PORTUGALIAE MATHEMATICA
Vol. 57 Fasc. 4 – 2000

ASYMPTOTIC SEPARATION IN BILINEAR MODELS

E. Gonçalves, C.M. Martins and N. Mendes-Lopes

Abstract: This paper presents a generalization of a non-classical decision procedure for simple bilinear models with a general error process, proposed by Gonçalves, Jacob and Mendes-Lopes [11]. This decision method involves two hypotheses on the model and its consistence is obtained by establishing the asymptotic separation of the sequences of probability laws defined by each hypothesis. Studies on the rate of convergence in the diagonal case are presented and an exponential decay is obtained. Simulation experiments are used to illustrate the behaviour of the power and level functions in small and moderate samples when this procedure is used as a test.

1 – Introduction

The asymptotic separation of two families of probability laws is a probabilistic notion that can be implemented in the statistical inference of stochastic processes to construct new kinds of convergent tests or estimators.

Let $X=(X_t,\, t\in\mathbb{Z})$ be a real valued stochastic process with a law which belongs to a set of parametric laws $(P_\theta,\, \theta\in\Theta)$. Let $\{\Theta_1,\Theta_2\}$ be a partition of Θ. Following Geffroy [7], we say that the two families of laws $(P_\theta,\, \theta\in\Theta_1)$ and $(P_\theta,\, \theta\in\Theta_2)$ are uniformly asymptotically separated if there exists a sequence of Borel sets of $\mathbb{R}^T$, $(A_T,\, T\in\mathbb{N})$, such that

$$\begin{cases} \inf\limits_{\theta\in\Theta_1} P_\theta^T(A_T) \underset{T\to+\infty}{\longrightarrow} 1, \\[2ex] \sup\limits_{\theta\in\Theta_2} P_\theta^T(A_T) \underset{T\to+\infty}{\longrightarrow} 0\ , \end{cases}$$

where P_θ^T denotes the probability law of $(X_1,X_2,...,X_T)$.

Received: July 28, 1999; *Revised*: April 1, 2000.
AMS Subject Classification: 62M10, 62F03.
Keywords: Time series; Asymptotic separation; Bilinear models; Test.

In Geffroy [7,8], we find conditions on the two families of laws under which the uniform asymptotic separation is stated; moreover, he has obtained a uniform lower bound (resp., upper bound) of $P_\theta^T(A_T)$, $\theta \in \Theta_1$ (resp., of $P_\theta^T(A_T)$, $\theta \in \Theta_2$) which lead to the rate of convergence of these sequences.

So, if the sets A_T only depend on the observable part of the process, we may use this procedure to construct sequences of convergent tests of the hypothesis $\theta \in \Theta_1$ against the alternative $\theta \in \Theta_2$, with acceptance regions A_T (Tassi and Legait [21]).

This was done by Moché [16], who applied Geffroy's results to test the signal function, $\theta \in \Theta$, of a model $X = (X_t,\, t \in \mathbb{Z})$ such that $X_t = \theta(t) + \varepsilon_t$, where $\varepsilon = (\varepsilon_t,\, t \in \mathbb{Z})$ is a classical white noise. The regions A_T involved were constructed using the variational distance between the conditional laws $P_{\theta_1}^{\underline{X}_{T-1}}$ and $P_{\theta_2}^{\underline{X}_{T-1}}$, where $\underline{X}_{T-1}$ is the σ-field generated by $X_{T-1}, X_{T-2}, ...$, and θ_1, θ_2 are particular functions θ such that $\theta_1 \in \Theta_1$ and $\theta_2 \in \Theta_2$.

If we deal with more general models like, for instance, ARMA models, Geffroy results are not directly applied. Nevertheless, the procedure used to construct regions A_T leads to convergent tests of simple hypotheses for those kind of models; in particular, consistent tests had been obtained by Massé and Viano [15] for AR(1) models with independent error processes and by Gonçalves, Jacob and Mendes-Lopes [9] and Gonçalves and Mendes-Lopes [10] for AR(p) models with a general non-independent error process.

We point out that, if the procedure is used as a test, the test obtained is not a Neyman–Pearson classical one; in fact, as the size is not fixed *a priori*, we do not privilege any one of the test hypotheses and, so, a symmetrical role is attributed to both.

In this paper, with the aim of proposing a decision procedure to distinguish between simple bilinear models and error processes, we consider the bilinear model $X = (X_t,\, t \in \mathbb{Z})$ defined by

$$X_t = \varphi\, X_{t-k}\, \varepsilon_{t-l} + \varepsilon_t\,, \qquad k > 0, \quad l > 0\ ,$$

where $\varepsilon = (\varepsilon_t,\, t \in \mathbb{Z})$ is a strictly stationary and ergodic error process with certain conditions on the conditional laws of ε_t; we state the asymptotic separation of the two families of probability laws associated with the hypotheses $\varphi = 0$ and $\varphi = \beta$ ($\beta \neq 0$, fixed). We construct separation sets A_T using the variational distance between two particular conditional distributions of X_t when the parameter values are $\varphi = 0$ and $\varphi = \beta$. In the diagonal case, we also obtain an explicit upper bound for the probability of $\overline{A}_T$ when $\varphi = 0$, which converges exponentially to zero.

As the separation sets depend on the non-observable error process of the model in this case, in order to use these results as a test procedure, we have to show that they remain true when the error process is replaced by the residual process. This problem has been treated by simulation and the results obtained lead us to conjecture its truthfulness. A simulation study is then presented in the last paragraph for models with Cauchy error processes.

The proposal developed here opens a way to a new methodology of test for bilinear models. In fact, to our knowledge the tests currently available for bilinear models are based on classical methods, like Lagrange multiplier ones (e.g., Saikkonen and Luukonnen [19]). Rank-based tests have also been proposed for this class of models (e.g., Benghabrit and Hallin [2,3]) as well as a nonparametric test for a class of models that includes the bilinear order one process (Diebolt and Ngatchou Wandji [5]).

Concerning our proposal, we point out the easy implementation of this test methodology, as the construction of the sets A_T is directly based on the process trajectories. Moreover, it has the great advantage of being applicable to models with general error processes; in fact, even the usual hypotheses of existence of second order moments are unnecessary here. We note that the upper bound obtained for the probability of $\overline{A}_T$ allows us to calculate the minimum number T for which the test has level at least equal to α, $\alpha \in\,]0,1[$, arbitrarily fixed.

This paper generalizes the results obtained in Gonçalves, Jacob and Mendes-Lopes [11] in which a consistent decision procedure for an order one bilinear diagonal model is proposed and an associated rate of convergence is evaluated.

2 – General properties and hypotheses

Consider the simple bilinear model $X = (X_t,\, t \in \mathbb{Z})$ defined by

$$X_t = \varphi\, X_{t-k}\, \varepsilon_{t-l} + \varepsilon_t\,, \qquad k > 0, \quad l > 0\;, \tag{1}$$

where φ is a real number and $\varepsilon = (\varepsilon_t,\, t \in \mathbb{Z})$ is a strictly stationary and ergodic process such that $E|\log|\varepsilon_t|| < \infty$ and $E(\log|\varepsilon_t|) + \log|\varphi| < 0$. Under these conditions, we obtain

$$X_t \,=\, \varepsilon_t + \sum_{n=1}^{\infty} \varphi^n\, \varepsilon_{t-nk} \prod_{j=0}^{n-1} \varepsilon_{t-l-jk} \quad (a.s.)\,, \qquad t \in \mathbb{Z}\;.$$

Then, from Quinn [18] and Azencott and Dacunha-Castelle [1, pp. 30/32], there exists a strictly stationary and ergodic solution to (1). Assuming

$E|\log|X_t|| < \infty$, model (1) is invertible if $E(\log|X_t|)+\log|\varphi| < 0$ and we obtain

$$\varepsilon_t \;=\; X_t + \sum_{n=1}^{\infty} (-\varphi)^n X_{t-nl} \prod_{j=0}^{n-1} X_{t-k-jl} \quad (a.s.)\,, \qquad t\in\mathbb{Z}\;.$$

Denoting by $\underline{X}_t$ and $\underline{\varepsilon}_t$ the σ-fields generated by $(X_t, X_{t-1}, ...)$ and $(\varepsilon_t, \varepsilon_{t-1}, ...)$ respectively, we conclude that $\underline{X}_t = \underline{\varepsilon}_t$, in view of the two equalities above.

