistique Q       38.17239 niveau de signific.       0.99436
    calcul de phi3 avec H0 (a,0,1) :     154.60677
```

On constate ici qu'il y a 0 retard sur l'endogène $DDZ1$ le modèle n'est donc pas autorégressif on utilise le test de Durbin et Watson .

- On commence par l'étude de l'autocorrélation du modèle. La statistique de DW =1.96 , il n'y a pas d'autocorrélation d'ordre 1. La statistique de Ljung-Box Q(63-1) a un niveau de significativité de .9936 on décide donc la non autocorrélation d'ordre supérieur à 1.
- On commence par le modèle avec tendance et constante et on regarde tout d'abord si sous ces deux hypothèses de tendance+constante il y a ou non une racine unitaire. Pour cela on fait le test de "Student",$\rho - 1$ a pour t=- 17.58 la table de D.F. pour n=1000 donne la borne ta=-3.41 pour le modèle avec tendance et constante, On décide donc très nettement H1 on n'a pas de RU.

4.4.5 Conclusion:

La série Z1 a donc 2 et seulement deux RU.
Ces études sont un peu longues et fastidieuses. On pourrait simplifier en notant par exemple, que si une série n'a pas de tendance alors sont écart non plus et passer alors au modèle avec constante sur la série écart , mais je n'aime pas trop cela car si la série a des breaks une incohérence dans les tendances est le signe de ce break dans la série. Il est donc préférable de tout effectuer même si c'est à mourir d'ennui !!!!