Working Paper 93-01
Statistics and Econometrics Series 01
January 1993

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (341) 624-9849

A GOODNESS-OF-FIT TEST FOR ARMA MODELS BASED ON THE STANDARDIZED SAMPLE SPECTRAL DISTRIBUTION OF THE RESIDUALS

Santiago Velilla*

Abstract

Given the residuals $\hat{\epsilon}, \ldots, \hat{\epsilon}_n$ from the fit of a causal and invertible ARMA(p,q) model $\phi(B)X_t = \theta(B)\epsilon_t$, we introduce a stochastic process derived from the standardized sample spectral of $\hat{\epsilon}, \ldots, \hat{\epsilon}_n$ to construct a goodness-of-fit procedure. We show that the goodness-of-fit statistics considered have a proper limiting distribution which is free of unknown parameters and which, unlike some well-known goodness-of-fit statistics based on the residuals, does not depend on the sample size.

Key Words

ARMA Models; Goodness-of-fit; Residuals; Spectral Density; Weak Convergence.

*Departamento de Estadística y Econometría, Universidad Carlos III de Madrid. Research partially supported by DGICYT grant nº PB90-0266.

1. Introduction. Consider a zero mean stationary process $\{X_t\}$ which follows a causal and invertible ARMA(p,q) model of the form

$$\phi(B)X_t=\theta(B)\varepsilon_t, \tag{1}$$

where: a) $\{\varepsilon_t\}$ is a zero mean white noise sequence with variance σ^2; b) $\phi(B)$ and $\theta(B)$ are polynomials given by

$$\phi(B)=1-\phi_1B- \ldots -\phi_pB^p,$$

$$\theta(B)=1+\theta_1B- \ldots +\theta_qB^q;$$

and, c) B is the backward shift operator $BX_t=X_{t-1}$. Given a finite observable series $(X_1, \ldots ,X_n)$ we can obtain the maximum likelihood estimators $\hat{\phi}$, $\hat{\theta}$, $\hat{\sigma}^2$ of the parameters $\phi=(\phi_1, \ldots ,\phi_p)'$, $\theta=(\theta_1, \ldots ,\theta_q)'$ and σ^2 and form the residual process

$$\hat{\varepsilon}_t=\hat{\theta}^{-1}(B)\hat{\phi}(B)X_t,$$

where $X_t\equiv 0$, for $t\leq 0$, and $\hat{\phi}(B)=1-\hat{\phi}_1B- \ldots -\hat{\phi}_pB^p$, $\hat{\theta}(B)=1+\hat{\theta}_1B+ \ldots +\hat{\theta}_qB^q$.

The residuals $(\hat{\varepsilon}_1, ,,,,\hat{\varepsilon}_n)$ are a natural building block for checking the adecuacy of the fitted model by means of its autocorrelation function $\hat{r}_k$ defined, for $|k|\leq n-1$,

$$\hat{r}_k=\hat{C}_k/\hat{C}_0 \tag{2}$$

where $\hat{C}_k=\hat{C}_{-k}=\sum_{t=1}^{n-k}\hat{\varepsilon}_t\hat{\varepsilon}_{t+k}$, $0\leq k\leq n-1$. For example, a well-known statistic for testing the goodness-of-fit of an ARMA model is the Box-Pierce (Box and Pierce (1970)) statistic

$$Q=n\sum_{k=1}^{m}\hat{r}_k^2, \tag{3}$$

where m is a function of the sample size $m=m_n$. The "asymptotic" distribution of Q is chi squared with m-(p+q) degrees of freedom. In order to improve the chi-square approximation, Ljung and Box (1978) recommended the modification

$$Q_1=n(n+2)\sum_{k=1}^{m}(n-k)^{-1}\hat{r}_k^2. \tag{4}$$

Both (3) and (4) use a time domain approach. The aim of this paper is to consider $(\hat{\varepsilon}_1, ,,,,\hat{\varepsilon}_n)$ in the frecuency domain to propose a new goodness-of-fit test procedure. This new method is based on a stochastic process which depends on the standardized sample spectral density of the residuals. We show that this process converges weakly, under appropriate regularity conditions, to the brownian bridge in [0,1]. All suggested test statistics are functionals of this stochastic process and therefore, unlike the statistics Q and Q_1 above, the asymptotic distribution of every test statistic is properly a limiting distribution and does not depend on the sample size.

The organization of this paper is as follows. Section 2 introduces the new criterion and establishes notation and basic assumptions. Section 3 contains the derivation of the relevant asymptotic results which are obtained by using repeatedly a convergence lemma established in section 2. Section 4 studies the issues involved in the practical application of the theory developed in section 3. Section 5 is devoted to comparisons with previous criteria and section 6 presents a short numerical illustration.

