Professor Régis BOURBONNAIS, PhD
LEDa, Université Paris-Dauphine, France
E-mail: regis.bourbonnais@dauphine.fr
Mara Magda MAFTEI, PhD
The Bucharest Academy of Economic Studies
E-mail: mmmaftei@yahoo.com

ARFIMA PROCESS: TESTS AND APPLICATIONS AT A WHITE NOISE PROCESS, A RANDOM WALK PROCESS AND THE STOCK EXCHANGE INDEX CAC 40

***Abstract*.** *The assumption of linearity is implicitly accepted in the process which generates a time series condition submitted to a ARIMA. That is why, in this paper, we shall discuss the research of long memory in the processes: the fractional ARIMA models, denoted as ARFIMA, where d and D, the degree of differentiation of the filters is not integer. After presenting the characteristics of the ARFIMA process, we shall discuss the long-memory tests (statistics rescaled Range Lo and R/S* Moody and Wu). Finally three examples and tests on a white noise process, a random walk model and the stock index of Paris Stock Exchange (CAC40) will illustrate the method.*

Key-words: *long-memory test, non stationary processes, ARIMA process, ARFIAM process.*

JEL Classification: C12, C30, C32.

1. The ARFIMA process

The ARMA processes are processes of short memory in the sense where the shock at a given moment is not sustainable and does not affect the future evolution of time series. Infinite memory processes such as *DS* (Difference Stationary) processes have an opposite behaviour: the effect of a shock is permanent and affects all future values of the time series (R. Bourbonnais, M. Terraza, 2010). This dichotomy is inadequate to account for long-term phenomena as shown by the works of Hurst (1956) in the field of hydrology.

The long memory process, but not infinite, is an intermediary case, in that the effect of a shock has lasting consequences for future values of the time series, but it will find its "natural" equilibrium level (Mignon V. 1997). This type of behaviour has been formalized by Mandelbrot and Wallis (1968) and Mandelbrot and Van Ness (1968) starting from fractional Brownian motions and from fractional Gaussian noises. From these studies Granger and Joyeux (1980) and Hoskins (1981) define the fractional ARIMA process as ARFIMA. More recently these processes have been extended to seasonal cases (Ray 1993, Porter-Hudak 1990) and are noted as SARFIMA process.

1.1. Definitions

Let us remember that a real process x_t from Wold : $x_t = \sum_{j=0}^{\infty} \psi_j a_{t-j}$ with $\psi_0 =$ 1, $\psi_j \in$ R and a_t is *i.i.d.*(0, σ_a^2) is stationary under the condition that $\sum_{j=0}^{\infty} \psi_j^{\ 2} < \infty$.

The stationary process x_t is a long memory if $\sum_{j=0}^{\infty} \left| \psi_j \right| = \infty$.

Let us consider a process centred on x_t , t = 1, …, n. We say that x_t is a stationary integrated process, noted ARFIMA (p, d, q) if it is written:

$\phi_p(B)(1-B)^d x_t = \theta_q(B) a_t$ with:

- $\phi_p(B)$ and $\theta_q(B)$ are respectively polynomial operators in B of parties AR(p) and MA(q) of the process,
- a_t is *i.i.d.*(0, σ_a^2),
- $d \in$ R.

$(1-B)^d$ is called fractional difference operator and is written starting from the time series expansion:

$$(1-B)^d = \sum_{j=0}^{\infty} C_j^d (-B)^j = 1 - dB - \frac{d(1-d)}{2} B^2 - \cdots - \frac{d(1-d)\cdots(j-d-1)}{j\,!} B^j - \cdots$$

$$= \sum_{j=0}^{\infty} \pi_j B^j$$

With $\pi_j = \dfrac{\Gamma(j-d)!}{\Gamma(j+1)\Gamma(-d)}$ j = 0, 1, and Γ is the Eulerian function.

Be it the process ARFIMA (0, d, 0) : $(1-B)^d x_t = a_t$ also called process FI(d).

It is this process that contains the long-term components, the party ARMA brings together the short-term components.