Hereafter we assume these general hypotheses concerning the stationarity, ergodicity, and invertibility of model (1). We also take $m=\min\{k,l\}$, $M=\max\{k,l\}$, and define the process $Y=(Y_t,\, t\in\mathbb{Z})$ by $Y_t = X_{t-k+m}\varepsilon_{t-l+m}$.

So,

$$(2)\qquad Y_t = X_{t-k+m}\left(X_{t-l+m} + \sum_{n=1}^{\infty}(-\varphi)^n X_{t-(n+1)l+m}\prod_{j=0}^{n-1} X_{t-k-(j+1)l+m}\right) \;\; (a.s.),$$

is also strictly stationary and ergodic. We note that $X_t = \varphi\, Y_{t-m} + \varepsilon_t$, according to (1) and (2). In the following and whenever necessary we will explicit the dependence of Y_t from the parameter φ by writing $Y_t(\varphi)$.

3 – A consistent decision procedure

3.1. The decision regions

Let us consider the hypotheses

$$H_0\colon\;\varphi=0 \quad\text{ against }\quad H_1\colon\;\varphi=\beta\;\;(\beta\neq 0\;\text{ fixed})\;.$$

We construct a decision procedure to distinguish between the models related with these hypotheses by establishing the asymptotic separation of two families of probability laws associated to them.

For each $t\in\mathbb{Z}$, we assume that the law of ε_t given the past $\underline{\varepsilon}_{t-m}$ has a unique zero median and we denote by F_t the distribution function of this conditional law.

Let us denote the model distribution and the corresponding expectation by P_φ and E_φ respectively, when the parameter of the model is equal to φ. Let g be a symmetrical weight-function defined on $\mathbb{R}$, which is strictly positive and non-decreasing on $\mathbb{R}^+$ and P_φ-integrable. We use T observations of the process

X denoted by $x_1, x_2, ..., x_T$, $T > M$, to construct the decision procedure. This construction is slightly different in the cases $\beta > 0$ and $\beta < 0$.

Defining the sets

$$D = \Big\{(u,v) \in \mathbb{R}^2\colon\ u > 0,\ v < \tfrac{\beta}{2}\,u\Big\} \cup \Big\{(u,v) \in \mathbb{R}^2\colon\ u < 0,\ v > \tfrac{\beta}{2}\,u\Big\}, \quad \beta > 0 ,$$

$$D' = \Big\{(u,v) \in \mathbb{R}^2\colon\ u < 0,\ v < \tfrac{\beta}{2}\,u\Big\} \cup \Big\{(u,v) \in \mathbb{R}^2\colon\ u > 0,\ v > \tfrac{\beta}{2}\,u\Big\}, \quad \beta < 0 ,$$

we consider the following regions

$$A_T \;=\; \left\{x^{(T)}\colon \sum_{t=M+1}^{T} g(\beta\, y_{t-m}) \Big[2\, \mathbb{1}_D(y_{t-m}, x_t) - 1\Big] \;\geq\; 0\right\}, \qquad \beta > 0 ,$$

$$A'_T \;=\; \left\{x^{(T)}\colon \sum_{t=M+1}^{T} g(\beta\, y_{t-m}) \Big[2\, \mathbb{1}_{D'}(y_{t-m}, x_t) - 1\Big] \;\geq\; 0\right\}, \qquad \beta < 0 ,$$

where $x^{(t)} = (..., x_{t-1}, x_t) \in \prod_{-\infty}^{t} \mathbb{R}$ and y_t denotes the particular value of Y_t.

The asymptotic separation of P_0 and P_β will then be established using these sequences of Borel sets. These sets are easily interpreted in an heuristic way as separation regions if we note that the conditional median of the law of X_t given $\underline{\varepsilon}_{t-m}$ is $\varphi\, Y_{t-m}$ and if we adapt the arguments outlined in Gonçalves and Mendes-Lopes [10] for AR models.

3.2. The consistence of the decision procedure

The following theorem allows us to conclude the consistence of the announced procedure.

Theorem 3.1. *Let $X = (X_t,\, t \in \mathbb{Z})$ be a real valued process satisfying model (1) with $(\varepsilon_t,\, t \in \mathbb{Z})$ a strictly stationary and ergodic process such that, for each $t \in \mathbb{Z}$, the median of the conditional law of ε_t given $\underline{\varepsilon}_{t-m}$ is unique and equal to zero. Under the hypotheses presented in section 2 on the model (1), there is a sequence of Borel sets ensuring the asymptotic separation of the sequences of the probability laws of the model defined by the hypotheses H_0 and H_1.*

Proof: Let us suppose $\beta > 0$. For each $t = M{+}1, ..., T$, let us take

$$\Psi(y_{t-m}, x_t) \;=\; \Psi_t \;=\; g(\beta\, y_{t-m}) \Big[2\, \mathbb{1}_D(y_{t-m}, x_t) - 1\Big] \quad \text{and} \quad \overline{\Psi}_T \;=\; \frac{1}{T} \sum_{t=M+1}^{T} \Psi_t \;.$$

The ergodic theorem allows us to conclude that

$$\lim_{T\to+\infty} \overline{\Psi}_T \;=\; E_\varphi(\Psi_{M+1}) \qquad (P_\varphi\text{-}a.s.)\ .$$

Let us now study the sign of this limit under each one of the hypotheses H_0 and H_1. Denoting $d = M - m$ and using properties of conditional expectation, we have

$$\begin{aligned} E_\varphi(\Psi_{M+1}) \;&=\; E_\varphi\Big[E_\varphi(\Psi_{M+1}/\underline{X}_{d+1})\Big] \\ &=\; E_\varphi\Big\{g(\beta\,Y_{d+1})\,\Big[2\,E_\varphi\Big(\mathbb{1}_D(Y_{d+1},X_{M+1})\,/\,\underline{X}_{d+1}\Big)-1\Big]\Big\} \end{aligned} \tag{3}$$

and

$$\begin{aligned} E_\varphi\Big(\mathbb{1}_D(Y_{d+1},X_{M+1})\,/\,\underline{X}_{d+1}\Big) \;&=\; E_\varphi\Big[\mathbb{1}_{\mathbb{R}^+}(Y_{d+1})\ \mathbb{1}_{]-\infty,\frac{\beta}{2}Y_{d+1}[}(X_{M+1}) \\ &\qquad +\ \mathbb{1}_{\mathbb{R}^-}(Y_{d+1})\ \mathbb{1}_{]\frac{\beta}{2}Y_{d+1},+\infty[}(X_{M+1})\,/\,\underline{X}_{d+1}\Big] \\ &=\; \mathbb{1}_{\mathbb{R}^+}(Y_{d+1})\ P_\varphi\Big(X_{M+1}<\tfrac{\beta}{2}\,Y_{d+1}\,/\,\underline{X}_{d+1}\Big) \\ &\qquad +\ \mathbb{1}_{\mathbb{R}^-}(Y_{d+1})\ P_\varphi\Big(X_{M+1}>\tfrac{\beta}{2}\,Y_{d+1}\,/\,\underline{X}_{d+1}\Big)\ . \end{aligned}$$

Thus, under H_0, the previous equality becomes

$$E_0\Big(\mathbb{1}_D(Y_{d+1},X_{M+1})\,/\,\underline{X}_{d+1}\Big) \;=\; \begin{cases} F_{M+1}(\frac{\beta}{2}\,Y^-_{d+1}), & \text{if }\ Y_{d+1}>0, \\ 1-F_{M+1}(\frac{\beta}{2}\,Y_{d+1}), & \text{if }\ Y_{d+1}<0, \\ 0, & \text{if }\ Y_{d+1}=0\ , \end{cases}$$

where $F_t(x^-) = P(\varepsilon_t < x/\underline{\varepsilon}_{t-m})$.

