ANNALES D'ÉCONOMIE ET DE STATISTIQUE. – N° 12 — 1988

Agrégation de processus autorégressifs d'ordre 1

Esmeralda GONÇALVES, Christian GOURIÉROUX *

RÉSUMÉ. – Dans cet article nous nous intéressons à l'agrégation de processus autorégressifs d'ordre 1. Nous déterminons le modèle de série temporelle satisfait par la série agrégée et caractérisons toutes les séries pouvant s'interpréter comme de tels agrégats.

Nous étudions plus en détail le cas où la distribution d'hétérogénéité du coefficient de régression est une loi bêta. En particulier nous proposons un test d'homogénéité et discutons le sens du biais d'hétérogénéité.

Aggregating Autoregressive Processes

ABSTRACT. – The aim of this paper is the aggregation of AR (1) processes. We determine the dynamic models satisfied by the aggregated series and we characterize all the series which may be interpreted as such an aggregate.

We study more carefully the case of a bêta heterogeneity distribution. In particular we propose an homogeneity test and we discuss the sign of the heterogeneity bias.

* E. GONÇALVES : Université de Coïmbra; C. GOURIÉROUX : CEPREMAP, Paris.

1 Introduction

Les modèles usuels de séries temporelles comme les représentations autorégressives (ou autorégressives moyennes mobiles) sont en pratique utilisées aussi bien dans le cadre d'études microéconomiques (par exemple dans la formalisation des erreurs d'un modèle panel) que macroéconomiques (détermination des multiplicateurs dynamiques). Cependant une telle pratique n'est mathématiquement justifiée que si la classe de séries temporelles retenue possède certaines propriétés d'invariances vis-à-vis de l'agrégation.

Dans cet article, nous nous intéressons essentiellement à l'agrégation de processus autorégressifs d'ordre 1, suivant ainsi l'exemple étudié par GRANGER [1980] qui a donné l'idée d'un certain nombre de comportements du processus agrégé.

Dans le paragraphe 2, nous déterminons le modèle de série temporelle satisfait par la série agrégée. Ceci permet de caractériser toutes les séries temporelles linéaires pouvant s'interpréter comme agrégats de séries autorégressives d'ordre 1.

Dans le paragraphe 3, nous étudions plus en détail le cas où la distribution d'hétérogénéité du coefficient de régression est une loi bêta. On voit alors que les séries agrégées satisfont des écritures hypergéométriques et sortent donc du cadre ARMA traditionnel. En particulier il apparaît que des agrégats de séries autorégressives stationnaires peuvent être non stationnaires ou être stationnaires avec des coefficients moyennes mobiles décroissant lentement (mémoire longue).

Dans le paragraphe 4, nous comparons les corrélations temporelles de la série agrégée avec la moyenne des corrélations temporelles des séries désagrégées et vérifions que les premières sont toujours plus grandes que les secondes, lorsque les liaisons temporelles sont positives. Les écarts entre ces deux types de corrélations fournissent alors des mesures du degré d'hétérogénéité adaptées au cas des séries temporelles.

Finalement nous examinons dans le paragraphe 5 les implications statistiques de ces résultats, proposons un test d'homogénéité et discutons le sens du biais dû à l'oubli de la composante individuelle.

2 Agrégation de processus autorégressifs d'ordre 1

2.a. Le principe

Nous suivrons la procédure usuelle d'agrégation de séries spatio-temporelles telle qu'elle est appliquée par exemple par GRANGER [1980] au cas des processus autorégressifs d'ordre 1. Nous considérons une suite de processus autorégressifs d'ordre 1, avec des coefficients de régression différents et des erreurs pouvant être corrélées entre elles. Cette corrélation éventuelle est prise en compte par l'intermédiaire d'une décomposition de l'erreur en effet temporel et effet croisé. De façon plus précise, on pose :

$$X_{it} = \varphi_i X_{i,t-1} + c_i \varepsilon_t + \eta_{it}, \qquad t \geqslant 0, \quad i = 1, 2, \ldots \tag{1}$$

avec $X_{i,-1} = 0$.

Les bruits (ε_t), (η_{1t}), (η_{2t})... sont indépendants entre eux, centrés, de variances respectives $\sigma^2 = V\varepsilon_t$ et $\mu^2 = V\eta_{it}$, $\forall i$.

Remarquons que nous avons retenu une écriture autorégressive pour les seuls indices temporels positifs. Ceci permet de considérer simultanément le cas de processus asymptotiquement stationnaires, lorsque $|\varphi_i| < 1$ et celui de processus non stationnaires, lorsque $|\varphi_i| \geqslant 1$.

Afin de mener à bien la procédure d'agrégation sur les individus i, nous ferons également des hypothèses sur les coefficients structurels φ_i, c_i. On suppose que ces valeurs peuvent être considérées comme résultant de tirages indépendants dans une même loi (dite distribution d'hétérogénéité), que ces tirages sont indépendants des valeurs prises par les bruits et que les variables φ et c sont indépendantes entre elles. Le paramètre c qui permet d'introduire une hétéroscédasticité est contraint à être strictement positif.

La moyenne calculée sur les n premières séries est :

$$\begin{aligned}
\bar{X}_{n,t} &= \frac{1}{n} \sum_{i=1}^{n} X_{i,t} \\
&= \frac{1}{n} \sum_{i=1}^{n} \sum_{k=0}^{t} \varphi_i^k [c_i \varepsilon_{t-k} + \eta_{i,t-k}] \\
&= \sum_{k=0}^{t} \left[\frac{1}{n} \sum_{i=1}^{n} \varphi_i^k c_i\right] \varepsilon_{t-k} + \sum_{k=0}^{t} \left[\frac{1}{n} \sum_{i=1}^{n} \varphi_i^k \eta_{i,t-k}\right].
\end{aligned}$$

Faisant tendre n vers l'infini, on obtient (p. s.) :

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \varphi_i^k c_i = \mathrm{E}(\varphi^k c) = \mathrm{E}(\varphi^k)(\mathrm{E}\, c),$$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \varphi_i^k \eta_{i,t-k} = \mathrm{E}(\varphi^k \eta_{i,t-k}) = \mathrm{E}(\varphi^k)\mathrm{E}(\eta_{i,t-k}) = 0.$$

On voit alors qu'à la limite le processus agrégé : $\bar{X}_t = \lim_{n\to\infty} \bar{X}_{n,t}$ satisfait (p. s.) une relation du type :

$$(2) \qquad \bar{X}_t = \sum_{k=0}^{t} \mathrm{E}(\varphi^k)\mathrm{E}\, c\, \varepsilon_{t-k} = \sum_{k=0}^{t} \mathrm{E}(\varphi^k)\bar{\varepsilon}_{t-k},$$

avec $\bar{\varepsilon}_t = \mathrm{E}\, c\, \varepsilon_t$.

Nous avons mené cette agrégation de façon heuristique, sans justifier les divers passages à la limite. Il faut évidemment que certaines conditions de régularité soient satisfaites, en particulier que les divers moments $\mathrm{E}\,\varphi^k$ existent.

2.b. **Classification des séries agrégées**

On peut distinguer plusieurs cas selon que la série agrégée est ou non (asymptotiquement) stationnaire et, lorsque la série est stationnaire, en fonction de la vitesse avec laquelle les coefficients moyennes mobiles tendent vers zéro. Ceci conduit aux propriétés suivantes.

PROPRIÉTÉ 3 : (i) La série agrégée est non (asymptotiquement) stationnaire si

$$\sum_{k=0}^{\infty} (\mathrm{E}\,\varphi^k)^2 = +\infty.$$

(ii) Elle est (asymptotiquement) stationnaire dans le cas contraire.