- When $d < ½$, the process is stationary and it has an infinite moving average representation.

$$x_t = (1-B)^{-d} a_t = \sum_{j=0}^{\infty} \psi_j a_{t-j} = \psi(B) a_t \text{ with } \psi_j = \frac{\Gamma(d+j)}{\Gamma(d)\Gamma(j+1)}$$ or the

function $\Gamma(h)$ is as such : $\Gamma(h) = (h-1)! = \begin{cases} \int_O^{\infty} t^{h-1} e^{-t} dt & \text{if } h > 0 \\ \infty & \text{if } h = 0 \\ \frac{\Gamma(1+h)}{h} & \text{if } h < 0 \end{cases}$

and $\Gamma(1/2) = \pi^{1/2}$

- When $d > -½$, the process is invert and has infinite autoregressive representation:

$$(1-B)^d x_t = \pi(B) x_t = \sum_{j=0}^{\infty} \pi_j x_{t-j} = a_t \text{ with } \pi_j = \frac{\Gamma(j-d)}{\Gamma(-d)\Gamma(j+1)}$$

The asymptotic value of the coefficients ψ_j et π_j : $\lim_{j\to\infty} \psi_j \approx \frac{j^{d-1}}{\Gamma(d)}$ and $\lim_{j\to\infty} \pi_j \approx \frac{j^{-d-1}}{\pi(-d)}$ decrease with a hyperbolic rhythm at a rate which is lower than the exponential rate of the process ARMA. The FAC has the same type of behaviour which allows to characterize the process FI(d).

- Finally if (Hoskins 1981):

$0 < d < ½$, the process FI(d), is a long-memory process

$d < 0$, the process FI(d) is an anti–persistent process,

$– ½ < d < 0$, the process FI(d) is not of long-memory, but it does not have the behaviour of ARMA. This intermediate case called anti-persistent by Mandelbrot corresponds to alternations of increases and decreases in the process. This behaviour is also called the "Joseph effect" by reference to the Bible.

The process FI(d), thus stationary is invert for $– ½ < d < ½$.

1.2. Long-memory tests

a) The "*Rescaled Range*" statistics [1]

The statistics *R/S* was introduced in 1951 in a study related to the debits of the Nile by the hydrologist Harold Edwin Hurst. His purpose is to find the intensity of an aperiodic cyclical component in a time series considered one of the aspect of the long-term dependence (long memory) developed by Mandelbrot.

Be it x_t a time series producing a stationary random process with $t = 1, \ldots, n$ and $x_t^* = \sum_{u=1}^{t} x_u$ the cumulated time series. The statistics R/S noted Qn is the extent Rn of partial sums of standard deviations of the series from its mean divided by its standard deviation Sn:

$$Q_n = \frac{R_n}{S_n} = \frac{\max\limits_{1\le k\le n}\sum_{j=1}^{k}\left(x_j - \bar{x}_n\right) - \min\limits_{1\le k\le n}\sum_{j=1}^{k}\left(x_j - \bar{x}_n\right)}{\left[\frac{1}{n}\sum_{j=1}^{n}\left(x_j - \bar{x}_n\right)^2\right]^{1/2}}$$

The first term of the numerator is the maximum k of partial sums of the k in standard deviation of x_j from its average. This term (max) is always positive or zero. By definition the second term (in min) is always negative or null. Therefore Rn is always positive or null.

The statistics $\tilde{Q}_n$ is always non-negative.

The statistics H of Hurst applied to a time series x_t is based on the division of time into intervals of length d, for given d we obtain $(T + 1)$ sections of time. The statistics H- is calculated on each section (Mandelbrot) using the previous method of Hurst taking into account the gap operated on the time scale. In this case:

$R_n(t, d) = \max\limits_{0\le u\le d}[\Delta(u)] - \min\limits_{0\le u\le d}[\Delta(u)]$ where $[\Delta(u)]$ is the linear interpolation of $x_t^* = \sum_{s=1}^{t} x_s$ between t et $t + d$; be it $\Delta(u) = \left[x_{t+u}^* - x_t^*\right] - \frac{u}{d}\left[x_{t+d}^* - x_t^*\right]$ it is about the expression brought at the difference d of $\sum_{j=1}^{k}\left(x_j - \bar{x}_n\right)$ used in order to calculate R_n.

The standard deviation is then written: $S_n(t, d) = \left[\frac{1}{d}\left[\sum_{0\le u\le d} x_{t+u}^2\right] - \frac{1}{d^2}\left[\sum_{0\le u\le d} x_{t+u}\right]\right]$.