2. Background and motivation. Given the errors $(\varepsilon_1, \ldots, \varepsilon_n)$, we define the standardized sample spectral density

$$I_n(\lambda)=(2\pi C_0)^{-1}\left|\sum_{t=1}^{n}\varepsilon_t \exp(i\lambda t)\right|^2=(2\pi)^{-1}\sum_{k=-(n-1)}^{k=n-1} r_k \cos(\lambda t), \quad -\pi\le\lambda\le\pi, \tag{5}$$

where r_k is the autocorrelation function of $(\varepsilon_1, \ldots, \varepsilon_n)$ and $C_0=\sum_{t=1}^{n}\varepsilon_t^2$. When $(\varepsilon_1, \ldots, \varepsilon_n)$ are observable, a common building block for testing for white noise of $\{\varepsilon_t\}$ is the process

$$Z_n(\lambda)=n^{1/2}[F_n(\lambda)-F_0(\lambda)], \quad 0\le\lambda\le\pi, \tag{6}$$

where $F_n(\lambda)=2\int_0^{\lambda} I_n(u)du$, $F_0(\lambda)=\lambda/\pi=2\int_0^{\lambda} f_0(u)du$, and $f_0(\lambda)=(2\pi)^{-1}$ is the standardized spectral density of a white noise. See, for example, Bartlett (1955) or, more recently, Durlauf (1990) and Anderson (1991). By introducing the change of variable $\lambda=\pi t$, $0\le t\le 1$, and dividing (6) by $\sqrt{2}$, we can write the process (6) in the form

$$U_n(t)=(\sqrt{2}/\pi)n^{1/2}\sum_{k=1}^{n-1} r_k \frac{\sin(k\pi t)}{k}, \quad 0\le t\le 1. \tag{7}$$

Durlauf (1990) shows that, under appropriate regularity conditions, the process $\{U_n(t), 0\le t\le 1\}$ converges weakly to the brownian bridge in $0\le t\le 1$,

$$U(t)=(\sqrt{2}/\pi)\sum_{k=1}^{\infty} u_k \frac{\sin(k\pi t)}{k}, \tag{8}$$

where $\{u_k\}_{1\le k}$ is a sequence of i.i.d. N(0,1) random variables. The goodness-of-fit statistics are nonnegative continuous functionals $h(\varepsilon_1, \ldots,\varepsilon_n;t)=h[U_n(t)]$ whose limiting distribution is that of the random variable $h[U(t)]$.

For the case of model (1), the $(\varepsilon_1, \ldots,\varepsilon_n)$ are not observable. However, if the fitted model is appropriate, the residuals $(\hat{\varepsilon}_1, \ldots,\hat{\varepsilon}_n)$ should behave in a manner consistent with the model and, therefore, in the light of (7), we could consider the process

$$\hat{U}_n(t)=(\sqrt{2}/\pi)n^{1/2}\sum_{k=1}^{m} \hat{r}_k \frac{\sin(k\pi t)}{k}, \quad 0\le t\le 1, \tag{9}$$

where m is a function of the sample size $m=m_n$, as a suitable building block for goodness-of-fit purposes and consider statistics of the form $h[\hat{U}_n(t)]$.

Observe that the process (9) is a reexpression of the process $\hat{Z}_n(\lambda)=n^{1/2}[\hat{F}_n(\lambda)-F_0(\lambda)]$, $0\le\lambda\le\pi$, where $\hat{F}_n(\lambda)=2\int_0^{\lambda}\hat{I}_n(u)du$ and

$$\hat{I}_n(\lambda)=(2\pi)^{-1}\sum_{k=-m}^{m} \hat{r}_k\cos(\lambda k), \quad -\pi\le\lambda\le\pi,$$

is a truncated version of the standardized sample spectral density of $(\hat{\varepsilon}_1, \ldots,\hat{\varepsilon}_n)$. However, as shown below (see theorem 3.4), the process $\{\hat{U}_n(t), 0\le t\le 1\}$ converges weakly, under some regularity conditions, to a zero mean gaussian process with covariance function depending on (ϕ,θ). Therefore, the asymptotic distribution of $h[\hat{U}_n(t)]$ depends on unknown parameters and is of limited practical use. To eliminate this dependence we propose, in the following, a modification to $\hat{U}_n(t)$.

One of the key elements for the weak convergence of the process (7) to the brownian bridge (8), is that, for every fixed integer k, $n^{1/2}R_k=n^{1/2}(r_1,$

$\ldots,r_k)' \xrightarrow{D} N_k(0,I_k)$, when $n \longrightarrow \infty$. However, if $n^{1/2}\hat{R}_k = n^{1/2}(\hat{r}_1, \ldots, \hat{r}_k)'$ by an argument in Box and Pierce (1970), $n^{1/2}\hat{R}_k = n^{1/2}(\hat{r}_1, \ldots, \hat{r}_k)' \xrightarrow{D} N_k(0, I_k - H_k)$, where H_k is the kxk orthogonal projection matrix $H_k = X_k(X_k'X_k)^{-1}X_k'$, and X_k is the kx(p+q) matrix

$$X_k = \begin{vmatrix} 1 & 0 & \ldots\ldots & 0 \\ a_1 & 1 & \ldots\ldots & . \\ a_2 & a_1 & \ldots\ldots & 0 \\ \vdots & \vdots & & . \\ \vdots & \vdots & & 1 \\ & & & . \\ a_{k-1} & a_{k-2} & \ldots\ldots & a_{k-(p+q)} \end{vmatrix}.$$