Il est de plus possible de classer les séries stationnaires selon que les corrélations temporelles sont ou non élevées. Ainsi on introduit souvent la notion de processus stationnaire à mémoire longue. Si $X_t = \sum_{k=0}^{\infty} a_k \varepsilon_{t-k}$ est un processus stationnaire, dont les coefficients moyennes mobiles satisfont la contrainte $\sum_{k=0}^{\infty} a_k^2 < \infty$, le processus est dit à mémoire longue si $\sum_{k=0}^{\infty} |a_k| = +\infty$, à mémoire courte sinon.

Cette définition de la mémoire est la plus naturelle; d'autres définitions fondées sur le comportement à l'origine de la densité spectrale ont été proposées dans la littérature (voir par exemple KUNSCH [1987]). Avec la définition retenue, nous avons la propriété suivante.

PROPRIÉTÉ 4 : Si la série agrégée est (asymptotiquement) stationnaire :

(i) elle est à mémoire longue, si $\sum_{k=0}^{\infty} |\mathrm{E}\,\varphi^k| = +\infty$;

(ii) elle est à mémoire courte, si $\sum_{k=0}^{\infty} |\mathrm{E}\,\varphi^k| < +\infty$.

Lorsque les séries désagrégées ne sont pas toutes (p. s.) asymptotiquement stationnaires, on a $\mathrm{P}[|\varphi| \geqslant 1] > 0$. On en déduit que $\mathrm{E}\,\varphi^{2k} \geqslant \mathrm{P}(|\varphi| \geqslant 1)$ et en particulier que $\sum_{k=0}^{\infty} (\mathrm{E}\,\varphi^k)^2 = +\infty$.

PROPRIÉTÉ 5 : Une condition nécessaire pour que la série agrégée soit asymptotiquement stationnaire est que presque toutes les séries désagrégées le soient, c'est-à-dire que la loi de φ soit concentrée sur $]-1, 1[$.

Comme nous le verrons dans le chapitre suivant cette condition n'est pas suffisante. Il est donc possible que la propriété de stationnarité soit perdue par agrégation.

On a de plus les diverses conditions suffisantes de stationnarité.

PROPRIÉTÉ 6 : (i) Si la loi de φ est concentrée sur $]-1, 1[$:

- une condition suffisante pour que la série agrégée soit (asymptotiquement) stationnaire est que $\mathrm{E}\left(\frac{1}{1-\varphi^2}\right) < \infty$;

- une condition suffisante pour que la série agrégée soit à mémoire courte est que $\mathrm{E}\left(\frac{1}{1-|\varphi|}\right) < \infty$.

(ii) Si la loi de φ est concentrée sur $[0, 1[$, la série est à mémoire courte si et seulement si $\mathrm{E}\left(\frac{1}{1-\varphi}\right) < \infty$.

2.c. Caractérisation des séries agrégées

Étant donnée une série $\left[\mathrm{X}_t = \sum_{k=0}^{t} a_k\,\bar{\varepsilon}_{t-k},\ t \in \mathrm{N}\right]$, on peut se demander dans quels cas elle peut être considérée comme provenant d'une agrégation de processus autorégressifs d'ordre 1. Il faut pour cela chercher des conditions pour que les coefficients a_k puissent s'écrire $a_k = \mathrm{E}\,\varphi^k$. De telles conditions sont classiques.

Propriété 7 : La série $\left[X_t = \sum_{k=0}^{t} a_k \bar{\varepsilon}_{t-k},\ t \in \mathrm{N}\right]$, est une série agrégée de processus autorégressifs d'ordre 1, si et seulement si on a :

$$\forall n,\ \forall \alpha_0, \alpha_1 \ldots \alpha_n \in \mathrm{R}, \quad \sum_{i=0}^{n} \sum_{j=0}^{n} \alpha_i \alpha_j a_{i+j} \geqslant 0.$$

Preuve : Voir Choquet [1969], p. 279.

Cette condition correspond simplement à la positivité de la matrice de Hankel infinie :

$$\begin{pmatrix} a_0 & a_1 & a_2 \ldots \\ a_1 & a_2 & \\ a_2 & & \\ \vdots & & \end{pmatrix}$$

Elle est donc facile à vérifier en pratique à partir des mineurs principaux, puisqu'elle équivaut à :

$$\det \begin{pmatrix} a_0 & a_1 & \ldots & a_i \\ a_1 & & & \vdots \\ \vdots & & & \vdots \\ a_i & \ldots & \ldots & a_{2i} \end{pmatrix} \geqslant 0, \qquad \forall i \geqslant 0.$$

Ainsi la première contrainte effective est, normalisant a_0 à 1 :

$$\det \begin{bmatrix} a_0 & a_1 \\ a_1 & a_2 \end{bmatrix} = a_2 - a_1^2 = \mathrm{E}\,\varphi^2 - (\mathrm{E}\,\varphi)^2 \geqslant 0.$$

Il est clair que l'ensemble des sous-déterminants peut être considéré comme une mesure vectorielle d'hétérogénéité des coefficients de régression au niveau désagrégé, la première composante étant identique à la variance de φ. De plus il est parfois possible de retrouver sans ambiguïté la loi du coefficient de régression φ.

Propriété 8 : Lorsque la condition de la propriété 7 est satisfaite, il existe une loi unique pour φ telle que $a_k = \mathrm{E}\,\varphi^k$, $\forall k$ si et seulement si la série $\sum_{k=0}^{\infty} \frac{a_k s^k}{k!}$ est absolument convergente pour une valeur strictement positive s.

Preuve : Voir Cramer [1945].

On peut alors se demander dans quel cas cette loi ne charge que les valeurs φ avec $|\varphi| < 1$. Nous avons vu que si $\mathrm{P}(|\varphi| \geqslant 1) > 0$, les coefficients $a_k = \mathrm{E}\,\varphi^k$ ne tendent pas vers zéro.

Inversement si $\mathrm{P}(|\varphi| < 1) = 1$, on a :

$$\forall \varepsilon > 0, \quad \exists \eta : 1 > \eta > 0 : \mathrm{P}[|\varphi| > 1 - \eta] \leqslant \varepsilon.$$

On a alors : $E|\varphi|^k \leqslant (1-\eta)^k(1-\varepsilon)+\varepsilon, \ \forall k$.

On en déduit : $\overline{\lim_k} E|\varphi|^k \leqslant \varepsilon$ et ceci étant valable pour toute valeur ε positive $\overline{\lim} E|\varphi|^k = 0$.

PROPRIÉTÉ 9 : Lorsque la condition de la propriété 7 est satisfaite, les lois possibles de φ ne chargent que les valeurs $|\varphi| < 1$ si et seulement si $\lim_{k \to \infty} a_k = 0$.

Exemple 10 : A titre d'exemple considérons un processus autorégressif pur $X_t = \varphi_1 X_{t-1} + \ldots + \varphi_p X_{t-p} + \bar{\varepsilon}_t$, $t \geqslant 0$, avec $X_{-1} = \ldots = X_{-p} = 0$. Dans quels cas peut-il s'interpréter comme un agrégat d'autorégressifs d'ordre 1 ? L'écriture moyenne mobile de ce processus est : $X_t = \sum_{k=0}^{t} \left(\sum_{l=1}^{p} \mu_l z_l^k \right) \bar{\varepsilon}_{t-k}$, en appelant z_l, $l=1, \ldots, p$ les inverses des racines (supposées distinctes) du polynôme $\varphi(z) = 1 - \varphi_1 z \ldots - \varphi_p z^p$ et μ_l, $l=1, \ldots, p$, les solutions du système de Cramer :

$$\left\{ \begin{array}{c} \sum_{l=1}^{p} \mu_l = 1, \\ \sum_{l=1}^{p} \mu_l z_l^k - \sum_{j=1}^{k} \varphi_j \sum_{l=1}^{p} \mu_l z_l^{k-j} = 0, \qquad k=1, \ldots, p-1. \end{array} \right.$$