We can calculate R_n/S_n for each of $(T + 1)$ section of d length but also their arithmetic average. We can also demonstrate (Mandelbrot–Wallis) that the

[1] We may consult for this paragraph Mignon (1997).

intensity of the long-term dependence is given by the coefficient H situated between zero and one in the relationship:

$Q_n = \frac{R_n}{S_n} \underset{d \to \infty}{\approx} cd^H$. Be it $\log(R_n/S_n) = \log c + H \log d$,

H is the estimator of OLS (Ordinary Least Square) in this relation.

In real life, we build M fictionary samples and we choose M arbitrary starting points of the time series. This starting point is given by : : $t = \frac{n(j-1)}{M} + 1$ and the length of the j sample is: $l = n\frac{(M-j-1)}{M}$.

The linear adjustment of the cloud[2] obtained leads to the estimation of the exponent Hurst.

The r^2 of the cloud depends on the initial difference obtained. We remember (Hernard, Moullard, Strauss–Khan, 1978, 1979) for H the one which gives the maximum r^2 maximum for an initial difference d.

The interpretation of the H values is the following:

If $0 < H < ½ \Rightarrow$ anti-persistent process,

If $H = ½ \Rightarrow$ a simply random process or ARMA process. There is a long term dependence absence.

If $½ < H < 1 \Rightarrow$long-term process, the dependence is even stronger as H tends towards 1.

b) The statistics of Lo

The statistic of the exponent Hurst can not be tested because it is too sensitive to the short term dependence. Lo (1991) shows that the analysis proposed by Mandelbrot can be concluded towards the presence of long memory, while the time series has only a short-term dependence. Indeed, in this case, the exponent Hurst by analysing R/S is biased upward. Lo proposes a new modified statistics *R/S* noted: $\widetilde{Q}_n = \frac{R_n}{\hat{\sigma}_n(q)}$

$$\widetilde{Q}_n = \frac{\max\limits_{1 \le k \le n} \sum\limits_{j=1}^{k} \left(x_j - \bar{x}_n\right) - \min\limits_{1 \le k \le n} \sum\limits_{j=1}^{k} \left(x_j - \bar{x}_n\right)}{\left[\frac{1}{n} \sum\limits_{j=1}^{n} \left(x_j - \bar{x}_n\right)^2 + \frac{2}{n} \sum\limits_{j=1}^{q} \omega_j(q) \left(\sum\limits_{i=j+1}^{n} \left(x_j - \bar{x}_n\right)\left(x_{i-j} - \bar{x}_n\right)\right)\right]^{1/2}}$$

This statistics is different from the previous one Q_n by its denominator, which takes into account not only the variances of individual terms but also the auto-covariance weighted accordingly to differences of q as related to:

[2] For which the first points are removed (transitory phase).

$$\omega_j(q) = 1 - \left| \frac{j}{q+1} \right| \text{ where } q < n.$$

Andrews–Lo (1991) proposed the following rule for q:

$q = [k_n]$ = whole party of k_n $\quad k_n = \left[\left(\frac{3\text{n}}{2} \right)^{1/3} \left(\frac{2\hat{\rho}}{1-\hat{\rho}} \right)^{2/3} \right]$ $\hat{\rho}$ is the estimation of the autocorrelation coefficient of order 1 and in this case $\omega_j = 1 - \left| \frac{j}{k_n} \right|$.

Lo proves that under the hypothesis H_0 : $x_t \Rightarrow i.i.d.(0,\ \sigma_x^2)$ and for n which tends towards the infinity, the asymptotic distribution of $\widetilde{Q}_n$ converges step by step towards V: $\frac{1}{\sqrt{n}} \widetilde{Q}_n \to V$ where V is the rank of a Brownian bridge, a process with independent Gaussian increases constrained to unity and for which $H = ½$.