The nullity and the uniqueness of the conditional median and the fact that β is strictly positive imply

$$E_0\Big(\mathbb{1}_D(Y_{d+1},X_{M+1})\,/\,\underline{X}_{d+1}\Big) \;>\; \frac{1}{2} \qquad (P_0\text{-}a.s.)\ ,$$

as $Y_t \neq 0$, $P_\varphi - a.s.$, $\forall\, t\in\mathbb{Z}$. Then, as g is strictly positive,

$$g(\beta\,Y_{d+1})\,\Big[2\,E_0\Big(\mathbb{1}_D(Y_{d+1},X_{M+1})\,/\,\underline{X}_{d+1}\Big)-1\Big] \;>\; 0 \qquad (P_0\text{-}a.s.)\ .$$

Now, from (3) we have, under H_0,

$$\lim_{T\to+\infty}(a.s.)\,\overline{\Psi}_T \;>\; 0\ ,$$

which implies

$$\lim_{T\to+\infty} \mathbb{1}_{\{\overline{\Psi}_T \geq 0\}} \;=\; 1 \;.$$

Finally, the dominated convergence theorem gives us

$$\lim_{T\to+\infty} P_0(A_T) \;=\; 1 \;.$$

Let us now verify that, under H_1, $\lim\limits_{T\to\infty} P_\varphi(A_T) = 0$. In fact, under H_1 we obtain

$$\begin{aligned}
E_\beta\Big(\mathbb{1}_D(Y_{d+1}, X_{M+1}) \,/\, \underline{X}_{d+1}\Big) \;=\; & \mathbb{1}_{\mathbb{R}^+}(Y_{d+1})\; P_\beta\Big(\beta\, Y_{d+1} + \varepsilon_{M+1} < \tfrac{\beta}{2}\, Y_{d+1} \,/\, \underline{X}_{d+1}\Big) \\
& + \mathbb{1}_{\mathbb{R}^-}(Y_{d+1})\, P_\beta\Big(\beta Y_{d+1} + \varepsilon_{M+1} > \tfrac{\beta}{2} Y_{d+1} / \underline{X}_{d+1}\Big) \\
=\; & \begin{cases} F_{M+1}(-\frac{\beta}{2}\, Y_{d+1}^{-}), & \text{if } \; Y_{d+1} > 0, \\ 1 - F_{M+1}(-\frac{\beta}{2}\, Y_{d+1}), & \text{if } \; Y_{d+1} < 0, \\ 0, & \text{if } \; Y_{d+1} = 0 \;. \end{cases}
\end{aligned}$$

Then

$$E_\beta\Big[\mathbb{1}_D(Y_{d+1}, X_{M+1}) \,/\, \underline{X}_{d+1}\Big] \;<\; \frac{1}{2} \;,$$

taking into account the uniqueness and the nullity of the conditional median and the fact that β is greater than zero. Therefore

$$g(\beta\, Y_{d+1}) \,\Big[2\, E_\beta\Big(\mathbb{1}_D(Y_{d+1}, X_{M+1}) \,/\, \underline{X}_{d+1}\Big) - 1\Big] \;<\; 0 \qquad (P_\beta\text{-}a.s.) \;.$$

From this inequality, we deduce

$$\lim_{T\to+\infty} (a.s.)\, \overline{\Psi}_T \;<\; 0$$

and, finally,

$$\lim_{T\to+\infty} P_\beta(A_T) \;=\; 0 \;,$$

by the dominated convergence theorem.

In the case $\beta < 0$, we obtain an analogous result considering the set D' and the Borel sequence $(A'_T,\, T > M)$. ∎

3.3. Implementation of the decision procedure as a test

As mentioned in the introduction, the theoretical result established in section 3.2 can be used as a consistent test of the hypotheses

$$H_0\colon\ \varphi=0 \quad \text{against} \quad H_1\colon\ \varphi=\beta \ \ (\beta\neq 0,\ \text{fixed})\ ,$$

namely when the definition of the regions A_T does not involve the non-observable error process; this happens, in particular, when we take $Y_t(\varphi)$ equal to $Y_t(\beta)$ in these regions. We remark that, in this case, the asymptotic separation of P_0 and P_β follows in the same way. So, we obtain clearly a convergent test of the hypothesis H_0 against H_1: in the case $\beta>0$, for example, the noise hypothesis is accepted if $\sum_{t=M+1}^{T} g(\beta\, y_{t-m})\,[2\,\mathbb{1}_D(y_{t-m},x_t)-1] \geq 0$, and rejected otherwise.

However, if we replace in the same expression of $Y_t(\varphi)$ the parameter φ with an estimate $\hat\varphi_T$, as it is usual, the convergence of the procedure is not necessarily guaranteed. This is a point of future research. The following theorem gives a contribution to this question as it ensures the consistence of the corresponding estimator of $Y_t(\varphi)$ when $\hat\varphi_T$ is a consistent estimator of φ.

Theorem 3.2. *If φ belongs to a bounded set B, included in the stationarity region of X_t, and $\hat\varphi_T$ is a consistent estimator of φ, then*

$$\begin{aligned} Y_t(\hat\varphi_T) &= X_{t-k+m}(\hat\varphi_T)\;\varepsilon_{t-l+m} \\ &= \left[\sum_{i=1}^{+\infty} \hat\varphi_T^{\,i}\;\varepsilon_{t-(i+1)k+m}\ \prod_{j=0}^{i-1}\varepsilon_{t-l-(j+1)k+m} + \varepsilon_{t-k+m}\right]\varepsilon_{t-l+m} \end{aligned}$$

converges a.s. to $Y_t(\varphi)$ when T tends to $+\infty$.

Proof: Let us consider the sequence of functions $(f_{t,n}(\varphi),\ n\in\mathbb{N})$ defined by

$$f_{t,n}(\varphi) \;=\; \sum_{i=1}^{n} \varphi^i\,\varepsilon_{t-ik}\ \prod_{j=0}^{i-1}\varepsilon_{t-l-jk} + \varepsilon_t\ .$$

As stated in section 2, we have

$$\lim_{n\to+\infty} f_{t,n}(\varphi) \;=\; f_t(\varphi) \;=\; \sum_{i=1}^{+\infty} \varphi^i\,\varepsilon_{t-ik}\ \prod_{j=0}^{i-1}\varepsilon_{t-l-jk} + \varepsilon_t \qquad (a.s.)\ .$$

If this convergence is uniform, then $f_t(\varphi)$ is a continuous function, as $f_{t,n}(\varphi)$ is continuous, for each $n \in \mathbb{N}$. We have

$$\begin{aligned}\sup_{\varphi\in B} |f_{t,n}(\varphi) - f_t(\varphi)| \;&=\; \sup_{\varphi\in B} \left| \sum_{i=n+1}^{+\infty} \varphi^i\, \varepsilon_{t-ik} \prod_{j=0}^{i-1} \varepsilon_{t-l-jk} \right| \\ &\le\; \sup_{\varphi\in B} \sum_{i=n+1}^{+\infty} |\varphi^i|\, |\varepsilon_{t-ik}| \prod_{j=0}^{i-1} |\varepsilon_{t-l-jk}| \\ &\le\; \sum_{i=n+1}^{+\infty} a^i\, |\varepsilon_{t-ik}| \prod_{j=0}^{i-1} |\varepsilon_{t-l-jk}| \;,\end{aligned}$$

with $a = \max\{|\inf B|, |\sup B|\}$. This last expression converges to zero when n tends to $+\infty$, as it is the rest of an $a.s.$ convergent series to $\tilde{X}_t$ defined by

$$\tilde{X}_t \;=\; a\, \tilde{X}_{t-k}\, |\varepsilon_{t-l}| + |\varepsilon_t| \;,$$

as this model satisfies the stationarity condition $E(\log|\varepsilon_t|) + \log|a| < 0$.