La forme des coefficients moyenne-mobile montre immédiatement que s'il s'agit d'un agrégat d'autorégressifs d'ordre 1, la loi de φ est nécessairement discrète chargeant les valeurs z_l, $l=1, \ldots, p$ avec des poids μ_l, $l=1, \ldots, p$. L'interprétation en terme d'agrégat est ainsi possible lorsque :

(i) toutes les racines du polynôme autorégressif sont réelles,

(ii) les solutions μ_l, $l=1, \ldots, p$ sont toutes positives. Ainsi pour un processus AR (2) ce système se réduit à :

$$\left\{ \begin{array}{c} \mu_1 + \mu_2 = 1 \\ \mu_1 z_1 + \mu_2 z_2 - \varphi_1(\mu_1 + \mu_2) = 0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} \mu_1 = \dfrac{z_2 - \varphi_1}{z_2 - z_1}. \\ \mu_2 = \dfrac{\varphi_1 - z_1}{z_2 - z_1} \end{array} \right.$$

Désignant par z_2 la plus grande des inverses des deux racines, la condition est simplement : z_1, z_2 réelles telles que $z_2 \geqslant \varphi_1 \geqslant z_1$, avec au moins une inégalité stricte.

3 Un exemple d'agrégation

3.a. Agrégation avec loi bêta

Nous allons détailler les résultats du chapitre précédent lorsque la loi du coefficient de régression φ est une loi bêta transformée par homothétie : $B(p, q; \alpha)$, $p > 0$, $q > 0$, $\alpha > 0$. Cette loi retenue par Granger dans son article de 1980 paraît en effet adaptée à ce problème, bien qu'elle présente l'inconvénient de ne charger que les valeurs positives. Cette loi est continue sur $[0, \alpha]$, de densité :

$$g(\varphi) = \frac{\varphi^{p-1}(1-\varphi/\alpha)^{q-1}}{\alpha^p B(p, q)} \underset{[0, \alpha]}{1}(\varphi),$$

avec :

$$B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}.$$

Les moments de cette loi sont :

$$E(\varphi^k) = \alpha^k \frac{\Gamma(p+k)}{\Gamma(p)} \frac{\Gamma(p+q)}{\Gamma(p+q+k)}.$$

La représentation moyenne mobile du processus agrégé est donc :

$$\bar{X}_t = \sum_{k=0}^{t} \frac{\Gamma(p+k)}{\Gamma(p)} \frac{\Gamma(p+q)}{\Gamma(p+q+k)} \alpha^k L^k \bar{\varepsilon}_t.$$

Introduisons le processus

$$\approx{\varepsilon}_t = \begin{cases} \bar{\varepsilon}_t, \text{ si } t \geqslant 0 \\ 0, \text{ sinon} \end{cases}$$

On a :

$$\bar{X}_t = \left[\sum_{k=0}^{\infty} \frac{\Gamma(p+k)}{\Gamma(p)} \frac{\Gamma(p+q)}{\Gamma(p+q+k)} \alpha^k L^k \right] \overset{\approx}{\varepsilon}_t,$$

$$\bar{X}_t = \left[\sum_{k=0}^{\infty} \frac{\Gamma(p+k)}{\Gamma(p)} \frac{\Gamma(p+q)}{\Gamma(p+q+k)} \frac{\Gamma(1+k)}{\Gamma(1)} \frac{\alpha^k L^k}{k!} \right] \overset{\approx}{\varepsilon}_t,$$

(9) $$\bar{X}_t = F[p; 1; p+q; \alpha L] \overset{\approx}{\varepsilon}_t,$$

où F désigne la fonction hypergéométrique ABRAMOWITZ-STEGUN [1965]. On obtient ainsi par agrégation de processus autorégressifs d'ordre 1, une classe de séries temporelles sur laquelle nous allons maintenant donner quelques éléments.

3.b. Classification des séries agrégées

(i) Si $\alpha > 1$, nous savons que les séries désagrégées ne sont pas presque toutes stationnaires et la série agrégée est donc non stationnaire.

(ii) Il nous reste à considérer le cas $\alpha \leqslant 1$.

D'après la formule de Stirling, on a : $\dfrac{\Gamma(p+k)}{\Gamma(p+q+k)} \approx k^{-q}$, lorsque k tend vers l'infini. Ceci permet d'obtenir un équivalent du coefficient moyenne mobile :

$$a_k = \mathrm{E}\,\varphi^k \approx \frac{\Gamma(p+q)}{\Gamma(p)}\,\alpha^k k^{-q}.$$

On en déduit alors la classification suivante pour les séries agrégées :

Valeurs des paramètres		Type de série agrégée
$\alpha > 1$		non stationnaire
$\alpha = 1$	$q \leqslant \frac{1}{2}$	non stationnaire
	$\frac{1}{2} < q \leqslant 1$	stationnaire, mémoire longue
	$1 < q$	stationnaire, mémoire courte
$\alpha < 1$		stationnaire, mémoire courte

On notera qu'il serait aussi possible d'effectuer une distinction parmi les séries stationnaires à mémoire courte : la décroissance des coefficients moyennes mobiles est en effet exponentielle pour $\alpha < 1$, mais seulement hyperbolique pour $\alpha = 1$, $q > 1$.

3.c. Quelques agrégats particuliers

Comme la moyenne d'une loi $\mathrm{B}(p, q, 1)$ est égale à : $\dfrac{p}{p+q} = d$ et sa variance à : $\dfrac{pq}{(p+q)^2\,(p+q+1)}$, il est intéressant d'effectuer un changement de paramètres pour ne faire apparaître que des paramètres facilement interprétables. Ainsi on pourrait retenir :

α donnant la valeur maximale des coefficients,

$d = \dfrac{p}{p+q}$ donnant la moyenne de la loi au paramètre d'homothétie près,

$\gamma = \dfrac{1}{p+q}$, qui permet de faire varier la variance de la loi à moyenne et coefficient d'homothétie donnés.

On voit alors que certains modèles classiques sont retrouvés comme cas particulier du modèle hypergéométrique (9) :

$$\bar{X}_t = F[d/\gamma;\ 1;\ 1/\gamma;\ \alpha L]\,\overset{\approx}{\varepsilon}_t.$$

(i) Si $\gamma = 0$, c'est-à-dire si $p+q = +\infty$, la variance de φ est nulle et le modèle se réduit à un modèle autorégressif d'ordre 1 :

$$\bar{X}_t = [1 - d\alpha L]^{-1}\,\overset{\approx}{\varepsilon}_t.$$

Lorsque $\gamma = 0$, on ne peut évidemment identifier séparément les deux autres paramètres.

(ii) Si $\gamma = p+q = 1$, ce qui implique $0 < p < 1$, on a :

$$\begin{aligned}
\bar{X}_t &= F[p;\ 1;\ 1;\ \alpha L]\,\overset{\approx}{\varepsilon}_t \\
&= \sum_{k=0}^{\infty} \frac{\Gamma(p+k)}{\Gamma(p)} \frac{\alpha^k L^k}{k!}\,\overset{\approx}{\varepsilon}_t \\
&= \sum_{k=0}^{\infty} \frac{-p(-p-1)\ldots(-p-k+1)}{k!}(-1)^k \alpha^k L^k\,\overset{\approx}{\varepsilon}_t, \\
\Leftrightarrow \quad \bar{X}_t &= (1-\alpha L)^{-p}\,\overset{\approx}{\varepsilon}_t.
\end{aligned}$$

Le modèle obtenu est un autorégressif fractionnaire [voir GRANGER-JOYEUX [1980], MANDELBROT-VAN NESS [1968], GONÇALVES [1987]...] :

$$(1-\alpha L)^p \bar{X}_t = \overset{\approx}{\varepsilon}_t, \qquad \text{avec} \quad 0 < p < 1.$$

Si $\alpha = 1$, on retrouve en particulier le modèle ARIMA $(0, p, 0)$. Ce modèle est non stationnaire si $\alpha = 1$ et $p = 1 - q \geqslant \dfrac{1}{2}$.

