Goodness-of-fit tests of normality for the innovations in ARMA models

(abbreviated title: Testing the residuals in ARMA)

Gilles R. Ducharme

and

Pierre Lafaye de Micheaux

Laboratoire de probabilités et statistique, cc051
Université Montpellier II
Place Eugène Bataillon
34095, Montpellier, Cedex 5
France

Key Words: ARMA process, Gaussian white noise, Goodness-of-fit test, Normality of residuals, Smooth test.

Summary

In this paper, we propose a goodness-of-fit test of normality for the innovations of an ARMA(p, q) model with known mean or trend. This test is based on the data driven smooth test approach and is simple to perform. An extensive simulation study is conducted to study the behavior of the test for moderate sample sizes. It is found that our approach is generally more powerful than existing tests while holding its level throughout most of the parameter space and thus, can be recommended. This agrees with theoretical results showing the superiority of the data driven smooth test approach in related contexts.

1 Introduction

Let $(Y_t,\, t \in \mathbb{Z})$ be a stationary process. In this paper, we consider the case where $E(Y_t)$ is known or has been estimated using information outside of the data set. Thus, without loss of generality, we set $E(Y_t) = 0$. Consider the framework where $(Y_t,\, t \in \mathbb{Z})$ obeys the causal and invertible finite order ARMA(p, q) model

$$Y_t - \boldsymbol{\varphi}^{\mathsf{T}} \boldsymbol{Y}_{t-1}^{(p)} = \boldsymbol{\theta}^{\mathsf{T}} \boldsymbol{\epsilon}_{t-1}^{(q)} + \epsilon_t \tag{1.1}$$

where $(\epsilon_t,\, t \in \mathbb{Z})$ is an innovation process of random variables with mean 0 and autocovariance $E(\epsilon_t \epsilon_{t+h}) = \sigma^2 < \infty$ (unknown) if $h = 0$ and 0 otherwise and where

$$\boldsymbol{\varphi} = \begin{bmatrix} \varphi_1 \\ \vdots \\ \varphi_p \end{bmatrix}, \ \boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_q \end{bmatrix}, \ \boldsymbol{Y}_{t-1}^{(p)} = \begin{bmatrix} Y_{t-1} \\ \vdots \\ Y_{t-p} \end{bmatrix}, \ \boldsymbol{\epsilon}_{t-1}^{(q)} = \begin{bmatrix} \epsilon_{t-1} \\ \vdots \\ \epsilon_{t-q} \end{bmatrix}.$$

A sample $\{Y_1, ..., Y_T\}$ is observed and model (1.1) is fitted by standard methods, for example the unconditional Gaussian maximum likelihood approach (see Brockwell and Davis (1991), p. 256-257), yielding the estimator $\hat{\boldsymbol{\beta}} = (\hat{\boldsymbol{\varphi}}^{\mathsf{T}}, \hat{\boldsymbol{\theta}}^{\mathsf{T}}, \hat{\sigma})^{\mathsf{T}}$ of $\boldsymbol{\beta} = (\boldsymbol{\varphi}^{\mathsf{T}}, \boldsymbol{\theta}^{\mathsf{T}}, \sigma)^{\mathsf{T}}$.

If it can be safely assumed that the distribution of the $(\epsilon_t,\, t \in \mathbb{Z})$ generating the Y_t's is of a given form, in particular independent identically distributed (*i.i.d.*) normal (Gaussian) random variables, then better inference can be obtained from the fitted model. For example, such an assumption is helpful to get accurate confidence or tolerance bounds for a predicted Y_{T+h}. Moreover, under this Gaussian assumption, $\hat{\boldsymbol{\beta}}$ is asymptotically efficient. It is thus important to have a tool to check the null hypothesis

$$H_0: \text{ the } \epsilon_t\text{'s are } i.i.d. \sim N(0, \sigma^2). \tag{1.2}$$

As pointed out by Pierce and Gray (1985) and Brockett et al. (1988), other reasons may motivate a test of (1.2). One such reason is to check the fit of the structural part of (1.1). Indeed, the process of fitting a model to data often reduces to finding the model whose residuals behave most like a sample of *i.i.d.* Gaussian variables. In this context, rejection of (1.2) may indicate lack-of-fit of the entertained ARMA model. We will not elaborate further here on this possibility and assume, in the sequel, that model (1.1) is not underspecified. Note however that there exist specific tests for detecting lack-of-fit (for a recent review, see Koul and Stute (1999).