which depends on the coefficients (a_k) which satisfy

$$a(B) = [\tilde{\phi}(B)]^{-1} = \sum_{k=0}^{\infty} a_k B^k, \qquad (10)$$

where $\tilde{\phi}(B)$ is the polynomial $\tilde{\phi}(B) = \phi(B)\theta(B) = 1 - \tilde{\phi}_1 B - \ldots - \tilde{\phi}_{p+q} B^{p+q}$. We have $I_k - H_k = C_k C_k'$, where $C_k = C_k(\phi, \theta)$ is a kx[k-(p+q)] suborthogonal matrix whose columns span, in $\mathbb{R}^k$, the orthogonal complement of the linear manifold spanned by the columns of X_k, i.e. such that $C_k'C_k = I_{k-(p+q)}$ and $C_k'X_k = 0$. Therefore, $n^{1/2}\hat{C}_k'\hat{R}_k \xrightarrow{D} N_{k-(p+q)}(0, I_{k-(p+q)})$, where $\hat{C}_k = C_k(\hat{\phi}, \hat{\theta})$. For the integer sequence $m = m_n$, define

$$\hat{S}_m = \hat{C}_m'\hat{R}_m = (\hat{s}_{p+q+1,m}, \ldots, \hat{s}_{m,m})',$$

and modify $\hat{U}_n(t)$ to

$$\hat{V}_n(t) = (\sqrt{2}/\pi) n^{1/2} \sum_{k=1}^{m-(p+q)} \hat{s}_{k+p+q,m} \frac{\sin(k\pi t)}{k}, \qquad 0 \le t \le 1. \qquad (11)$$

This modification turns out to be effective because as established in section 3 below the weak limit of $\{V_n(t),\ 0 \le t \le 1\}$ is, under the model (1), the brownian bridge $\{U(t),\ 0 \le t \le 1\}$. The asymptotic distribution of every continuous functional $h[\hat{V}_n(t)]$ is then free of unknown parameters.

3. Asymptotic theory.

3.1 Regularity conditions and auxiliary lemmas. The following lemmas are needed for our theoretical development.

LEMMA 3.1. (Box and Pierce). Let $m=m_n$ and consider the m vectors $\hat{R}_m=(\hat{r}_1, \ldots, \hat{r}_m)'$ and $R_m=(r_1, \ldots, r_m)'$. If we assume:

C.1) $\sup_{k\geq m_n-(p+q)} |a_k|=O(n^{-1/2})$; and

C.2) $m_n=O(n^{1/b})$, $b>2$,

we have,

$$\max_{1\leq k\leq m_n} |\hat{r}_k-\xi_{km}|=O_p(n^{-1}),$$

as $n\longrightarrow\infty$, where $\xi_m=(\xi_{1m}, \ldots, \xi_{mm})'=(I_m-H_m)R_m$.

LEMMA 3.2. Let $\{A_n(t),\ 0\leq t\leq 1\}_{n\geq 1}$ be a sequence of stochastic processes and let $\{A(t),\ 0\leq t\leq 1\}$ be a process such that, for every integer d, we can write

$$A_n(t)=A_n^d(t)+R_n^d(t);$$
$$A(t)=A^d(t)+R^d(t),$$

such that

(*i*) For every d, $A_n^d(t)\xrightarrow[w]{}A^d(t)$ as $n\longrightarrow\infty$; and

(*ii*) For every $\varepsilon>0$,

$$\overline{\lim_d}\ \overline{\lim_n}\ P[\sup_{0\leq t\leq 1} |R_n^d(t)|>\varepsilon]=\overline{\lim_d}\ P[\sup_{0\leq t\leq 1} |R^d(t)|>\varepsilon]=0.$$

Then

$$A_n(t)\xrightarrow[w]{}A(t) \text{ as } n\longrightarrow\infty.$$

Proof. The proof follows easily from the necessary and sufficient conditions for the weak convergence of $A_n(t)$ to $A(t)$ given in Billingsley (1968) theorems 8.1 and 8.2. ∎

Besides the regularity conditions C.1) and C.2), we will assume:

C.3) The error process $\{\varepsilon_t\}$ in model (1) is i.i.d with $E[\varepsilon_1^8]<\infty$.

C.4) $\|(\hat{C}_m - C_m)\hat{R}_m\|_\infty = O_p(1/n)$, where $\|A\|_\infty$ denotes the maximum absolute value of the elements of the matrix A.

3.2. Weak convergence of the process $\hat{V}_n(t)$. We establish in this section the main result of this paper.

THEOREM 3.3. Under conditions C.1)—C.4), we have

$$\hat{V}_n(t) \xrightarrow[w]{} U(t).$$

Proof. Consider the three processes

$$\hat{V}_n(t)=(\sqrt{2}/\pi)n^{1/2}\sum_{k=1}^{m-(p+q)} \hat{s}_{k+p+q,m}\frac{\sin(k\pi t)}{k}, \qquad 0\le t\le 1;$$

$$V_n(t)=(\sqrt{2}/\pi)n^{1/2}\sum_{k=1}^{m-(p+q)} s_{k+p+q,m}\frac{\sin(k\pi t)}{k}, \qquad 0\le t\le 1;$$

$$W_n(t)=(\sqrt{2}/\pi)n^{1/2}\sum_{k=1}^{m-(p+q)} t_{k+p+q,m}\frac{\sin(k\pi t)}{k}, \qquad 0\le t\le 1;$$

where

$$(s_{p+q+1,m}, \ldots, s_{m,m})'=C'_m\hat{R}_m, \qquad (t_{p+q+1,m}, \ldots, t_{m,m})'=C'_mR_m.$$

The claim of the theorem will follow if we prove:

a) $V_n(t) \xrightarrow[w]{} U(t)$ and

b) $\sup_{0\le t\le 1}|\hat{V}_n(t)-V_n(t)|=o_p(1)$.