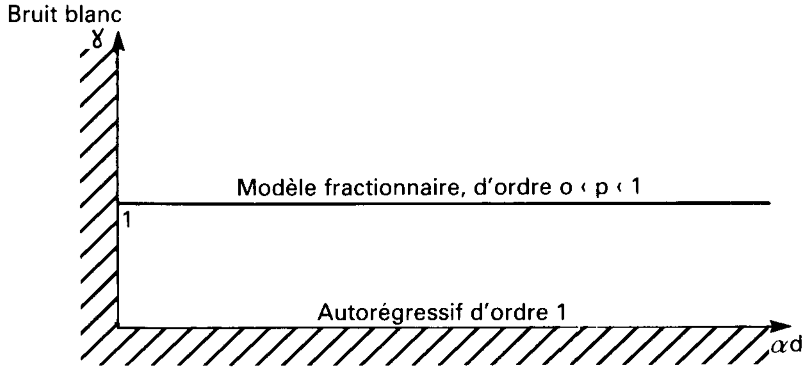
(iii) Finalement remarquant que :

$$E\varphi = \alpha d$$

$$V\varphi = \alpha^2 d(1-d)\frac{\gamma}{1+\gamma},$$

on voit que $E\varphi = V\varphi = 0 \Leftrightarrow \alpha d = 0$.

Le cas limite du bruit blanc correspond donc à la contrainte : $\alpha d = 0$. Dans l'espace $(\gamma, \alpha d)$ ces divers cas particuliers apparaissent de la façon suivante :



3.d. **Agrégations successives**

Nous avons vu que par agrégation de processus autorégressifs d'ordre 1, on obtenait un processus tel que :

$$\bar{X}_t = \sum_{k=0}^{\infty} E(\varphi^k)\tilde{\tilde{\varepsilon}}_{t-k}.$$

En particulier, dès que la loi de φ est non dégénérée, on a $E\varphi^2 \neq (E\varphi)^2$ et le processus $\bar{X}$ n'a pas de représentation autorégressive d'ordre 1. La classe des processus AR(1) ne possède donc aucune propriété de stabilité par agrégation.

On peut maintenant regarder si de telles propriétés peuvent dans certains cas exister pour les processus à représentation hypergéométrique.

La question s'écrit mathématiquement : existe-t-il une loi de probabilité pour p, q, α telle que :

$$E(F[p;\ 1;\ p+q;\ \alpha L])$$

soit de la forme $F[\bar{p};\ 1;\ \bar{p}+\bar{q};\ \bar{\alpha}L]$?

Si nous notons :

$$G(p,\ q,\ \alpha;\ x) = \frac{1}{B(p,\ q)} \int_0^x \frac{\varphi^{p-1}(1-\varphi/\alpha)^{q-1}}{\alpha^p} d\varphi, \qquad 0 \leqslant x \leqslant \alpha,$$

on obtient en intégrant par parties :

$$G(p, q, \alpha; x) = \frac{p}{p+q} G(p+1, q, \alpha; x) + \frac{q}{p+q} G(p, q+1, \alpha; x).$$

Cette relation montre qu'étant donnés des processus hypergéométriques : $\bar{X}_{jt} = F(p_j; 1; p_j+q_j; \alpha L)\tilde{\tilde{\varepsilon}}_t$, $j=1, \ldots, J$ avec $p_j+q_j=\delta$ indépendant de j et $p_{j+1}=p_j+1$, on peut trouver des poids β_j, $j=1, \ldots, J$ de façon que le processus agrégé $\sum_{j=1}^{J} \beta_j \bar{X}_{jt}$ admette aussi une représentation hypergéométrique de paramètres $\bar{p} = \underset{j}{\mathrm{Min}}\, p_j$, $\bar{p}+\bar{q}=\delta$, $\bar{\alpha}=\alpha$.

4 Étude des corrélations

4.a. Accroissement des corrélations

L'étude des liens existant entre la série agrégée et les séries désagrégées peut aussi se faire en regardant les fonctions d'autocorrélation. Au niveau désagrégé la fonction d'autocorrélation est :

$$\rho_i(h) = \varphi_i^h, \qquad i=1, 2, \ldots, \quad h=0, 1, \ldots$$

et donc la moyenne de ces fonctions d'autocorrélation est :

$$\bar{\rho}(h) = E\,\varphi^h, \qquad h=0, 1, \ldots$$

Calculons maintenant la fonction d'autocorrélation associée à la série agrégée. Ce calcul est possible lorsque cette série est stationnaire, ce que nous supposons. On obtient :

(11) $$\tilde{\tilde{\rho}}(h) = \frac{\sum_{k=0}^{\infty} E(\varphi^k)\,E(\varphi^{k+h})}{\sum_{k=0}^{\infty} [E(\varphi^k)]^2}, \qquad \forall h \in N.$$

Si la loi du coefficient de régression charge uniquement les réels positifs, les fonctions $\varphi \to \varphi^k$, $k \geqslant 0$ et $\varphi \to \varphi^h$, $h \geqslant 0$ sont croissantes et donc :

$$\mathrm{Cov}(\varphi^k, \varphi^h) = E(\varphi^{k+h}) - E(\varphi^k)\,E(\varphi^h) \geqslant 0.$$

On en déduit que :

$$\tilde{\tilde{\rho}}(h) \geqslant E\,\varphi^h = \bar{\rho}(h).$$

PROPRIÉTÉ 12 : Si la loi de φ ne charge que les valeurs positives, les corrélations au niveau agrégé sont supérieures aux moyennes des corrélations désagrégées.

4.b. Cas des processus fractionnaires

Considérons par exemple une distribution d'hétérogénéité de type $B(p, q; \alpha)$ avec $p+q=1$, qui correspond à un processus agrégé satisfaisant : $(1-\alpha L)^p \bar{X}_t = \approx{\varepsilon}_t$.

Nous avons : $E\varphi^k = \frac{\Gamma(p+k)}{\Gamma(p)} \frac{\alpha^k}{k!} = \frac{p(p+1)\ldots(p+k-1)}{k!} \alpha^k$.

D'autre part $\left(\text{sous la condition de stationnarité } \alpha < 1 \text{ ou } \alpha = 1, \; 0 < p < \frac{1}{2}\right)$ les moments croisés du processus agrégé se calculent à partir des sommes :

$$\sum_{k=0}^{\infty} E(\varphi^k) E(\varphi^{k+h})$$

$$= \sum_{k=0}^{\infty} \frac{\Gamma(p+k)}{\Gamma(p)} \frac{\Gamma(p+k+h)}{\Gamma(p)} \frac{\alpha^{2k+h}}{k!(k+h)!}$$

$$= \alpha^h \frac{\Gamma(p+h)}{\Gamma(p)\Gamma(h+1)} \sum_{k=0}^{\infty} \frac{\Gamma(p+k)}{\Gamma(p)} \frac{\Gamma(p+k+h)}{\Gamma(p+h)} \frac{\Gamma(h+1)}{\Gamma(k+h+1)} \frac{\alpha^{2k}}{k!}$$

$$= \alpha^h \frac{\Gamma(p+h)}{\Gamma(p)\Gamma(h+1)} F[p, p+h, h+1; \alpha^2].$$

On en déduit :

$$\approx{\rho}(h) = \frac{\alpha^h \Gamma(p+h)}{\Gamma(p)\Gamma(h+1)} \frac{F[p, p+h, h+1; \alpha^2]}{F[p, p, 1; \alpha^2]},$$

c'est-à-dire :

$$\frac{\approx{\rho}(h)}{\bar{\rho}(h)} = \frac{F[p, p+h, h+1; \alpha^2]}{F[p, p, 1; \alpha^2]}. \tag{13}$$

Ce rapport est comme nous l'avons vu auparavant supérieur à 1. En particulier si $\alpha = 1$ on obtient (en utilisant la propriété : $F[a, b, c; 1] = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}$ (ABRAMOWITZ-STEGUN [1965], 15.1.B) :

$$\frac{\approx{\rho}(h)}{\bar{\rho}(h)} = \frac{\Gamma(1-p)}{\Gamma(h+1-p)} h! = \frac{h!}{(1-p)(2-p)\ldots(h-p)}.$$

Forme particulière (13) du rapport $\dfrac{\tilde{\tilde{\rho}}(h)}{\bar{\rho}(h)}$ permet d'étudier facilement son évolution en fonction de h.