So, if $\hat{\varphi}_T \to \varphi$ $(a.s.)$ then $Y_t(\hat{\varphi}_T) = f_{t-k+m}(\hat{\varphi}_T)\, \varepsilon_{t-l+m}$ converges $a.s.$ to $Y_t(\varphi)$ when T tends to $+\infty$. ∎

4 – Convergence rate of $P_0(A_T)$ in the diagonal case

Restricting ourselves to diagonal models, the convergence results concerning the asymptotic separation, presented in the previous section, may be completed by the knowledge of the convergence rate of $P_\varphi(A_T)$, $\varphi = 0$ and $\varphi = \beta$. In this section we study the convergence rate of $P_0(A_T)$.

So, let us consider now model (1) with $k = l = m = M$ and $d = 0$, under the general hypotheses of stationarity, ergodicity and invertibility presented in section 2.

In this section, we assume that the error process ε is such that

> $\forall\, t \in \mathbb{Z}$, $\varepsilon_t = \eta_{t-1}\, Z_t$, where
>
> η_{t-1} is a strictly positive and measurable function of $\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots$ with $0 < l_1 \le \eta_t \le l_2$;
>
> $(Z_t,\, t \in \mathbb{Z})$ is an independent and identically distributed sequence of real random variables with symmetrical distribution and unique zero median, and Z_t independent of $\underline{\varepsilon}_{t-1}$.

We remark that the form imposed here on the error process includes, in particular, conditionally heteroscedastic models like ARCH (Engle [6]), GARCH (Bollerslev [4]) or GTARCH (Gonçalves and Mendes-Lopes [10]) models.

We also point out that the law of ε_t given $\underline{\varepsilon}_{t-1}$ is symmetrical; moreover, the influence of $\underline{\varepsilon}_{t-1}$ on ε_t, specified by this formulation, is an influence on the variance of ε_t, when it exists.

Let us denote by F the distribution function of a random variable Z identically distributed with Z_t, $t \in \mathbb{Z}$.

The following result establishes an exponential rate for the convergence to one of $P_0(A_T)$, $(T \to +\infty)$, using recursively an inequality of Hoeffding [13] and an inequality of moments (Martins [14]). These two results, called inequality A and B respectively, are summarized in the appendix.

Theorem 4.1. *Under the previous conditions on the model and supposing the weight-function g defined by*

$$g(x) \,=\, 2\,F\Big(\Big(\tfrac{|x|}{2\,l_2}\Big)^-\,\Big) - 1 \,=\, 2\,P\Big(Z < \tfrac{|x|}{2\,l_2}\Big) - 1\,, \qquad x \in \mathbb{R}\,,$$

we have

$$P_0(\overline{A}_T) \,\leq\, \left\{ E_0\left[\exp\left[-\tfrac{1}{2}\left(2\,F\Big(\Big(\tfrac{|\beta\, l_1^2\, Z^2|}{2\,l_2}\Big)^-\,\Big) - 1\right)^2\,\right]\right]\right\}^{T-k}, \qquad T > k\;.$$

Proof: As the proof is long and technical we present it in the appendix to improve the readability of the paper.

A corresponding upper bound for the probability of A_T when $\varphi = \beta$ is in study. The underlying difficulties come, certainly, from the fact that the model under the alternative is much more complex than under the null. Nevertheless, we note that the simulation study presented in the next section allows us to conjecture the same kind of behaviour for $P_\beta(A_T)$.

5 – Application to the statistical study of bilinear models

The asymptotic separation of two probabilistic models is particularly interesting when it can be used as a statistical decision rule, either in estimation or in test theory.

The aim of this section is to illustrate the suitability and usefulness of this decision procedure within the scope of tests for simple bilinear models.

So, a simulation study is done considering a real valued process $X = (X_t,\, t\in\mathbb{Z})$ following a bilinear model

$$X_t \;=\; \varphi\, X_{t-1}\, \varepsilon_{t-1} + \varepsilon_t$$

where $(\varepsilon_t,\, t\in\mathbb{Z})$ is a sequence of i.i.d. random variables following the standard Cauchy distribution, for all t, and supposing $\varphi \in\,]0, 0.53[$.

Under these conditions, this model is strictly stationary, ergodic and invertible.

In fact, we can prove that $(\varepsilon_t,\, t\in\mathbb{Z})$ is a strictly stationary and ergodic error process; moreover, $E(\log|\varepsilon_t|) \le \frac{2}{\pi}$ and $E(\log|X_t|) \le \frac{2}{\pi}$. So, if $|\varphi| < 0.53$, the quantities $E(\log|\varepsilon_t|) + \log|\varphi|$ and $E(\log|X_t|) + \log|\varphi|$ are simultaneously negative.

Furthermore, it is obvious that the error process has the form considered in section 4 with $l_1 = l_2 = 1$.

The decision procedure will be used to test the hypotheses

$$H_0\colon\; \varphi = 0 \qquad \text{against} \qquad H_1\colon\; \varphi = \beta, \quad (\beta > 0, \text{ fixed}) \;.$$

In this case, the separation set

$$A_T \;=\; \left\{ x^{(T)}\colon\; \sum_{t=2}^{T} g(\beta\, y_{t-1}) \left[2\, \mathbb{1}_D(y_{t-1}, x_t) - 1 \right] \;\ge\; 0 \right\}, \qquad T > 1 \;,$$

is the acceptance region of the test. The values of Y_t were taken as

$$\begin{cases} \hat{y}_t \;=\; x_t \left(\displaystyle\sum_{k=1}^{t-1} (-\varphi)^k\, x_{t-k}^2 \prod_{i=1}^{k-1} x_{t-i} + x_t \right), \qquad t = 2, ..., T \;, \\ \hat{y}_1 \;=\; x_1^2 \;; \end{cases}$$

this is the same as if we took the observations before time 1 equal to zero. The first proposed value $\hat{y}_1$ is obtained using the definition of y_1, $y_1 = x_1\, \varepsilon_1$, and the model equation $\varepsilon_t = -\varphi\, x_{t-1}\, \varepsilon_{t-1} + x_t$ with $x_0 = 0$.

To evaluate the importance of the weight-function present in the test statistics, we take firstly the observations equally weighted ($g = g_1 = 1$); afterwards, we consider the weight-function used to establish the rate of convergence of this test level sequence, which takes here the form $g(x) = g_2(x) = \frac{2}{\pi}\, \mathrm{arctg}(\frac{|x|}{2})$. These two functions are symmetrical, non-decreasing on $\mathbb{R}^+$ and P_φ-integrable. Moreover, $g(\beta Y_t) > 0$, $\forall\, t\in\mathbb{Z}$, P_φ-*a.s.*, as ε is absolutely continuous.

To evaluate the behaviour of our test when H_0 is true, a simulation study is done, for $T = 20$ and $T = 50$, taking $\varphi = 0$ in the model. The size of the test is estimated, testing this model against six alternatives ($\beta = 0.01$, $\beta = 0.05$,

$\beta = 0.1$, $\beta = 0.2$, $\beta = 0.3$ and $\beta = 0.5$). For each one of these alternatives, we calculate the proportion of rejections of H_0 in 60 replications of the model; in table 1, we record the 95% confidence regions, corresponding to samples of size 30, for this proportion.