The distribution of the random variable V is given by Kennedy (1976) and Siddiqui (1976): $F_V(v) = 1 + 2\sum_{k=1}^{\infty} \left(1 - 4k^2 v^2\right) e^{-2(k\,v)^2}$

The critical values of this symmetrical distribution the most commonly used are:

$P(V<\nu)$	0.005	0.025	0.05	0.1	0.2	0.3	0.4	0.5	0.543	0.6	0.7	0.8	0.9	0.95	0.975	0.995
ν	0.721	0.809	0.861	0.927	1.018	1.09	1.157	1.223	$\sqrt{\pi/2}$	1.294	1.374	1.473	1.620	1.747	1.862	2.098

The calculation of H is done as above and Lo analyzes the behaviour of $\widetilde{Q}_n$ under alternative long-term dependency. He then shows that:

$$V = \frac{1}{\sqrt{n}} \widetilde{Q}_n \xrightarrow{P} \begin{cases} \infty \text{ pour } H \in [0.5;1] \\ 0 \text{ pour } H \in [0;0.5] \end{cases}$$

Under the hypothesis of H_0, there is a short memory in the time series ($H \in [0{,}5;1]$). For an acceptance threshold at 5% H_0 is accepted if $\nu \in [0{,}809\,;\,1{,}862]$.

He concludes that: "For the values of H between 0,5 and 1 the acceptance threshold of long memory at 10% is ν > 1,620. For the values of H between 0,5 and 1 the acceptance threshold of the anti-persistent hypothesis at 10% is ν > 0,861."

We can verify that there is a relation between the values d and the ARFIMA processes and H of the exponent Hurst ($d = H – 0{,}5$).

c) The statistics *R/S** modified by Moody and Wu (1996)

The statistics Lo fix some flaws of the R/S traditional statistics of Hurst when the number of observations is too low.

Moody and Wu show on the application of exchange rates:

- There is an error in estimating the extent of *R* related to short-term dependencies in the series. The value of the traditional *R/S* statistics led to accepting the existence of a long-term component absent in the generating process. When the number of observations is important, the statistics of Lo corrects this error.
- For a small number of observations, the statistics of Lo and the exponent Hurst are poorly estimated.
- For a number of important observations, the right line corresponding to the Lo statistic is independent of *q*: the traditional statistics and those of Lo have the same exponent Hurst.

Moody and Wu suggest introducing a different denominator *S** in the statistics:

$$S^{*} = \left[\left[1 + 2\sum_{j=1}^{q} \omega_j(q)\frac{n-j}{n^2}\right]\hat{\sigma}^2(n) + \frac{2}{n}\sum_{j=1}^{q}\omega_j^2(q)\left(\sum_{i=j+1}^{n}\left(x_j - \bar{x}_n\right)\left(x_{i-j} - \bar{x}_n\right)\right)\right]^{1/2}$$

or: $\hat{\sigma}^2(n) = \frac{1}{n-1}\sum_{j=1}^{n}\left(x_j - \bar{x}_n\right)^2$ is the estimation of the variance.

$\omega_j(q) = 1 - \frac{j}{q+1}$ are weights such as $q < 1$ so that the denominator of the statistics be positive. For $q = 0$ the statistics of Moody and of Wu lead neither to the traditional statistics nor to that of Lo.

Applications: simulation and calculation of the statistics of Hurst, Lo and of Moody and Wu for a white noise, for a random walk model and for CAC40 (index stock exchange Paris).

1.3. Application for a white noise of 1 000 observations.

We have simulated a white noise of 1 000 observations included between (–100) and (+100) and we have calculated the statistics of Hurst, Lo and of Moody and Wu.

For the statistics of Hurst we have in mind: an initial gap of 20, 10 samples and a threshold point of 200, which allows to interpolate *H* on the thirty most significant values, be it 36% observations.

Results

Threshold point	50	100	200	230
H	0.203	0.302	0.481	0.503
R^2	0.629	0.761	0.863	0.853

For the statistics of Lo we have 3 samples and the interpolation is realised on 25% estimations.

Results

Threshold point	50	110	150	210
H	0.241	0.384	0.428	0.503
R^2	0.679	0.883	0.897	0.871
ν	0.893	1.03	1.11	1.07

Whatever the method, the most reliable values of *H* are those for which R^2 tends to one. This is the case for the gap between 200 and 300 (Hurst) and 150 and 210 (Lo). We note that *H* tends to 0.5 in accordance with the theory.

For values of *H* tending towards 0.5, the variable ν must be between 0.861 and 1.620 to accept the hypothesis of zero memory: this is the case with this exercise. Hurst's method evaluates the memory around about 200 times the periodicity, whereas that of Lo estimates it at just 150.