Ainsi pour $\alpha < 1$, nous avons :

$$\frac{\tilde{\tilde{\rho}}(h)}{\bar{\rho}(h)} = \sum_{k=0}^{\infty} \beta_k(h)\,\alpha^{2k}$$

avec $\dfrac{\beta_k(h+1)}{\beta_k(h)} = \dfrac{p+k+h}{p+h}\;\dfrac{h+1}{k+h+1} \geqslant 1$, car $p < 1$.

De même, si $\alpha = 1$; un calcul direct donne :

$$\frac{\tilde{\tilde{\rho}}(h+1)}{\bar{\rho}(h+1)} \Big/ \frac{\tilde{\tilde{\rho}}(h)}{\bar{\rho}(h)} = \frac{h+1}{h+1-p} > 1 \quad \text{car } 0 < p < \frac{1}{2}.$$

On en déduit que le rapport $\dfrac{\tilde{\tilde{\rho}}(h)}{\bar{\rho}(h)}$ est fonction croissante de h.

D'autre part, utilisant l'équivalence asymptotique :

$$\mathrm{F}[a,\, b+\lambda,\, c+\lambda,\, z] \approx \frac{1}{(1-z)^a}, \qquad \text{si } |z| < 1 \text{ et si } \lambda \text{ tend vers l'infini,}$$

on voit que ce rapport est asymptotiquement donné par :

$$(14) \qquad \lim_{h\to\infty} \frac{\tilde{\tilde{\rho}}(h)}{\bar{\rho}(h)} = \frac{1}{(1-\alpha^2)^p\,\mathrm{F}(p,\, p,\, 1;\, \alpha^2)}, \qquad \text{si} \quad \alpha < 1;$$

Si $\alpha = 1$ le comportement asymptotique de ce rapport se déduit directement de la formule de Stirling et on obtient :

$$\frac{\tilde{\tilde{\rho}}(h)}{\bar{\rho}(h)} \approx \Gamma(1-p)\,h^p \qquad \left(h \to \infty,\ \alpha = 1,\ 0 < p < \frac{1}{2}\right).$$

5 Test de l'hypothèse d'homogénéité

5.a. Dérivation de la statistique du multiplicateur de Lagrange

Plaçons nous de nouveau dans le cas d'une agrégation de processus autorégressifs d'ordre 1 effectuée selon une loi bêta. Nous avons vu que la série agrégée satisfait une représentation hypergéométrique :

$$\bar{X}_t = F\left(\frac{d}{\gamma};\ 1;\ \frac{1}{\gamma};\ \alpha L\right)\approx{\varepsilon}_t,$$

où γ est un paramètre positif directement lié à la variance du coefficient d'autorégression. Dans ce paragraphe, on se propose de développer un test d'homogénéité, c'est-à-dire un test de l'hypothèse nulle $H_0=\{\gamma=0\}$ contre l'alternative $H_1=\{\gamma>0\}$. Nous avons noté précédemment que si $\gamma=0$, X est un AR(1) de coefficient αd : α et d ne sont pas identifiables; afin de résoudre cette difficulté, nous développons le test sous la contrainte $\alpha=1$, c'est-à-dire pour le modèle : $\bar{X}_t = F\left(\frac{d}{\gamma};\ 1;\ \frac{1}{\gamma};\ L\right)\overset{\approx}{\varepsilon}_t$. On suppose de plus les composantes du bruit indépendantes de même loi. Nous fondons le test sur le principe du multiplicateur de Lagrange. Notant $\Phi_{d,\gamma}(L)\,\bar{X}_t=\overset{\approx}{\varepsilon}_t$ la forme autorégressive du processus, la statistique du multiplicateur de Lagrange fait intervenir l'estimateur du maximum de vraisemblance des paramètres calculé sous l'hypothèse nulle :

$$\hat{d}_T^0 = \frac{\sum_{t=1}^{T} \bar{X}_t \bar{X}_{t-1}}{\sum_{t=1}^{T} \bar{X}_{t-1}^2}$$

(puisque sous H_0, d est le coefficient d'un autorégressif d'ordre 1).

$$\hat{\gamma}_T^0 = 0,$$

$$\hat{\sigma}_T^{0\,2} = \frac{1}{T}\sum_{t=1}^{T}(\bar{X}_t - \hat{d}_T^0 \bar{X}_{t-1})^2 = \frac{1}{T}\sum_{t=1}^{T}(\overset{\approx}{\varepsilon}_t^0)^2,$$

où $\overset{\approx}{\varepsilon}_t^0$ désigne le résidu sous l'hypothèse nulle.

Elle s'exprime alors en fonction des scores estimés sous contraintes :

$$\partial\Phi(L)\,\tilde{\bar{X}}_t\ \partial\Phi(L)\,\tilde{\bar{X}}_t$$

$$Z_{1t} = \frac{\partial\Phi_{\hat{d}_T^0,0}(L)\,\tilde{\bar{X}}_t}{\partial d} \quad \text{et} \quad Z_{2t} = \frac{\partial\Phi_{\hat{d}_T^0,0}(L)\,\tilde{\bar{X}}_t}{\partial \gamma},$$

où $\tilde{\bar{X}}_t=\bar{X}_t$, si $t \geqslant 0$, $\tilde{\bar{X}}_t=0$ sinon. Elle est donnée par (voir HOSKING [1980]) :

$$\xi_{L.M} = \frac{1}{\hat{\sigma}_T^{0\,2}}\left[\sum_{t=1}^{T} Z_{2t}\,\overset{\approx}{\varepsilon}_t^0\right]^2 \Bigg/ \left(\sum_{t=1}^{T} Z_{2t}^2 - \frac{\left[\sum_{t=1}^{T} Z_{2t} Z_{1t}\right]^2}{\sum_{t=1}^{T} Z_{1t}^2}\right), \tag{15}$$

et suit asymptotiquement sous l'hypothèse d'homogénéité une loi du khi-deux à un degré de liberté.

Il nous reste à expliciter la forme des scores.