T	20	20	50	50
β	g_1	g_2	g_1	g_2
0.01	$]0.39, 0.44[$	$]0.19, 0.23[$	$]0.31, 0.35[$	$]0.06, 0.08[$
0.05	$]0.26, 0.31[$	$]0.05, 0.08[$	$]0.16, 0.19[$	$]0.003, 0.009[$
0.1	$]0.20, 0.24[$	$]0.03, 0.04[$	$]0.08, 0.11[$	$]0.0, 0.002[$
0.2	$]0.12, 0.16[$	$]0.005, 0.012[$	$]0.03, 0.05[$	0.0
0.3	$]0.09, 0.12[$	$]0.001, 0.005[$	$]0.014, 0.03[$	0.0
0.5	$]0.05, 0.07[$	0.0	$]0.003, 0.01[$	0.0

Table 1 – Proportion of rejections of H_0 when $\varphi = 0$.

We note that our test performs very well, even for quite small values of β and T ($\beta = 0.05$ and $T = 20$), when the observations are weighted according to g_2. When T increases ($T = 50$), the behaviour of the test is strongly improved even when the observations are not weighted (for $\beta \geq 0.1$ the estimation of the size is very good in both cases). Nevertheless, we must point out the significant influence of the weight function in the performance of the test.

In order to have an idea of the rate of convergence of the power of this test, we consider, for $T = 20$ and $T = 50$, the same values of β ($\beta = 0.01$, $\beta = 0.05$, $\beta = 0.1$, $\beta = 0.2$, $\beta = 0.3$ and $\beta = 0.5$) and we generate the bilinear models corresponding to $\varphi = \beta$; in each case, we record the proportion of rejections of the alternative hypothesis (true in this case) in 60 replications of the model. The 95% confidence regions for this proportion, obtained with samples of size 30, are presented in table 2.

T	20	20	50	50
β	g_1	g_2	g_1	g_2
0.01	$]0.36, 0.41[$	$]0.16, 0.19[$	$]0.31, 0.36[$	$]0.06, 0.07[$
0.05	$]0.24, 0.28[$	$]0.06, 0.08[$	$]0.15, 0.18[$	$]0.008, 0.02[$
0.1	$]0.15, 0.18[$	$]0.03, 0.05[$	$]0.05, 0.08[$	$]0.0, 0.004[$
0.2	$]0.07, 0.09[$	$]0.01, 0.08[$	$]0.01, 0.02[$	0.0
0.3	$]0.04, 0.06[$	$]0.005, 0.02[$	$]0.002, 0.008[$	0.0
0.5	$]0.01, 0.03[$	$]0.0, 0.002[$	0.0	0.0

Table 2 – Proportion of rejections of H_1 when $\varphi = \beta$.

Taking into account the similarity between the results presented in this table and in the previous one, we point out the symmetrical role given by our procedure to H_0 and H_1; in fact, the empirical behaviour of the two kinds of test errors is obviously the same.

Moreover, this analogous behaviour emphasizes the importance of the weight-function in the performance of the test and allows us to conjecture the same rate of convergence for its power function.

APPENDIX

I. Inequality A

Proposition (Hoeffding [13] inequality (4.16)). *Let $(\Omega, \mathcal{A}, \mu)$ be a probability space, U a real valued random variable on $(\Omega, \mathcal{A}, \mu)$ taking its values on the interval $[a, b]$ and $\mathcal{B}$ a sub-σ-field of $\mathcal{A}$ with a regular version $\mu_U^{\mathcal{B}}$ of the conditional law of U given $\mathcal{B}$. Then*

$$E_\mu^{\mathcal{B}}\left(e^{U - E_\mu^{\mathcal{B}}(U)}\right) \leq e^{\frac{1}{8}(b-a)^2} \qquad (\mu\text{-a.s.}). \ \blacksquare$$

II. Inequality B

Proposition (Martins [14]). *Let Y be a real valued random variable with a symmetrical distribution, P_Y. Consider a symmetrical function $g : \mathbb{R} \to \mathbb{R}_0^+$ that is non-increasing on $\mathbb{R}^+$ and such that $g(Y)$ is a r.v. with $E(g(Y)) < +\infty$. Let a, b and ν be real numbers.*

a) *Consider the function h defined by $h = a\, \mathbb{1}_{]-\infty,\nu[} + b\, \mathbb{1}_{[\nu,+\infty[}$, and suppose that one of the following cases occurs:*

- **i**) *$\nu \leq 0$ and $a \geq b$;*
- **ii**) *$P_Y(\{0\}) = 0$ and $\nu \geq 0$ and $a \leq b$;*
- **iii**) *$P_Y(\{0\}) \neq 0$ and $\nu > 0$ and $a \leq b$.*

Then

$$E[g(Y)\, h(Y)] \leq E[g(Y)]\, E[h(Y)] \ . \tag{4}$$

b) *Suppose now that h is defined as $h = a\,\mathbb{1}_{]-\infty,\nu]} + b\,\mathbb{1}_{]\nu,+\infty[}$. Then (4) is still true if one of the following situations is verified:*

i) *$\nu \geq 0$ and $a \leq b$;*

ii) *$P_Y(\{0\}) = 0$ and $\nu \leq 0$ and $a \geq b$;*

iii) *$P_Y(\{0\}) \neq 0$ and $\nu < 0$ and $a \geq b$.* ∎

This result is a generalization of a result of Massé and Viano [15] and its proof is based on the following lemma.

Lemma. *Let U and V be real valued random variables such that $E(U) < +\infty$ and V is bounded. Then*

$$E(UV) - E(U)\,E(V) \;=\; \int_{\mathbb{R}^2} \Big(F_{U,V}(u,v) - F_U(u)\,F_V(v)\Big)\,du\,dv \;,$$

where $F_{U,V}$, F_U, F_V denote the distribution functions of (U,V) and its margins respectively. ∎

The proof of this lemma is analogous to that of a similar result of Hoeffding (Suquet [20]), for U and V verifying $E(U^2) < +\infty$ and $E(V^2) < +\infty$.

III. Proof of Theorem 4.1

Only the case $\beta > 0$ is presented here, as the study for $\beta < 0$ is analogous. From the definition of A_T, we have

$$P_0(\overline{A}_T) \;=\; P_0\left[-\sum_{t=k+1}^{T} \Psi_t > 0\right] \;=\; P_0\left[\exp\left(-\sum_{t=k+1}^{T} \Psi_t\right) > 1\right] .$$

Then

$$\begin{aligned}
P_0(\overline{A}_T) \;&\leq\; P_0\left[\exp\left(-\sum_{t=k+1}^{T} \Psi_t\right) \geq 1\right] \\
&\leq\; E_0\left[\exp\left(-\sum_{t=k+1}^{T} \Psi_t\right)\right], \qquad \text{using Markov's inequality} \\
&=\; E_0\left\{E_0\left[\exp\left(-\sum_{t=k+1}^{T} \Psi_t\right) / \,\underline{X}_{T-1}\right]\right\} \\
&=\; E_0\left\{\exp\left(-\sum_{t=k+1}^{T-1} \Psi_t\right) E_0\Big[\exp(-\Psi_T) / \,\underline{X}_{T-1}\Big]\right\} . \qquad (5)
\end{aligned}$$

To prove the inequality stated in the theorem, we need to consider separately the cases $T \geq 2k+1$ and $k+1 \leq T \leq 2k$, but in fact, only the first case is important in terms of asymptotic studies. Let us then suppose $T \geq 2k+1$. Moreover, as the proof is quite long, we consider several steps in order to improve its understanding.

Step A. In this part, we obtain an upper bound for $E_0[\exp(-\Psi_T) \,/\, \underline{X}_{T-1}]$.