Step a). It can be seen that for *every* suborthogonal C_m matrix such that $C_mC'_m=I_m-Q_m$, we have $\|C'_m(\hat{R}_m-R_m)\|_\infty = \|C'_m\hat{R}_m - C'_mC_mC'_mR_m)\|_\infty \le m^{1/2}\|\hat{R}_m - C_mC'_mR_m)\|_\infty = m^{1/2}\|\hat{R}_m-\xi_m\|_\infty = m^{1/2}O_p(1/n)=o_p(n^{-3/4})$ by lemma 3.1 above. By Cauchy- Schwartz,

$\sup_{0\le t\le 1} |V_n(t)-W_n(t)| \le n^{1/2}m^{1/2}o_p(n^{-3/4})=o_p(1)$ and, therefore, the weak limit of $V_n(t)$ and $W_n(t)$, if they exist, are the same.

Fix now an integer $d>p+q$. We can take

$$C_m = \left(\begin{array}{c|c} C_d & A \\ \hline 0 & B \end{array} \right), \tag{12}$$

where C_d is a suborthogonal matrix of $d\times[d-(p+q)]$, 0 is of $(m-d)\times[d-(p+q)]$, A is of $d\times(m-d)$ and B is of $(m-d)\times(m-d)$. We then write

$$W_n(t)=W_n^d(t)+R_n^d(t),$$

where

$$W_n^d(t) = (\sqrt{2}/\pi)n^{1/2} \sum_{k=1}^{d-(p+q)} t_{k+p+q,m} \frac{\sin(k\pi t)}{k},$$

and

$$R_n^d(t) = (\sqrt{2}/\pi)n^{1/2} \sum_{k=d+1-(p+q)}^{m-(p+q)} t_{k+p+q,m} \frac{\sin(k\pi t)}{k}.$$

If $\{u_k\}_{1\le k}$ is a sequence of i.i.d. $N(0,1)$ random variables, we proof that $W_n^d(t) \xrightarrow[w]{} U^d(t)=(\sqrt{2}/\pi) \sum_{k=1}^{d-(p+q)} u_{k+p+q} \frac{\sin(k\pi t)}{k}$ and that, for every $\varepsilon>0$, $\overline{\lim}_d \overline{\lim}_n P[\sup_{0\le t\le 1} |R_n^d(t)|>\varepsilon]=0$. Since, by construction of the brownian bridge (8), for every $\varepsilon>0$, $\overline{\lim}_d P[\sup_{0\le t\le 1} |R^d(t)|>\varepsilon]=0$, where $R^d(t)=U(t)-U^d(t)$, weak convergence for $W_n(t)$ follows from the convergence lemma 3.2.

Under regularity condition C.3), $n^{1/2}R_d=n^{1/2}(r_1, \ldots, r_d)' \xrightarrow{D} N_d(0,I_d)$, and, since $C_d'C_d=I_{d-(p+q)}$, $n^{1/2}(t_{p+q+1,m}, \ldots, t_{d,m})'=n^{1/2}C_d'R_d \longrightarrow N_{d-(p+q)}(0,I_{d-(p+q)})$. Routine application of the conditions for weak convergence in theorem 8.2 of Billingsley (1968) yields $W_n^d(t) \xrightarrow[w]{} U^d(t)$.

We now write $(A',B')=(\alpha_{k+p+q,j})$ for $(d+1)-(p+q)\le k\le m-(p+q)$ and $1\le j\le m$ so that $t_{k+p+q,m}=\sum_{j=1}^{m} \alpha_{k+p+q,j}r_j$ and $R_n^d(t) = (\sqrt{2}/\pi)(C_0/n)^{-1}Q_n(t)$, where

$$Q_n^d(t)=n^{-1/2}\sum_{k=d+1-(p+q)}^{m-(p+q)} \beta_{k+p+q}\frac{\sin(k\pi t)}{k},$$

$\beta_{k+p+q}=\sum_{j=1}^{m}\alpha_{k+p+q,j}C_j$, and $C_j=\sum_{t=1}^{n-j}\varepsilon_t\varepsilon_{t+j}$. Since $C_0/n \longrightarrow E[\varepsilon_1^2]$, it suffices to prove that, given $\varepsilon>0$,

$$\overline{\lim_d}\ \overline{\lim_n}\ P[\sup_{0\le t\le 1}|Q_n^d(t)|>\varepsilon]=0.$$

But $|Q_n^d(t)|^2\le n^{-1}|\sum_{k=d+1-(p+q)}^{m-(p+q)}\beta_{k+p+q}\frac{\exp(ik\pi t)}{k}|^2$ and, therefore,

$$\sup_{0\le t\le 1}|Q_n^d(t)|^2\le Q_n^d=2\sum_{r=0}^{m-(d+1)}|Y_r|,$$

where $Y_r=n^{-1}\sum_{k=d+1-(p+q)}^{m-r-(p+q)}\frac{\beta_{k+p+q}\,\beta_{k+r+p+q}}{k(k+r)}$. We have

$$E[Y_r^2]\le n^{-2}\sum_{k=d+1-(p+q)}^{m-r-(p+q)}\ \sum_{l=d+1-(p+q)}^{m-r-(p+q)}|E[\frac{\beta_{k+p+q}\,\beta_{k+r+p+q}\,\beta_{l+p+q}\,\beta_{l+r+p+q}}{k(k+r)l(l+r)}]|.$$