(i) On a (voir annexe) :

$$\frac{\partial \Phi_{d,0}(\mathrm{L})}{\partial d} = -\mathrm{L}.$$

La variable Z_{1t} est donc :

$$(16) \qquad Z_{1t} = \frac{\partial \Phi_{\partial \Phi \hat{d}^0_{\mathrm{T}},0}(\mathrm{L})}{\partial d} \tilde{\tilde{X}}_t = -\tilde{\tilde{X}}_{t-1}.$$

(ii) Calculons de même l'autre composante du vecteur des scores. Après calcul (voir annexe), on trouve :

$$\frac{\partial \Phi_{d,0}(\mathrm{L})}{\partial \gamma} = d(d-1)\frac{\mathrm{L}^2}{1-d\mathrm{L}}.$$

La variable Z_{2t} est donc :

$$Z_{2t} = \hat{d}^0_{\mathrm{T}}(\hat{d}^0_{\mathrm{T}}-1)\frac{\mathrm{L}^2}{1-\hat{d}^0_{\mathrm{T}}\mathrm{L}}\tilde{\tilde{X}}_t,$$

$$Z_{2t} = \hat{d}^0_{\mathrm{T}}(\hat{d}^0_{\mathrm{T}}-1)(1-\hat{d}^0_{\mathrm{T}}\mathrm{L})^{-1}\tilde{\tilde{X}}_{t-2}.$$

(iii) La statistique du multiplicateur de Lagrange s'écrit finalement :

$$(17) \quad \xi_{\mathrm{L.M}} = \frac{1}{\hat{\sigma}^{02}_{\mathrm{T}}} \times \frac{\left[\sum_{t=1}^{\mathrm{T}} (1-\hat{d}^0_{\mathrm{T}}\mathrm{L})^{-1}\tilde{\tilde{X}}_{t-2}.\tilde{\tilde{\varepsilon}}^0_t\right]^2}{\sum_{t=1}^{\mathrm{T}} [(1-\hat{d}^0_{\mathrm{T}}\mathrm{L})^{-1}\tilde{\tilde{X}}_{t-2}]^2 - \left[\sum_{t=1}^{\mathrm{T}} (1-\hat{d}^0_{\mathrm{T}}\mathrm{L})^{-1}\tilde{\tilde{X}}_{t-2}.\tilde{\tilde{X}}_{t-1}\right]^2 \Big/ \sum_{t=1}^{\mathrm{T}} \tilde{\tilde{X}}^2_{t-1}}$$

5.b. Une statistique de test asymptotiquement équivalente

Une autre forme de cette statistique équivalente sous l'hypothèse nulle s'en déduit en approchant le dénominateur. On a après calcul :

$$\sum_{t=1}^{\mathrm{T}} [(1-\hat{d}^0_{\mathrm{T}}\mathrm{L})^{-1}\tilde{\tilde{X}}_{t-2}]^2 \ \# \ \mathrm{T}\,\hat{\sigma}^{02}_{\mathrm{T}} \frac{1+\hat{d}^{02}_{\mathrm{T}}}{(1-\hat{d}^{02}_{\mathrm{T}})^3},$$

$$\sum_{t=1}^{\mathrm{T}} \tilde{\tilde{X}}_{t-1}(1-\hat{d}^0_{\mathrm{T}}\mathrm{L})^{-1}\tilde{\tilde{X}}_{t-2} \ \# \ \mathrm{T}\,\hat{\sigma}^{02}_{\mathrm{T}} \frac{\hat{d}^0_{\mathrm{T}}}{(1-\hat{d}^{02}_{\mathrm{T}})^2},$$

$$\sum_{t=1}^{\mathrm{T}} \tilde{\tilde{X}}^2_{t-1} \ \# \ \mathrm{T}\frac{\hat{\sigma}^{02}_{\mathrm{T}}}{1-\hat{d}^{02}_{\mathrm{T}}}.$$

Le dénominateur peut donc être approché par :

$$\mathrm{T}\,\hat{\sigma}_{\mathrm{T}}^{02}\,\frac{1+\hat{d}_{\mathrm{T}}^{02}}{(1-\hat{d}_{\mathrm{T}}^{02})^3}-\frac{\mathrm{T}^2\,\hat{\sigma}_{\mathrm{T}}^{04}\,\hat{d}_{\mathrm{T}}^{02}(1-\hat{d}_{\mathrm{T}}^{02})}{(1-\hat{d}_{\mathrm{T}}^{02})^4\,\mathrm{T}\,\hat{\sigma}_{\mathrm{T}}^{02}}$$
$$=\mathrm{T}\,\frac{\hat{\sigma}_{\mathrm{T}}^{02}}{(1-\hat{d}_{\mathrm{T}}^{02})^3}[1+\hat{d}_{\mathrm{T}}^{02}-\hat{d}_{\mathrm{T}}^{02}]=\frac{\mathrm{T}\,\hat{\sigma}_{\mathrm{T}}^{02}}{(1-\hat{d}_{\mathrm{T}}^{02})^3}.$$

Une statistique équivalente à $\xi_{\mathrm{L.M}}$ sous l'hypothèse nulle est donc :

(18)
$$\xi_{\mathrm{L.M}}^{*}=\frac{(1-\hat{d}_{\mathrm{T}}^{02})^3}{\hat{\sigma}_{\mathrm{T}}^{04}}\,\frac{1}{\mathrm{T}}\left[\sum_{t=1}^{\mathrm{T}}(1-\hat{d}_{\mathrm{T}}^{0}\,\mathrm{L})^{-1}\,\tilde{\tilde{\mathrm{X}}}_{t-2}\,.\,\tilde{\tilde{\varepsilon}}_t^{0}\right]^2.$$

5.c. L'aspect unilatéral du test

Les statistiques que nous venons de calculer sont celles habituellement utilisées dans le cas d'un test bilatéral de $\mathrm{H}_0=\{\gamma=0\}$ contre $\mathrm{H}_1=\{\gamma\neq 0\}$. Dans notre application il faut tenir compte de l'aspect unilatéral du test. Ceci conduit à modifier les statistiques de test en prenant :

(19)
$$\mathrm{T}_{\mathrm{LM}}=s\sqrt{\xi_{\mathrm{LM}}}\qquad \text{et}\qquad \mathrm{T}_{\mathrm{LM}}^{*}=s\sqrt{\xi_{\mathrm{LM}}^{*}},$$

où s désigne le signe de $\sum_{t=1}^{\mathrm{T}}(1-\hat{d}_{\mathrm{T}}^{0}\,\mathrm{L})^{-1}\,\tilde{\tilde{\mathrm{X}}}_{t-2}\,.\,\tilde{\tilde{\varepsilon}}_t^{0}$. Ainsi on a :

$$\mathrm{T}_{\mathrm{LM}}^{*}=\frac{(1-\hat{d}_{\mathrm{T}}^{02})^{3/2}}{\hat{\sigma}_{\mathrm{T}}^{02}}\,\frac{1}{\sqrt{\mathrm{T}}}\sum_{t=1}^{\mathrm{T}}(1-\hat{d}_{\mathrm{T}}^{0}\,\mathrm{L})^{-1}\,\tilde{\tilde{\mathrm{X}}}_{t-2}\,.\,\tilde{\tilde{\varepsilon}}_t^{0}.$$

Ces deux statistiques ont asymptotiquement une loi normale centrée réduite sous l'hypothèse nulle $\mathrm{H}_0=\{\gamma=0\}$.

Afin de déterminer la région du rejet, nous allons maintenant étudier leur comportement sous une suite d'alternatives locales du type $\mathrm{H}_{1\,\mathrm{T}}=\left\{\gamma=\frac{\gamma_0}{\sqrt{\mathrm{T}}}\right\}$, avec $\gamma_0>0$.