Applying the Hoeffding inequality with $U = -\Psi_T = -\Psi(Y_{T-k}, X_T)$ and $\mathcal{B} = \underline{X}_{T-1}$, we obtain

$$E_0\Big\{e^{-\Psi(Y_{T-k},X_T)+E_0[\Psi(Y_{T-k},X_T)\,/\,\underline{X}_{T-1}]} \,/\, \underline{X}_{T-1}\Big\} \leq e^{1/2g^2(\beta Y_{T-k})}, \qquad P_0\text{-}a.s.\ ,$$

which is equivalent to

$$e^{E_0[\Psi(Y_{T-k},X_T)\,/\,\underline{X}_{T-1}]}\, E_0\Big\{e^{-\Psi(Y_{T-k},X_T)} \,/\, \underline{X}_{T-1}\Big\} \leq e^{1/2g^2(\beta Y_{T-k})}, \quad P_0\text{-}a.s.\ . \tag{6}$$

On the other hand, we have

$$E_0\Big[\Psi(Y_{T-k}, X_T) \,/\, \underline{X}_{T-1}\Big] = g(\beta\, Y_{T-k}) \Big[2\, E_0\Big(\mathbb{1}_D(Y_{T-k}, X_T) \,/\, \underline{X}_{T-1}\Big) - 1\Big] \tag{7}$$

with

$$E_0\Big(\mathbb{1}_D(Y_{T-k}, X_T) \,/\, \underline{X}_{T-1}\Big) = F\Big(\Big(\tfrac{\beta\,|Y_{T-k}|}{2\,\eta_{T-1}}\Big)^{-}\Big), \qquad P_0\text{-}a.s.\ . \tag{8}$$

In fact, as in section 3,

$$E_0\Big(\mathbb{1}_D(Y_{T-k}, X_T) \,/\, \underline{X}_{T-1}\Big) = \begin{cases} P_0\Big(X_T < \frac{\beta}{2}\, Y_{T-k} \,/\, \underline{X}_{T-1}\Big), & Y_{T-k} > 0, \\ P_0\Big(X_T > \frac{\beta}{2}\, Y_{T-k} \,/\, \underline{X}_{T-1}\Big), & Y_{T-k} < 0, \\ 0, & Y_{T-k} = 0\ . \end{cases}$$

But, under H_0, $X_T = \varepsilon_T$; so, using the definition of ε_t and the symmetrical property of the common distribution of $(Z_t,\, t \in \mathbb{Z})$, we obtain (8).

From (7) and (8), we have

$$E_0\Big[\Psi(Y_{T-k}, X_T) \,/\, \underline{X}_{T-1}\Big] = g(\beta\, Y_{T-k}) \Big[2\, F\Big(\Big(\tfrac{\beta\,|Y_{T-k}|}{2\,\eta_{T-1}}\Big)^{-}\Big) - 1\Big], \qquad P_0\text{-}a.s.\ ,$$

which allows us to conclude that inequality (6) is equivalent to

$$\begin{aligned} E_0\Big\{e^{-\Psi(Y_{T-k},X_T)} \,/\, \underline{X}_{T-1}\Big\} &\leq \\ &\leq \exp\Big\{\tfrac{1}{2}\, g^2(\beta\, Y_{T-k}) - g(\beta\, Y_{T-k}) \Big[2\, F\Big(\Big(\tfrac{\beta\,|Y_{T-k}|}{2\,\eta_{T-1}}\Big)^{-}\Big) - 1\Big]\Big\}, \qquad P_0\text{-}a.s.\ , \\ &\leq \exp\Big\{\tfrac{1}{2}\, g^2(\beta\, Y_{T-k}) - g(\beta\, Y_{T-k}) \Big[2\, F\Big(\Big(\tfrac{\beta\,|Y_{T-k}|}{2\, l_2}\Big)^{-}\Big) - 1\Big]\Big\}, \qquad P_0\text{-}a.s.\ , \end{aligned}$$

in view of the condition $\eta_t \le l_2$, $t \in \mathbb{Z}$. The minimum of this upper bound is obtained considering the weight-function $g(x) = 2\,F((\frac{|x|}{2l_2})^-) - 1$, which is symmetrical, strictly positive and non-increasing on $\mathbb{R}^+$ and P_φ-integrable. Let us also remark that $g(x)$ is the distance in variation between the Gaussian laws $N(0,l_2)$ and $N(x,l_2)$. Using this definition of g, we obtain

$$E_0\Big\{e^{-\Psi(Y_{T-k},X_T)} \,/\, \underline{X}_{T-1}\Big\} \;\le\; \exp\Big(-\tfrac{1}{2}\, g^2(\beta\, Y_{T-k})\Big)\,, \qquad P_0\text{-}a.s.\ . \tag{9}$$

Step B. After inserting the result just obtained in (5), we use $k-1$ times a technique analogous to the one used in step A.

Inequality (9) together with (5) gives us

$$P_0(\overline{A}_T) \;\le\; E_0\left[\exp\Big(-\sum_{t=k+1}^{T-1} \Psi_t\Big)\, \exp\Big(-\tfrac{1}{2}\, g^2(\beta\, Y_{T-k})\Big)\right]\,. \tag{10}$$

To facilitate notation, let us define

$$S_n = \exp\left(-\sum_{t=k+1}^{n} \Psi_t\right), \quad n \ge k+1 \quad \text{and} \quad G(x) = \exp\Big(-\tfrac{1}{2}\, g^2(\beta\, x)\Big), \quad x \in \mathbb{R}\ .$$

If $k \ge 2$, the expectation in (10) is then equal to

$$E_0\Big\{S_{T-2}\, G(Y_{T-k})\, E_0\Big[\exp(-\Psi_{T-1}) \,/\, \underline{X}_{T-2}\Big]\Big\}\ .$$

Applying the Hoeffding inequality again and repeating the procedure, we have, if $k \ge 3$,

$$P_0(\overline{A}_T) \;\le\; E_0\Big\{S_{T-3}\, G(Y_{T-k})\, G(Y_{T-k-1})\, E_0\Big[\exp(-\Psi_{T-2}) \,/\, \underline{X}_{T-3}\Big]\Big\}\ .$$

Using the same technique $k-3$ times more, which corresponds to $k-1$ applications of the Hoeffding inequality as in step A, we arrive at

$$P_0(\overline{A}_T) \;\le\; E_0\left\{S_{T-k} \prod_{i=0}^{k-2} G(Y_{T-k-i})\, E_0\Big[\exp(-\Psi_{T-(k-1)}) \,/\, \underline{X}_{T-k}\Big]\right\}\ . \tag{11}$$

Applying the Hoeffding inequality once more, now with $U = -\Psi_{T-(k-1)}$ and $\mathcal{B} = \underline{X}_{T-k}$, we obtain

$$E_0\{e^{-\Psi_{T-(k-1)}} \,/\, \underline{X}_{T-k}\} \;\le\; G(Y_{T-2k+1})\,, \qquad P_0\text{-}a.s.\ .$$

This inequality and (11) give us the following upper bound for $P_0(\overline{A}_T)$:

$$P_0(\overline{A}_T) \;\leq\; E_0\left\{ S_{T-k} \prod_{i=0}^{k-1} G(Y_{T-k-i}) \right\} .$$

This is equivalent to

$$P_0(\overline{A}_T) \;\leq$$

$$\leq\; E_0\left\{ S_{T-k-1} \prod_{i=1}^{k-1} G(Y_{T-k-i})\, E_0\Big[G(Y_{T-k}) \exp(-\Psi_{T-k}) \,/\, \underline{X}_{T-k-1} \Big] \right\} . \tag{12}$$

Note that (12) is still true when $T = 2\,k+1$, stipulating $S_k = 1$.