Simulation results (Tests of Lo and Moody–Wu)

White noise (*n* = 1000) : exponent Hurst and statistics *R/S* modified (Lo)

q	0	2	4	6	8
H	0.484	0.426	0.403	0.364	0.341
V	1.042	1.027	1.012	1.007	0.999

White noise (*n* = 1000): exponent Hurst and statistics *R/S* modified (Moody–Wu)

q	0	2	4	6	8
H	0.489	0.447	0.440	0.422	0.418
V	1.041	1.025	1.00	1.004	0.995

Random walk (n = 1000): exponent Hurst and statistics *R/S* modified (Lo)

q	0	2	4	6	8
H	–0.028	–0.032	–0.063	–0.087	–0.111
V	11.8	7.82	6.06	5.13	4.53

Random walk (n = 1000): exponent Hurst and statistics *R/S* modified (Moody–Wu)

q	0	2	4	6	8
H	0.023	–0.020	–0.044	–0.062	–0.08
V	13.51	7.79	6.03	5.09	4.50

1.4. Application to a random walk of 1 000 observations.

We have simulated a random market $x_t = x_1 + \sum_{i=1}^{1000} a_i$ with $a_t \Rightarrow i.i.d.(0, 1)$ and $x_1 = 1$ on 1 000 observations and we calculated the statistics of Hurst and of Lo.

For the statistics of Hurst we have chosen an initial difference of 25 and of 10 samples. For interpolation, we used gaps superiors at 120 or 72 estimations on 220 (32%).

Results

Threshold point	50	70	110	200
H	0.683	0.817	0.889	0.903
R^2	0.914	0.975	<u>0.986</u>	0.970

The calculation of the statistics of Lo is done under the same conditions but with a unique sample:

Results

Threshold point	50	100	150	210
H	0.789	0.885	0.903	0.892
R^2	0.978	<u>0.993</u>	0.989	0.986
ν	6.92	7.52	7.62	7.57

The most reliable value of H by the method of Hurst is 0.889, which can therefore conclude towards the presence of a long memory. The one given by Lo is made for a gap between 100 and 150, which is about H = 0.9.The value ν is superior to 1.620 and it confirms the structure of long-term dependency.

These calculations obtained from a non-stationary series show the misunderstanding that can be made with these tests.

1.5. The statistics of Hurst and of Lo on the data of CAC40 known for 1109 days

Finally, on a series of CAC40 (index of Paris Stock Exchange), we have used the method GPH and the maximum of likelihood in order to estimate the order d of the generating process FI(d) of the raw time series and of first differences.

The generating process of CAC40 contains a unitary root. Statistics H of Hurst calculated on the raw series are of around 0.9 for a threshold point between 50 and 60 and have a value equivalent to a gap as by comparing to the statistics of Lo (ν= 7.14 superior to 1.62).

We could infer the existence of a positive dependence between 40 and 70 values. In fact when the generating process is stationary by the transition to first differences, the statistics of Hurst and Lo are respectively of 0.46 and of 0.45 and the value of the coefficient ν is of 1.294 less than 1.620. We can then conclude that there is no long-term dependency in the CAC40 series in first differences and that the results issued from the raw series are not consistent with the assumptions of applying tests.

The calculations are made with the software Gauss and TSM under Gauss. The results are the following:

Estimated GPH

	Raw series	Differential series	Standard deviation (Differential series)
No window	1.130	–0.096	0.161
Rectangular	1.079	–0.166	0.178
Bartlett	1.048	–0.126	0.103
Daniell	1.043	–0.126	0.126
Tukey	1.060	–0.139	0.112
Parzen	1.0602	–0.123	0.092
Bartlett–Priestley	1.067	–0.151	0.138

Estimation by maximum likelihood
series in level

Number of observations = 1109
Number of estimated parameters: 2
Value of the likelihood function = –5 202.185

Parameter	Estimation	Standard deviation	t–statistics	Prob.
d	1.155	0.019	62.28	0.000
Sigma	26.413	0.0561	47.10	0.000

Series in first differences

Number of observations = 1108
Number of estimated parameters: 2
Value of the likelihood function = –4 950.552

Parameter	Estimation	Standard deviation	*t*–statistics	Prob.
d	0.034	0.029	1.199	0.231
Sigma	21.097	0.448	47.08	0.000

These results show that the time series has a unit root, as with or without window, the GPH estimator is close to 1 as well as the one of the maximum likelihood.