Pour cette suite le modèle initial :

$$\frac{1}{\mathrm{F}\left(\dfrac{d}{\gamma_{\mathrm{T}}};\,1;\,\dfrac{1}{\gamma_{\mathrm{T}}};\,\mathrm{L}\right)}\bar{\mathrm{X}}_t=\Phi_{d,\,\gamma_{\mathrm{T}}}(\mathrm{L})\,\bar{\mathrm{X}}_t=\tilde{\tilde{\varepsilon}}_t,$$

peut être approché par :

$$\left(\Phi_{d,\,0}(\mathrm{L})+\frac{\partial\Phi_{d,\,0}(\mathrm{L})}{\partial\gamma}\,\frac{\gamma_0}{\sqrt{\mathrm{T}}}\right)\bar{\mathrm{X}}_t=\tilde{\tilde{\varepsilon}}_t,$$

ou par :

$$\bar{X}_t=\left(\frac{1}{\Phi_{d,0}(L)}-\frac{\partial\Phi_{d,0}(L)/\partial\gamma}{\Phi^2_{d,0}(L)}\frac{\gamma_0}{\sqrt{T}}\right)\overset{\approx}{\varepsilon}_t,$$

$$\Leftrightarrow\quad \bar{X}_t=\left[(1-dL)^{-1}+d(1-d)L^2(1-dL)^{-3}\frac{\gamma_0}{\sqrt{T}}\right]\overset{\approx}{\varepsilon}_t$$

$$\Leftrightarrow\quad \bar{X}_t-d\bar{X}_{t-1}=\overset{\approx}{\varepsilon}_t+\frac{d(1-d)L^2}{(1-dL)^2}\frac{\gamma_0}{\sqrt{T}}\overset{\approx}{\varepsilon}_t.$$

Sous cette suite d'alternatives locales, l'estimateur contraint du paramètre d peut être approché par :

$$\hat{d}^0_T=\frac{\sum_t\bar{X}_t\bar{X}_{t-1}}{\sum_t\bar{X}^2_{t-1}}$$

$$=d+\frac{\frac{1}{T}\sum_t\left(\overset{\approx}{\varepsilon}_t+\frac{d(1-d)L^2}{(1-dL)^2}\frac{\gamma_0}{\sqrt{T}}\overset{\approx}{\varepsilon}_t\right)\bar{X}_{t-1}}{\frac{1}{T}\sum_t\bar{X}^2_{t-1}}+o_P(1)$$

$$=d+o_P(1).$$

Le numérateur de la statistique T_{LM} ou T^*_{LM} peut être approché par :

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(1-\hat{d}^0_T L)^{-1}\overset{\approx}{\bar{X}}_{t-2}\overset{\approx}{\varepsilon}{}^0_t$$

$$=\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(1-dL)^{-1}\bar{X}_{t-2}(\bar{X}_t-d\bar{X}_{t-1})+o_P(1)$$

$$=\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(1-dL)^{-1}\bar{X}_{t-2}\left(\overset{\approx}{\varepsilon}_t+\frac{d(1-d)}{(1-dL)^2}\frac{\gamma_0}{\sqrt{T}}\overset{\approx}{\varepsilon}_{t-2}\right)+o_P(1)$$

$$=\frac{1}{T}\sum_{t=1}^{T}(1-dL)^{-1}\bar{X}_{t-2}\gamma_0\frac{d(1-d)}{(1-dL)^2}\overset{\approx}{\varepsilon}_{t-2}+o_P(1),$$

(car $\overset{\approx}{\varepsilon}_t$ est indépendant des valeurs $\bar{X}_{t-2}$, $\bar{X}_{t-3}\ldots$).

Ne conservant que le terme d'ordre supérieur dans l'écriture de $\bar{X}_{t-2}$, on voit que le numérateur est pour la suite d'alternatives locales équivalent à :

$$\frac{1}{T}\sum_{t=1}^{T}(1-dL)^{-2}\overset{\approx}{\varepsilon}_{t-2}\gamma_0\frac{d(1-d)}{(1-dL)^2}\overset{\approx}{\varepsilon}_{t-2}+o_P(1)$$

$$=\gamma_0 d(1-d)V((1-dL)^{-2}\overset{\approx}{\varepsilon}_{t-2}).$$

Comme $\gamma_0>0$ et que $0<d<1$ le numérateur converge donc vers une constante positive.

La région critique du test unilatéral est donc de la forme :

$$W_T = \{T_{LM} > c\} \qquad \text{ou} \qquad W_T^* = \{T_{LM}^* > c\},$$

où c est déterminée par la condition asymptotique sur le risque de première espèce. Si ce risque est fixé à α_0, on doit avoir :

$$\lim_T P(W_T) = \alpha_0 \quad \Leftrightarrow \quad 1 - u(c) = \alpha_0 \quad \Leftrightarrow \quad c = u^{-1}(1-\alpha_0),$$

où u désigne la fonction de répartition de la loi normale centrée réduite.

PROPRIÉTÉ 20 : Le test du multiplicateur de Lagrange au seuil α_0 de l'hypothèse d'homogénéité dans le cadre d'un modèle hypergéométrique consiste à : accepter l'hypothèse d'homogénéité si :

$$T_{LM} < u^{-1}(1-\alpha_0) \qquad [\text{ou si } T_{LM}^* < u^{-1}(1-\alpha_0)],$$

à la refuser dans le cas contraire.

5.d. **Biais d'hétérogénéité**

Finalement, on peut se demander quelles seraient les propriétés des estimateurs contraints par l'hypothèse d'homogénéité, lorsque celle-ci n'est pas satisfaite.

L'estimateur contraint du paramètre d est comme nous l'avons vu donné par :

$$\hat{d}_T^0 = \frac{\sum_t \bar{X}_t \bar{X}_{t-1}}{\sum_t \bar{X}_{t-1}^2}$$

et il converge donc vers la corrélation théorique à l'ordre 1 du processus $\bar{X}$, c'est-à-dire vers $\tilde{\tilde{\rho}}(1)$.

On a vu d'autre part dans la propriété 12 que $\tilde{\tilde{\rho}}(1) \geqslant E\,\varphi$.

Comme d s'interprète justement comme la moyenne de φ, on en déduit l'existence d'un biais positif d'hétérogénéité.

PROPRIÉTÉ 21 : L'estimateur du maximum de vraisemblance de d calculé sous l'hypothèse d'homogénéité est biaisé, lorsqu'il y a hétérogénéité; ce biais est toujours positif.

Ainsi l'oubli de l'hétérogénéité conduit à une surestimation systématique de ce coefficient.

Ce biais peut être explicité en fonction des paramètres. On a :

$$\tilde{\tilde{\rho}}(1) - d = \frac{\sum_{i=0}^{\infty} a_i a_{i+1} - d \sum_{i=0}^{\infty} a_i^2}{\sum_{i=0}^{\infty} a_i^2},$$

où a_i désigne le i-ième coefficient du développement mobile de $\bar{X}$, c'est-à-dire :

$$a_i = \frac{\Gamma((d/\gamma)+i)}{\Gamma(d/\gamma)} \frac{\Gamma(1/\gamma)}{\Gamma((1/\gamma)+i)}.$$

Comme : $a_{i+1} = \dfrac{(d/\gamma)+i}{(1/\gamma)+i} a_i$, on a :

$$\tilde{\tilde{\rho}}(1) - d = \frac{\sum_{i=0}^{\infty} a_i (a_{i+1} - d a_i)}{\sum_{i=0}^{\infty} a_i^2}$$

$$= \frac{(1-d) \sum_{i=0}^{\infty} a_i^2 \dfrac{i}{(1/\gamma)+i}}{\sum_{i=0}^{\infty} a_i^2}.$$

On voit que puisque $0 \leqslant \tilde{\tilde{\rho}}(1) - d \leqslant 1 - d$, le biais tend à s'annuler lorsqu'on se rapproche de la non-stationnarité.