Step C. Let us now look at $E_0[G(Y_{T-k}) \exp(-\Psi_{T-k}) \,/\, \underline{X}_{T-k-1}]$.

In order to find an upper bound for

$$E_0\Big[G(Y_{T-k}) \exp(-\Psi_{T-k}) \,/\, \underline{X}_{T-k-1} \Big] \;=$$

$$=\; E_0\left\{ \exp\Big[-\tfrac{1}{2}\, g^2(\beta\, Y_{T-k}) \Big] \exp\Big[-\Psi(Y_{T-2k}, X_{T-k}) \Big] \,/\, \underline{X}_{T-k-1} \right\} \tag{13}$$

using inequality B, we consider the functions $\bar{g}$ and $\bar{h}_y$ defined by

- $\bar{g}(x) = G(x^2) = \exp[-\frac{1}{2}\, g^2(\beta x^2)]$, $x \in \mathbb{R}$,
- $\bar{h}_y(x) = \exp[-\Psi(y,x)]$, $x \in \mathbb{R}$, with $y \neq 0$ a fixed real number.

These functions satisfy the hypotheses of inequality B. In fact, taking into account the definition of g, it is simple to verify that $\bar{g}$ is strictly positive, symmetrical and decreasing on $\mathbb{R}^+$. Concerning $\bar{h}_y$ we have, for a fixed $y \in \mathbb{R}$,

$$\begin{aligned} \bar{h}_y(x) \;&=\; \exp\Big\{ -g(\beta\, y)\, [2\, \mathbb{1}_D(y,x) - 1] \Big\} \\ &=\; \begin{cases} \exp[-g(\beta y)], & (y,x) \in D \\ \exp[g(\beta y)], & \text{otherwise} \end{cases} \\ &=\; \begin{cases} \exp[-g(\beta y)], & xy < 0 \,\vee\, |x| < \frac{\beta}{2}\,|y| \\ \exp[g(\beta y)], & xy \geq 0 \,\wedge\, |x| \geq \frac{\beta}{2}\,|y| \;. \end{cases} \end{aligned}$$

If $y > 0$,

$$\begin{aligned} \bar{h}_y(x) \;&=\; \begin{cases} \exp[-g(\beta y)], & x < 0 \,\vee\, |x| < \frac{\beta}{2}\, y \\ \exp[g(\beta y)], & x \geq 0 \,\wedge\, |x| \geq \frac{\beta}{2}\, y \end{cases} \\ &=\; e^{-g(\beta y)}\, \mathbb{1}_{]-\infty, \frac{\beta}{2} y[}(x) \;+\; e^{g(\beta y)}\, \mathbb{1}_{[\frac{\beta}{2} y, +\infty[}(x) \;. \end{aligned}$$

If $y < 0$,

$$\bar{h}_y(x) \;=\; \begin{cases} \exp[-g(\beta y)], & x > 0 \vee |x| < -\frac{\beta}{2}\, y \\ \exp[g(\beta y)], & x \leq 0 \wedge x \leq \frac{\beta}{2}\, y \end{cases}$$

$$= e^{g(\beta y)}\, \mathbb{1}_{]-\infty,\frac{\beta}{2}y]}(x) \,+\, e^{-g(\beta y)}\, \mathbb{1}_{]\frac{\beta}{2}y,+\infty[}(x)\;.$$

So, if $y > 0$, inequality B applies, with $\nu = \frac{\beta}{2}\, y > 0$ and $a = e^{-g(\beta y)} < b = e^{g(\beta y)}$; if $y < 0$, inequality B applies, with $\nu = \frac{\beta}{2}\, y < 0$ and $a = e^{g(\beta y)} > b = e^{-g(\beta y)}$.

Let us now go back to the expectation in (13). Under H_0, that expectation is equal to

$$E_0\Big[\bar{g}(\varepsilon_{T-k})\, \bar{h}_{Y_{T-2k}}(\varepsilon_{T-k}) \,/\, \underline{\varepsilon}_{T-k-1}\Big]$$

which, from inequality B, is less than or equal to

$$E_0\Big[G(\varepsilon^2_{T-k}) \,/\, \underline{\varepsilon}_{T-k-1}\Big]\, E_0\Big[\exp(-\Psi_{T-k}) \,/\, \underline{\varepsilon}_{T-k-1}\Big]\;, \tag{14}$$

as $\Psi(Y_{T-2k}, \varepsilon_{T-k}) = \Psi_{T-k}$, under H_0.

On the other hand, under H_0, $G(Y_{T-k-i}) = G(\varepsilon^2_{T-k-i})$, $i \in \mathbb{N}$. Then, (12) and (14) allow us to obtain the following upper bound for $P_0(\overline{A}_T)$:

$$E_0\left[S_{T-k-1} \prod_{i=1}^{k-1} G(\varepsilon^2_{T-k-i})\, E_0\Big(G(\varepsilon^2_{T-k}) \,/\, \underline{\varepsilon}_{T-k-1}\Big)\, E_0\Big(e^{-\Psi_{T-k}} \,/\, \underline{\varepsilon}_{T-k-1}\Big)\right]\;. \tag{15}$$

Step D. Let us now concentrate on $E_0[G(\varepsilon^2_{T-k}) \,/\, \underline{\varepsilon}_{T-k-1}]$.

Using the definition of g and the fact that $\varepsilon_{T-k} = \eta_{T-k-1}\, Z_{T-k}$, we obtain

$$E_0\Big[G(\varepsilon^2_{T-k}) \,/\, \underline{\varepsilon}_{T-k-1}\Big] = E_0\left\{\exp\left[-\tfrac{1}{2}\left(2\,F\Big(\Big(\tfrac{\beta\,\eta^2_{T-k-1}\, Z^2_{T-k}}{2\, l_2}\Big)^{-}\Big) - 1\right)^2\right] \Big/\, \underline{\varepsilon}_{T-k-1}\right\}\;.$$

This expectation is less than or equal to

$$E_0\left\{\exp\left[-\tfrac{1}{2}\left(2\,F\Big(\Big(\tfrac{\beta\, l_1^2\, Z^2_{T-k}}{2\, l_2}\Big)^{-}\Big) - 1\right)^2\right] \Big/\, \underline{\varepsilon}_{T-k-1}\right\} \;=\; E_0\Big(G(l_1^2\, Z^2)\Big)\;,$$

as $\eta_{T-k-1} \geq l_1 > 0$ and Z_{T-k} is independent of $\underline{\varepsilon}_{T-k-1}$ and identically distributed with Z. From this and (15), we obtain

$$P_0(\overline{A}_T) \;\leq\; E_0\left[S_{T-k-1} \prod_{i=1}^{k-1} G(\varepsilon^2_{T-k-i})\, E_0\Big(G(l_1^2\, Z^2)\Big)\, E_0\Big(e^{-\Psi_{T-k}} \,/\, \underline{\varepsilon}_{T-k-1}\Big)\right]\;.$$

Note that $E_0(G(l_1^2\,Z^2))$ is a deterministic value. Therefore, the previous expectation is equal to

$$E_0\Big(G(l_1^2\,Z^2)\Big)\,E_0\left[S_{T-k-1}\,\prod_{i=1}^{k-1} G(\varepsilon^2_{T-k-i})\,E_0\Big(e^{-\Psi_{T-k}}\,/\,\underline{X}_{T-k-1}\Big)\right]. \tag{16}$$

Step E. Next, we firstly find an upper bound for $E_0(e^{-\Psi_{T-k}}\,/\,\underline{X}_{T-k-1})$, then we use it in (16) and obtain a new upper bound for $P_0(\overline{A}_T)$. We repeat this procedure with $E_0(e^{-\Psi_{T-k-1}}\,/\,\underline{X}_{T-k-2})$.