By lemma 1 in Grenander and Rosenblatt (1957, p. 186), and exploiting the fact that the rows of (A',B') are orthonormal, we get:

(*i*) $E[\beta_{k+p+q}^2]\le n$;

(*ii*) $E[\beta_{k+p+q}^2\,\beta_{k+r+p+q}^2]\le An^2$, $r\ge 1$;

(*iii*) $|E[\beta_{k+p+q}\beta_{l+p+q}\beta_{k+r+p+q}\beta_{l+r+p+q}]|\le Bnm$, $k<l$, $r\ge 1$.

By Cauchy-Schwartz it is easily seen that

$$E[Q_n^d]\le\sum_{k=d+1-(p+q)}^{m-(p+q)}k^{-2}+[A+2B(m^2/n)]^{1/2}\sum_{r=0}^{m-(d+1)}[\sum_{k=d+1-(p+q)}^{m-r-(p+q)}\frac{1}{k^2(k+r)^2}]^{1/2}.$$

By observing that $m=O(n^{1/b})$ and by recalling an argument given in Grenander and Rosenblatt (1957, p. 189), we get $\overline{\lim_d}\ \overline{\lim_n}\ E[Q_n^d]=0$. This completes the proof of step a).

Step b). By C.4), $\sup_{0\le t\le 1}|\hat{V}_n(t)-V_n(t)|\le(\sqrt{2}/\pi)n^{1/2}\sum_{k=1}^{m-(p+q)}|\hat{s}_{k+p+q,m}-s_{k+p+q,m}|/k$ $\le n^{1/2}m^{1/2}\ O_p(1/n)=o_p(1)$. ■

3.3. Weak convergence of the process $\hat{U}_n(t)$. For completeness, we study the convergence of the process $\{U_n(t),\ 0\le t\le 1\}$. We will need the additional regularity condition:

C.5) In matrix C_m in (12) above, $\|A\|_E=O(d^{-1/2})$ and $\|B-I_{m-d}\|_E=O(d^{-1/2})$, where $\|.\|_E$ denotes the euclidean norm of a matrix.

We also define the array $H=(h_{jk})_{1\le j,k}$,

$$h_{jk}=\sum_{r=1}^{p+q}\ \sum_{s=1}^{p+q} a_{j-r}\, a_{k-s}\, \tilde{\Gamma}^{rs}_{p+q},$$

where the coefficients $\{a_k\}$ are as defined in (11), $a_k=0$ for $k<0$, and $\tilde{\Gamma}^{rs}_{p+q}$ are the elements of the inverse of the $(p+q)x(p+q)$ matrix $\tilde{\Gamma}_{p+q}=(\sum_{k=0}^{\infty} a_k a_{k+|r-s|})$.

THEOREM 3.4. Under conditions C.1), C.2), C.3) and C.5), we have

$$\hat{U}_n(t)\xrightarrow[w]{} G(t).$$

The process $\{G(t),\ 0\le t\le 1\}$ is defined by

$$G(t)=(\sqrt{2}/\pi)\sum_{k=1}^{\infty} u_k\ p_k[t],$$

where a) $\{u_k\}_{1\le k}$ is a sequence of i.i.d. $N(0,1)$ random variables; b) $p_k[t]=\sum_{j=1}^{\infty}[\delta_{jk}-h_{jk}]\ \dfrac{\sin(j\pi t)}{j}$ ($\delta_{jk}=1$, $j=k$, and $\delta_{jk}=0$, $j\ne k$).

Proof. (Sketch). By lemma 3.1, $\sup_{0\le t\le 1}|\hat{U}_n(t)-Y_n(t)|=o_p(1)$, where

$$Y_n(t)=(\sqrt{2}/\pi)n^{1/2}\sum_{k=1}^{m} \xi_{km}\,\frac{\sin(k\pi t)}{k},\qquad 0\le t\le 1;$$

and $\xi_m=(\xi_{1m},\ \dots,\xi_{mm})'=C_m C'_m R_m=\left(\dfrac{C_d C'_d R_d}{0}\right)+\left(\dfrac{A}{B}\right)(A'\ \ B')R_m$. If

$$Y^d_n(t)=(\sqrt{2}/\pi)n^{1/2}\sum_{k=1}^{d} \alpha_k\,\frac{\sin(k\pi t)}{k},\qquad 0\le t\le 1,$$

where $(\alpha_1,\ \dots,\alpha_d)'=C_d C'_d R_d$ and we now take

$$G^d(t)=(\sqrt{2}/\pi)\sum_{k=1}^{d}u_k[\sum_{j=1}^{d}(\delta_{jk}-h_{jk})\frac{\sin(j\pi t)}{j}], \quad 0\le t\le 1,$$

it can be shown that $Y_n^d(t) \xrightarrow[w]{} G^d(t)$. By considering the representations

$$Y_n(t)=Y_n^d(t)+R_n^d(t),$$

$$G(t)=G^d(t)+R^d(t),$$

we need to prove: a) $\overline{\lim_d}\ \overline{\lim_n}\ P[\sup_{0\le t\le 1}|R_n^d(t)|>\varepsilon]=0$ and b) $\overline{\lim_d}$ $P[\sup_{0\le t\le 1}|R^d(t)|>\varepsilon]=0$.