When the time series is differentiated in order to become stationary, according to the theory, the hypothesis H_0 of nullity of the coefficient of fractional integration is accepted in both cases:

- estimation GPH $\left(\frac{d^{gph}}{Std - error} < 1.96 \right)$.
- maximum de likelihood (*cf.* the critical probabilities).

Finally, the relationship $H = - 0.5$ (H = Hurst statistic), can help us to verify that it leads to a result contradictory to the raw series (d ≈ 0.4 by the relation and $d \approx 1$ by calculation). For the differentiated series, we obtain $d \approx -0.037$ from H of Hurst and $H \approx 0.05$ by the statistics of Lo. These results are according to the estimations.

REFERENCES

[1]**Bourbonnais R., Terraza M. (2010),** ***Analyse des séries temporelles en économie*****,** Dunod, 3 ème Edt.;

[2]**Chen G., Abraham B., Peiris M.S. (1994),** ***Lag Window Estimation of the Degree of Differencing in Fractionally Integrated Time Series***; *Journal of Time Series Analysis*, Vol. 15;

[3]**Chung C. F. (1996),** ***A Generalized Fractionally Integrated Autotregressive Moving Average Process****; Journal of Time Series Analysis*, Vol. 17;

[4]**Geweke J., Porter-Hudak S. (1983),** ***The Estimation and Application of Long Memory Time Series Models*****;** *Journal of Time Series Analysis*, Vol. 4;

[5]**Granger C.W.J., Joyeux R. (1980),** ***An Introduction to Long Memory Time Series and Fractional Differencing*** ; *Journal of Time Series Analysis*, Vol. 1;

[6]**Gray H. L., Zang N. F., Woodward W. (1989),** ***A Generalized Fractional Process****; Journal of Time Series Analysis,* Vol. 10;

[7]**Hassler U. (1993),** ***Regression of Spectral Estimators with Fractionally Integrated Time Series*** *; Journal of Time Series Analysis*, Vol. 14;

[8]**Hassler U. (1994),** ***Specification of Long Memory in Seasonal Time Series*****;** *Journal of Time Series Analysis*, Vol. 15;

[9]**Hernad D., Mouillard M., Strauss-Kahn D. (1978),** ***Du bon usage de R/S***, *Revue de Statistique Appliquée*, Vol. 4;

[10]**Hernad D., Mouillard M., Strauss-Kahn D. (1979),** ***Forme typique du spectre et dépendance temporelle en économie***, *Revue de Statistique Appliquée*, Vol. 4;

[11]**Hosking J. R. M. (1981),** ***Fractional Differencing***; *Biometrika*, Vol. 68;

[12]**Hurst H. (1951),** ***Long Term Storage Capacity of Reservoirs***; *Transaction of the American society of civil engineers*, Vol. 116;

[13]**Hurvich C., Ray B. (1995),** ***Estimation of the Memory Parameter of Nonstationary or Noninvertible Fractionally Integrated Processes****; Journal of Time Series Analysis*, Vol. 16;

[14]**Lo A.W. (1991),** ***Long Term Memory in Stock Market Prices***; *Econometrica*, Vol. 59;

[15]**Mandelbrot B., Van Ness J.W. (1968),** ***Fractional Brownian Motion, Fractional Noises and Application****; SIAM Review*, Vol.10;

[16]**Mandelbrot B., Wallis J. (1968),** ***Noah, Joseph, and Operational Hydrology*****,** *Water resources research*, Vol. 4;

[17]**Mignon V. (1997),** ***Marchés financiers et modélisation des rentabilités boursières*****;** Economica Publishing House, Bucharest;

[18]**Moody J., Wu L. (1996),** ***Improved Estimates for the Relaxed Range and Hurst Exposant****; Neural Network in Engeneering, World Scientific*;

[19]**Porter-Hudak S. (1990),** ***An Application of the Seasonal Fractionally Differenced Model to the Monetary Aggregates***; *Journal of the American Statistical association*, Vol. 85;

[20]**Ray B. (1993),** ***Long Range Forecasting of IBM Product Revenues Using a Seasonal Fractionally Differenced ARMA Model***; *International Journal of Forecasting*, Vol. 9.