The Hoeffding inequality leads us to

$$E_0\Big(e^{-\Psi_{T-k}}\,/\,\underline{X}_{T-k-1}\Big)\;\le\;G(Y_{T-2k})\,,\qquad P_0\text{-}a.s.\ .$$

This inequality and (16) allow us to arrive at

$$P_0(\overline{A}_T)\;\le\;E_0\Big(G(l_1^2\,Z^2)\Big)\,E_0\left(S_{T-k-1}\,\prod_{i=1}^{k} G(\varepsilon^2_{T-k-i})\right), \tag{17}$$

which is equivalent to

$$\begin{aligned} P_0(\overline{A}_T)\;\le\;&E_0\Big(G(l_1^2\,Z^2)\Big)\;\times\\ &\times\;E_0\left\{S_{T-k-2}\,\prod_{i=2}^{k} G(\varepsilon^2_{T-k-i})\,E_0\Big[G(\varepsilon^2_{T-k-1})\,e^{-\Psi_{T-k-1}}\,/\,\underline{X}_{T-k-2}\Big]\right\}. \end{aligned} \tag{18}$$

Inequality B applied to the conditional expectation gives

$$\begin{aligned} &P_0(\overline{A}_T)\;\le\\ &\le\;\Big[E_0\Big(G(l_1^2\,Z^2)\Big)\Big]^2\,E_0\left[S_{T-k-2}\,\prod_{i=2}^{k} G(\varepsilon^2_{T-k-i})\,E_0\Big(e^{-\Psi_{T-k-1}}\,/\,\underline{X}_{T-k-2}\Big)\right] \end{aligned} \tag{19}$$

and, from the Hoeffding inequality, we have

$$E_0\Big(e^{-\Psi_{T-k-1}}\,/\,\underline{X}_{T-k-2}\Big)\;\le\;G(Y_{T-2k-1})\,,\qquad P_0\text{-}a.s.\ . \tag{20}$$

From (19) and (20), we deduce

$$P_0(\overline{A}_T)\;\le\;\Big[E_0\Big(G(l_1^2\,Z^2)\Big)\Big]^2\,E_0\left(S_{T-k-2}\,\prod_{i=2}^{k+1} G(\varepsilon^2_{T-k-i})\right). \tag{21}$$

Step F. Finally, we note that this inequality is of the same type as (17). Repeating the procedure $T-2\,k-1$ times, we have

$$P_0(\overline{A}_T) \;\leq\; \Big[E_0\Big(G(l_1^2\,Z^2)\Big)\Big]^{T-2k-1} E_0\left(S_{k+1}\prod_{i=T-2k-1}^{T-k-2} G(\varepsilon^2_{T-k-i})\right) . \tag{22}$$

A final application of inequality B and the Hoeffding inequality gives

$$P_0(\overline{A}_T) \;\leq\; \Big[E_0\Big(G(l_1^2\,Z^2)\Big)\Big]^{T-2k} E_0\left(\prod_{i=T-2k}^{T-k-1} G(\varepsilon^2_{T-k-i})\right) . \tag{23}$$

It is now easy to verify that

$$E_0\left(\prod_{i=T-2k}^{T-k-1} G(\varepsilon^2_{T-k-i})\right) \;\leq\; \Big[E_0\Big(G(l_1^2\,Z^2)\Big)\Big]^{k} ,$$

in view of the form of ε_t, the bounds of η_t and the independence of $(Z_t,\, t\in\mathbb{Z})$.

These two inequalities give us the upper bound for $P_0(\overline{A}_T)$ stated in the theorem.

To finalize, we point out that, if $k+1\leq T\leq 2\,k$, i.e., $T=2\,k-j$, $0\leq j\leq k-1$, we obtain the same result by using the Hoeffding inequality $k-j$ times. ∎

ACKNOWLEDGEMENT – We are grateful to the referee for helpful comments and suggestions. This work was partially supported by the Praxis/2/2.1/Mat/400/94 Project.

REFERENCES

[**1**] Azencott, R. and Dacunha-Castelle, D. – *Séries d'Observations Irrégulières, Modélisation et Prévision,* Masson, Paris, 1984.

[**2**] Benghabrit, Y. and Hallin, M. – Optimal rank-based tests against first-order superdiagonal bilinear dependence, *J. Statist. Plann. Inference,* 32 (1992), 45–61.

[**3**] Benghabrit, Y. and Hallin, M. – Rank-based tests for autoregressive against bilinear serial dependence, *Nonparametric Statist.,* 6 (1996), 253–272.

[**4**] Bollerslev, T. – Generalized autoregressive conditional heteroscedasticity, *Journal of Econometrics,* 31 (1986), 307–327.

[**5**] Diebolt, J. and Ngatchou Wandji, J. – A nonparametric test for generalized first-order autoregressive models, *Scand. J. Statist.,* 24 (1997), 241–259.

[6] Engle, R.F. – Autoregressive Conditional Heteroscedasticity with estimates of the variance of the UK inflation, *Econometrica,* 50 (1982), 987–1008.

[7] Geffroy, J. – Inégalités pour le niveau de signification et la puissance de certains tests reposant sur des données quelconques, *C. R. Acad. Sci. Paris, Ser. A,* 282 (1976), 1299–1301.

[8] Geffroy, J. – *Asymptotic separation of distributions and convergence properties of tests and estimators*, in: "Asymptotic Theory of Statistical Tests and Estimation" (I.M. Chakravarti, Ed.), Academic Press, 1980, 159–177.

[9] Gonçalves, E.; Jacob, P. and Mendes-Lopes, N. – A new test for arma models with a general white noise process, *Test,* 5(1) (1996), 187–202.

[10] Gonçalves, E. and Mendes-Lopes, N. – Some statistical results on autoregressive conditionally heteroscedastic models, *Journal of Statistical Planning and Inference,* 68 (1998), 193–202.

[11] Gonçalves, E.; Jacob, P. and Mendes-Lopes, N. – A decision procedure for bilinear time series based on the asymptotic separation, *Statistics,* 33 (2000), 333–348.

[12] Granger, C.W.J. and Andersen, A. – *An Introduction to Bilinear Time Series Models*, Vandenhoeck and Ruprecht, Göttingen, 1978.

[13] Hoeffding, W. – Probability inequalities for sums of bounded random variables, *J. Amer. Stat. Assoc.,* 58(1) (1963), 13–30.

[14] Martins, C.M. – *Moment inequalities for some functions of symmetrical distributions*, Actas do VII Congresso Anual da Sociedade Portuguesa de Estatística, 2000, in press.

[15] Massé, B. and Viano, M.C. – Explicit and exponential bounds for a test on the coefficient of an ar(1) model, *Stat. Prob. Letters,* 25(4) (1995), 365–371.

[16] Moché, R. – Quelques tests relatifs à la detection d'un signal dans un bruit Gaussian, *Publications de l'IRMA,* 19(III) (1989), Univ. Lille, France.

[17] Pham, T.D. and Tran, L.T. – On the first order bilinear time series model, *J. Appl. Prob.,* 18 (1981), 617–627.

[18] Quinn, B.G. – Stationarity and invertibility of simple bilinear models, *Stoch. Processes and their Applications,* 12 (1982), 225–230.

[19] Saikonnen, P. and Luukonen, R. – Lagrange multiplier tests for testing non linearities in time series models, *Scand. J. Statist.,* 15 (1988), 55–68.

[20] Suquet, C. – Introduction à l'association (preprint), *Publications de l'IRMA,* 34(XIII) (1994), Univ. Lille, France.

[21] Tassi, Ph. and Legait, S. – *Théorie des Probabilités en vue des Applications Statistiques*, Éditions Téchnip, Paris, 1990.

E. Gonçalves, C.M. Martins and N. Mendes-Lopes,
Dep. Matemática, Universidade de Coimbra
Coimbra – PORTUGAL