Step a) $\sup_{0\le t\le 1}|R_n^d(t)|\le(\sqrt{2}/\pi)n^{-1/2}(C_0/n)^{-1}|\sum_{j=1}^{m}\frac{e^{ij\pi t}}{j}D_j|\le$

$$(\sqrt{2}/\pi)(C_0/n)^{-1}n^{-1/2}[n^{-1/2}|\sum_{j=1}^{d}\frac{e^{ij\pi t}}{j}D_j|+n^{-1/2}|\sum_{j=d+1}^{m}\frac{e^{ij\pi t}}{j}D_j|]\le$$

$$(\sqrt{2}/\pi)(C_0/n)^{-1}[A_n^d+B_n^d],$$

say, where $D_j=\sum_{k=d+1-(p+q)}^{m-(p+q)}\alpha_{k+p+q,j}\beta_{k+p+q}$. We have $(A_n^d)^2\le n^{-1}[\sum_{j=1}^{d}j^{-2}][\sum_{j=1}^{d}D_j^2]$

and

$$E[D_j^2]=\sum_{k=d+1-(p+q)}^{m-(p+q)}\sum_{l=d+1-(p+q)}^{m-(p+q)}\alpha_{k+p+q,j}\alpha_{l+p+q,j}E[\beta_{k+p+q}\beta_{l+p+q}]\le$$

$$(n-1)\sum_{k=d+1-(p+q)}^{m-(p+q)}\alpha^2_{k+p+q,j}+m^2\sum_{k=d+1-(p+q)}^{m-(p+q)}\alpha^2_{k+p+q,j},$$

so that

$$\sum_{j=1}^{d}E[D_j^2]\le[(n-1)+m^2]\|A\|_E^2,$$

and from here and regularity condition C.5), $\overline{\lim}\ E[(A_n^d)^2]=0$.

We now observe that $B_n^d\le n^{-1/2}|\sum_{j=1}^{d}\frac{e^{ij\pi t^d}}{j}(D_j-\beta_j)|+n^{-1/2}|\sum_{j=d+1}^{m}\frac{e^{ij\pi t}}{j}\beta_j|=C_n^d+D_n^d$. By an argument used in the proof of theorem 3.3 $\overline{\lim_d}\ E[(D_n^d)^2]=0$. It can be shown that $\sum_{j=1}^{d} E[(D_j-\beta_j)^2]\leq[(n-1)+m^2]\|B-I_{m-d}\|_E^2$ so that, again by regularity condition C.5), $\overline{\lim_d}\ E[(C_n^d)^2]=0$.

Step b) We can write $R^d[t]=RA^d[t]+RB^d[t]$, where

$$RA^d[t]=-(\sqrt{2}/\pi)\sum_{k=d+1}^{\infty} u_k\ [\ \sum_{j=1}^{d} h_{jk}\frac{\sin(j\pi t)}{j}\ ],$$

and

$$RB^d[t]=-(\sqrt{2}/\pi)\sum_{k=d+1}^{\infty} u_k \frac{\sin(k\pi t)}{k} + (\sqrt{2}/\pi)\sum_{k=1}^{\infty} u_k\ [\ \sum_{j=d+1}^{\infty} h_{jk}\frac{\sin(j\pi t)}{j}\ ].$$

We have $RA^d=\sup_{0\leq t\leq 1}|RA^d[t]|^2\leq(4/\pi^2)\sum_{r=0}^{d-1}|Y_r|$, where $Y_r=\sum_{j=1}^{d-r}\frac{\beta_j\beta_{j+r}}{j(j+r)}$ and $\beta_j=\sum_{k=d+1}^{\infty} h_{jk}u_k$. It can be seen that $|E[\beta_j\beta_{j+r}\beta_1\beta_{1+r}]|\leq O(d^{-1})\sum_{k=d+1}^{\infty}|a_k|$, and therefore $E[Y_r^2]\leq O(1)\sum_{k=d+1}^{\infty}|a_k|\sum_{j=1}^{d-r}\frac{1}{j^2(j+r)^2}$. By Cauchy-Schwartz,

$$\overline{\lim_d}\ E[RA^d]\leq O(1)[\sum_{k=d+1}^{\infty}|a_k|]^{1/2}\sum_{r=0}^{d-1}[\sum_{j=1}^{d-r}\frac{1}{j^2(j+r)^2}]^{1/2}=0.$$

Write $RB^d[t]=A^d[t]+B^d[t]$. By Brownian bridge construction, $\overline{\lim_d}\ P[\sup_{0\leq t\leq 1}|A^d[t]|>\varepsilon]=0$. On the other hand, $B^d=\sup_{0\leq t\leq 1}|B^d[t]|^2\leq(2/\pi^2)[\sum_{j=d+1}^{\infty} j^{-2}][\sum_{j=d+1}^{\infty}\beta_j^2]$. Since $E[\beta_j^2]=\sum_{k=1}^{\infty} h_{jk}^2$ we have $\overline{\lim_d}\ E[B^d]=0$.