Dérivée de $\Phi_{d,\gamma}$ (L) par rapport à *d* en $\gamma=0$

On a :

$$\Phi_{d,\gamma}(\mathrm{L})=\frac{1}{\mathrm{F}(d/\gamma;\ 1;\ 1/\gamma;\ \mathrm{L})} \qquad \text{et} \qquad \Phi_{d,0}(\mathrm{L})=(1-d\mathrm{L}).$$

On en déduit :

$$\frac{\partial\Phi_{d,\gamma}(\mathrm{L})}{\partial d}=-\frac{(\partial/\partial d)\,\mathrm{F}(d/\gamma;\ 1;\ 1/\gamma;\ \mathrm{L})}{\mathrm{F}^2(d/\gamma;\ 1;\ 1/\gamma;\ \mathrm{L})}.$$

Le terme général du développement en série de la fonction hypergéométrique $\mathrm{F}\left(\frac{d}{\gamma};\ 1;\ \frac{1}{\gamma};\ \mathrm{L}\right)$ est :

$$\frac{\Gamma((d/\gamma)+i)}{\Gamma(d/\gamma)}\ \frac{\Gamma(1/\gamma)}{\Gamma((1/\gamma)+i)}\,\mathrm{L}^i.$$

On a :

$$\frac{\partial}{\partial d}\mathrm{Log}\frac{\Gamma((d/\gamma)+i)}{\Gamma(d/\gamma)}\ \frac{\Gamma(1/\gamma)}{\Gamma((1/\gamma)+i)}=\frac{1}{\gamma}[\psi((d/\gamma)+i)-\psi(d/\gamma)],$$

où $\psi(z)=\dfrac{\Gamma'(z)}{\Gamma(z)}$ est la fonction digamma. Utilisant les propriétés de la fonction digamma (Abramowitz-Stegun [1965], 6.3.6), on a :

$$\frac{\partial}{\partial d}\left[\mathrm{Log}\frac{\Gamma((d/\gamma)+i)}{\Gamma(d/\gamma)}\ \frac{\Gamma(1/\gamma)}{\Gamma((1/\gamma)+i)}\right]$$
$$=\frac{1}{\gamma}\left[\sum_{j=0}^{i-1}\frac{1}{(d/\gamma)+j}\right]=\sum_{j=0}^{i-1}\frac{1}{d+j\gamma}$$

et :

$$\frac{\partial}{\partial d}\left[\frac{\Gamma((d/\gamma)+i)}{\Gamma(d/\gamma)}\ \frac{\Gamma(1/\gamma)}{\Gamma((1/\gamma)+i)}\,\mathrm{L}^i\right]$$
$$=\frac{\Gamma((d/\gamma)+i)}{\Gamma(d/\gamma)}\ \frac{\Gamma(1/\gamma)}{\Gamma((1/\gamma)+i)}\sum_{j=0}^{i-1}\frac{1}{d+j\gamma}\,\mathrm{L}^i.$$

On en déduit la forme du score relatif à d, évalué en $\gamma=0$:

$$\frac{\partial\Phi_{d,0}(\mathrm{L})}{\partial d}=-(1-d\mathrm{L})^2\sum_{i=1}^{\infty} id^{i-1}\mathrm{L}^i$$
$$=-(1-d\mathrm{L})^2\frac{\mathrm{L}}{(1-d\mathrm{L})^2}=-\mathrm{L}.$$

Dérivée de $\Phi_{d,\gamma}(\mathrm{L})$ par rapport à γ en $\gamma=0$

On a :

$$\frac{\Phi_{d,\gamma}(\mathrm{L})-\Phi_{d,0}(\mathrm{L})}{\gamma}$$
$$=\frac{\dfrac{1}{\mathrm{F}(d/\gamma;1;1/\gamma;\mathrm{L})}-(1-d\mathrm{L})}{\gamma}$$
$$=\frac{1-\mathrm{F}(d/\gamma;\ 1;\ 1/\gamma;\ \mathrm{L})+d\mathrm{L}\mathrm{F}(d/\gamma;\ 1;\ 1/\gamma;\ \mathrm{L})}{\gamma\,\mathrm{F}(d/\gamma;\ 1;\ 1/\gamma;\ \mathrm{L})}$$
$$=\frac{-\sum_{i=1}^{\infty}\dfrac{\Gamma((d/\gamma)+i)}{\Gamma(d/\gamma)}\dfrac{\Gamma(1/\gamma)}{\Gamma((1/\gamma)+i)}\mathrm{L}^i+d\mathrm{L}\sum_{i=0}^{\infty}\dfrac{\Gamma((d/\gamma)+i)}{\Gamma(d/\gamma)}\dfrac{\Gamma(1/\gamma)}{\Gamma((1/\gamma)+i)}\mathrm{L}^i}{\gamma\,\mathrm{F}(d/\gamma;\ 1;\ 1/\gamma;\ \mathrm{L})}$$
$$=\frac{\sum_{i=1}^{\infty}\left[d\dfrac{\Gamma((d/\gamma)+i-1)}{\Gamma(d/\gamma)}\dfrac{\Gamma(1/\gamma)}{\Gamma((1/\gamma)+i-1)}-\dfrac{\Gamma((d/\gamma)+i)}{\Gamma(d/\gamma)}\dfrac{\Gamma(1/\gamma)}{\Gamma((1/\gamma)+i)}\right]\mathrm{L}^i}{\gamma\,\mathrm{F}(d/\gamma;\ 1;\ 1/\gamma;\ \mathrm{L})}$$
$$=\frac{\sum_{i=1}^{\infty}\dfrac{\Gamma((d/\gamma)+i-1)}{\Gamma(d/\gamma)}\dfrac{\Gamma(1/\gamma)}{\Gamma((1/\gamma)+i-1)}\left[d-\dfrac{(d/\gamma)+i-1}{(1/\gamma)+i-1}\right]\mathrm{L}^i}{\gamma\,\mathrm{F}(d/\gamma;\ 1;\ 1/\gamma;\ \mathrm{L})}$$
$$=\frac{\sum_{i=1}^{\infty}\dfrac{\Gamma((d/\gamma)+i-1)}{\Gamma(d/\gamma)}\dfrac{\Gamma(1/\gamma)}{\Gamma((1/\gamma)+i-1)}\dfrac{(i-1)(d-1)}{(1/\gamma)+i-1}\mathrm{L}^i}{\gamma\,\mathrm{F}(d/\gamma;\ 1;\ 1/\gamma;\ \mathrm{L})}$$
$$=(d-1)\frac{\sum_{i=1}^{\infty}\dfrac{\Gamma((d/\gamma)+i-1)}{\Gamma(d/\gamma)}\dfrac{\Gamma(1/\gamma)}{\Gamma((1/\gamma)+i-1)}\dfrac{i-1}{1+\gamma(i-1)}\mathrm{L}^i}{\mathrm{F}(d/\gamma;\ 1;\ 1/\gamma;\ \mathrm{L})}.$$

Faisant tendre γ vers zéro, on obtient :

$$\frac{\partial\Phi_{d,0}(\mathrm{L})}{\partial\gamma}=(d-1)(1-d\mathrm{L})\sum_{i=1}^{\infty}d^{i-1}(i-1)\mathrm{L}^i$$
$$=(d-1)(1-d\mathrm{L})\frac{d\mathrm{L}^2}{(1-d\mathrm{L})^2}$$
$$=\frac{d(d-1)\mathrm{L}^2}{1-d\mathrm{L}}.$$

• Références bibliographiques

ABRAMOWITZ, M. et STEGUN, I. (1965). – *Handbook of Mathematical functions*, Dover.

CHOQUET, G. (1969). – *Lectures on Analysis*, vol. II, Representation theory, W.A. Benjamin, Inc., New York.

CRAMER, H. (1945). – *Mathematical Methods of Statistics*, Princeton University Press.

GONÇALVES, E. (1987). – « Une Généralisation des Processus ARMA », *Annales d'Économie et de Statistique*, 5, p. 109-146.

GRANGER, C. (1980). – « Long Memory Relationships and the Aggregation of Dynamic Models », *Journal of Econometrics*, 14, p. 227-238.

GRANGER, C. et JOYEUX, R. (1980). – « An Introduction to Long Memory Time Series and Fractional Differencing », *Journal of Time Series Analysis*, 1, p. 15-29.

HOSKING, J. R. (1980). – « Lagrange Multiplier Tests of Time Series Models », *J. R. Statist. Soc.*, B, 42, 2, 170-181.

KUNSCH, J. (1987). – « Long Range Dependence », Bernoulli, vol. 1, p. 67-74.

MANDELBROT, B. B. et VAN NESS, J. W. (1968). – « Fractional Brownian Motions, Fractional Noises and Applications », *SIAM Review*, 10, p. 422-437.