Remark 3.5. Since the matrix H_m is symmetric and idempotent for every m, it is easy to proof that $h_{jk}=\sum_{s=1}^{\infty} h_{js}h_{sk}$ and, therefore, the process G(t) is a zero mean gaussian process with covariance function given by

$$\hat{C}(t,u)=C(t,u)-D(t,u),$$

where $C(t,u)=\min(t,u)-tu$, $D(t,u)=(2/\pi^2)\sum_{j=1}^{\infty}\sum_{k=1}^{\infty} h_{jk}\frac{\sin(j\pi t)}{j}\frac{\sin(k\pi u)}{k}$. ■

4. Testing the goodness-of-fit of an ARMA(p,q) model

4.1. Testing criteria

The test statistics to be considered will be of the form $D_n=D[\hat{V}_n(t)]$, where D[.] is a continuous nonnegative functional defined on $C_{[0,1]}$, the space of continous functions in [0,1] endowed with the sup norm. Extreme values of these functionals should indicate the inadecuacy of model (1). The asymptotic distribution of the random variable $D_n=D[\hat{V}_n(t)]$ is, by a straightforward application of the continuous mapping theorem, the distribution of the random variable D=D[U(t)])]. Approximate significance levels are then computed with reference to the distribution of the functional D=D[U(t)])]. Standard choices for D[.] and its associated tests statistics are:

a) Kolmogorov-Smirnov criterion:

$$D_{KS}[f(t)]=\sup_{0\leq t\leq 1}|f(t)|; \qquad D_{n,KS}=\sup_{0\leq t\leq 1}|\hat{V}_n(t)|;$$

b) Cramér-von Mises criterion:

$$D_{CvM}[f(t)]=\int_0^1[f(t)]^2dt; \qquad D_{n,CvM}=\int_0^1[\hat{V}_n(t)]^2dt;$$

c) Anderson-Darling criterion:

$$D_{AD}[f(t)]=\int_0^1[f(t)]^2u(t)dt, \qquad D_{n,AD}=\int_0^1[\hat{V}_n(t)]^2u(t)dt,$$

with u(t)=1/[t(1-t)].

For the functionals a), b) and c) above, tables can be seen in Shorack and Wellner (1987). It is interesting to observe that, for the Cramér-von Mises criterion, we get

$$D_{n,CvM}=(n/\pi^2)\sum_{k=1}^{m-(p+q)}(\hat{s}^2_{k+p+q,m}/k^2)$$

which is a statistic of similar structure as the Box-Pierce statistic Q in (3).

4.2. Computational issues. For computing the process (11) and the test statistics above we need:

(1) The MLE's $\hat{\phi}$, $\hat{\theta}$;

(2) The estimated residuals $(\hat{\varepsilon}_1, ,,,,\hat{\varepsilon}_n)$ and the associated correlation function $\hat{r}_k$, k=1, ...,m;

(3) The matrix $\hat{C}_m=C_m(\hat{\phi},\hat{\theta})$ which we compute as follows:

(*i*) We obtain the coefficients $\{a_k=a_k(\phi,\theta)\}$ by setting $a_0=1$ and using the recursive relation

$$a_k-\tilde{\phi}_1 a_{k-1}- \ldots -\tilde{\phi}_{p+q} a_{k-(p+q)}=0, \tag{16}$$

where $a_k=0$ for k<0. For example, for the AR(1) model, $X_t-\phi X_{t-1}=\varepsilon_t$, we have $a_k=\phi^k$ while for the AR(2) model, $X_t-\phi_1 X_{t-1}-\phi_2 X_{t-2}=\varepsilon_t$, we get $a_0=1$, $a_1=\phi_1$, $a_2=\Phi_1^2+\Phi_2$, $a_3=\Phi_1^3+2\phi_1\Phi_2$, and so on.

(*ii*) We compute $\hat{a}_k=a_k(\hat{\phi},\hat{\theta})$ and form the matrix $\hat{X}_m=X_m(\hat{\phi},\hat{\theta})$.

(*iii*) Because rank$[\hat{X}_m]$=p+q, only p+q eigenvalues of $\hat{X}_m\hat{X}_m'$ are different from zero. By the singular value decomposition of $\hat{X}_m$, we can chose for $\hat{C}_m$ the mx[m-(p+q)] matrix formed by the m-(p+q) eigenvectors of $\hat{X}_m\hat{X}_m'$ corresponding to the m-(p+q) null eigenvalues of $\hat{X}_m$.

5. Comparisons with previous criteria.

We compare briefly, in this section, the goodness-of-fit tests based on the process $\hat{V}_n(t)$ with the classical goodness-of-fit tests based in the statistics Q of Box and Pierce (1970) and Q_1 of Ljung and Box (1978). We compare also $\hat{V}_n(t)$ with other spectral goodness-of-fit criteria proposed in the literature.

With respect to Q and Q_1, it seems that the "averaging" used in (14) provides a proper limiting distribution for every $D_n=D[\hat{V}_n(t)]$ as $n\longrightarrow\infty$. This is in contrast with Q and Q_1 whose "asymptotic" distribution depends on the sample size through $m=m_n$. On the other hand, on the light of section 4.2, criteria based on $\hat{V}_n(t)$ are computationally more expensive but, however, not by a great amount.

Several other spectral based goodness-of-fit criteria have been proposed in the literature. See, e.g., Priestley (1981, chap. 6), Dzhaparidze (1986, chap. V) or the recent review contained in Anderson (1991). These previous methods are not based in the residuals and use directly the observations $(X_1, \ldots, X_n)$. With some exceptions, for example the Quenouille's goodness of fit test for autoregressive models described in Priestley (1981, p. 488), these methods are fully manageable only in the case when we assume a *completely* specified model (both the orders (p,q) and the parameters (ϕ,θ,σ^2) are known). This is a very restrictive framework for goodness-of-fit purposes because to check the appropriateness of an ARMA(p,q) model for the data, we should assume only that the data $X_1, \ldots, X_n$ are generated by an ARMA(p,q) process with unknown parameters (ϕ,θ,σ^2). It is important to remark that the process $\hat{V}_n(t)$ converges weakly to the brownian bridge for *every* model in the class of causal and invertible ARMA(p,q) models (1). Therefore, $\hat{V}_n(t)$ provides a sensible building block for goodness-of-fit under a more flexible framework. Recall also that the convergence of $\hat{V}_n(t)$ does not require normality of the errors $(\varepsilon_1, \ldots, \varepsilon_n)$.

6. Example. We illustrate, in this section, the proposed goodness-of-fit procedure with reference to the first n=100 simulated observations from the AR(2) model $X_t - .4X_{t-1} + .7X_{t-2} = \varepsilon_t$, $\{\varepsilon_t\} \sim N(0,1)$ which appear in the appendix of Priestley (1981). Given the data, we try to find out about the merits of the models AR(1) and AR(2) in providing a satisfactory fit for the data. We take m=10.

(*i*) For the AR(1) model $X_t - \phi X_{t-1} = \varepsilon_t$, we get $\hat{\phi}=.18$. Table 1.a) displays the autocorrelation function $\hat{r}_k$, k=1, ...,10, the sequence $\hat{a}_k$, k=0,1, ...,9, and the elements $\hat{s}_{k,10}$, k=2, ...,10.

k	0	1	2	3	4	5	6	7	8	9	10
$\hat{r}_k$	-	.131	-.687	-.368	.359	.494	-.080	-.438	-.095	.299	.128
$\hat{a}_k$	1.	.180	.030	.005	.001	.000	.000	.000	.000	.000	-
$\hat{s}_{k,10}$	-	-	.007	-.500	-.851	-.155	-.008	-.438	-.095	.299	.128

Table 1.a)

Table 1.b) displays the values of the statistics presented in section 4.1 which are all highly significant.

$D_{n,KS}$=1.83;	$D_{n,CvM}$=1.5335;	$D_{n,AD}$=11.24.

Table 1.b)

(*ii*) For the AR(2) model $X_t-\phi_1X_{t-1}-\phi_2X_{t-2}=\varepsilon_t$, we get $\hat{\phi}_1$=.326 and $\hat{\phi}_2$=-.762. Table 2.a) displays the autocorrelation function $\hat{r}_k$, k=1, ...,10, the sequence $\hat{a}_k$, k=0,1, ...,9, and the elements $\hat{s}_{k,10}$, k=3, ...,10.

k	0	1	2	3	4	5	6	7	8	9	10
$\hat{r}_k$	-	-.095	-.007	.009	-.093	.151	.054	-.101	-.109	.176	-.045
$\hat{a}_k$	1.	.326	-.656	-.462	.349	.466	-.114	-.392	-.041	.286	-
$\hat{s}_{k,10}$	-	-		.054	-.017	-.13	.162	.064	.030	.230	.023

Table 2.a)

Table 2.b) displays the values of the statistics presented in section 4.1. which confirm the adecuacy of a model AR(2).

$D_{n,KS}$=0.62;	$D_{n,CvM}$=0.0786;	$D_{n,AD}$=0.4352.

Table 2.b)

Acknowledgements. Research partially supported by Grant DGICYT PB90-0266 (Spain).

REFERENCES

ANDERSON, T.W. (1991), Goodness of fit tests for spectral distributions. Tech. Report No 27, Stanford University.

BARTLETT, M.S. (1955), *An Introduction to Stochastic Processes*, Cambridge University Press.

BILLINGSLEY, P. (1968), *Convergence of Probability Measures*, John Wiley, New York.

BOX, G.E.P. and PIERCE, D.A (1970), Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *J. A. S. A.* 65, 1509-1526.

DURLAUF, S.N. (1990), Spectral based testing of the martingale hypothesis, Technical Working Paper No. 90, NBER.

DZHAPARIDZE, K. (1986), *Parameter Estimation and Hypothesis Testing in Spectral Analysis of Stationary Time Series*, Springer Verlag, New York.

GRENANDER, U. and ROSENBLATT, M. (1957), *Statistical Analysis of Stationary Time Series*, John Wiley, New York.

LJUNG, G.M. and BOX, G.E.P. (1978), On a measure of lack of fit in time series models, *Biometrika*, 65, 297-303.

PRIESTLEY, M.B. (1981), *Spectral Analysis and Time Series*, Vols. 1 and 2, Academic Press, Mew York.

SHORACK, G.R. and WELLNER, J.A. (1987), *Empirical Processes with Applications to Statistics*, John Wiley, New York.