For the problem of testing (1.2), the few tests available fall roughly into two groups. Tests of the first group use the fact that for the ARMA (p, q) models, normality of the Y_t's induces normality of the ϵ_t's and vice versa. Thus a test of the hypothesis that a process (Y_t, $t \in \mathbb{Z}$) is Gaussian (Lomnicki (1961); Hinich (1982); Epps (1987)) can serve for problem (1.2). This presents the advantage of not requiring the values of p and q. But Gasser (1975) and Granger (1976) have shown, and Lutkepohl and Schneider (1989) have confirmed by simulation, that this approach may lose much power. This is because the central limit theorem forces the Y_t's to be close to normality even when (1.2) is false. Moreover, the adaptation of standard normality tests to dependent data is not an easy task. A small simulation study by Heuts and Rens (1986) has shown that, because of the serial correlation between the Y_t's, the finite null behavior of standard normality tests based on the empirical distribution function (EDF) of the Y_t's is different from what is obtained under *i.i.d.* data. The same problem appears for tests based on the third or fourth moment of Y_t (see Lomnicki (1961); Lutkepohl and Schneider (1989)) and for Pearson's chi-square test (Moore (1982)).

It thus appears better, when there are reasons to believe that a given ARMA(p, q) model holds, to "inverse filter" the data and compute the residuals $\hat{\epsilon}_t$ of the fitted model. These can then be subjected to some test of normality. Tests of the second group are based on this idea and some examples are listed in Hipel and McLeod (1994). However, these and other authors use such tests in conjunction with critical values for *i.i.d.* data. Since the residuals of an ARMA model are dependent, the null distribution of standard test statistics may be affected and critical values for *i.i.d.* data may no longer be valid. It turns out that for AR models, there is theoretical evidence that this dependence affects only slightly the critical values, at least when T is large. For an AR(p) model with unknown $E(Y_t)$, Pierce and Gray (1985) has shown that the asymptotic null distribution of any test statistic based on the EDF of the residuals coincides with that of the same statistic for *i.i.d.* data with mean and variance unknown. Thus one can insert the residuals from an AR(p) model into any of the standard EDF-based tests (Kolmogorov-Smirnov, Anderson-Darling) and if T is large, use the critical values given, for example, in Chapter 4 of D'Agostino and Stephens (1986), to obtain an asymptotically valid test strategy. In the same vein, Lee and Na (2002) have recently adapted the Bickel-Rosenblatt test to this AR setting. Beiser (1985) has found that for the AR(1) model, tests based on the skewness or kurtosis coefficient of the residuals (D'Agostino and Stephens (1986), p. 408) in conjunction with the critical points derived for *i.i.d.* data produce valid levels if T is large and the AR-parameter is not too close to its boundary. This has been confirmed by Lutkepohl and Schneider (1989). See also Anděl (1997).

For the general ARMA model, much less is known. Ojeda et al. (1997) show that tests based on quadratic forms in differences between sample moments and expected values of certain non-linear functions of the sample have the same asymptotic distribution under the ARMA model as under *i.i.d.* data. This suggests that a generalization of Pierce and Gray (1985) theorem to ARMA models could hold although, to our knowledge, no proof of this has been published. In accordance with this conjecture, the practice recommended in many textbooks (see for example, Brockwell and Davis (1991), p. 314; Hipel and McLeod (1994), p. 241) is to use standard normality tests in conjunction with critical values for *i.i.d.* data.

In this paper, we develop some tests designed specifically for problem (1.2) in the ARMA(p, q) context. Our approach is based on the smooth test paradigm introduced by Neyman (1937) and improved by the data driven technology introduced by Ledwina (1994) to select the best order for the test. This approach has been shown in the *i.i.d.* case to offer many advantages, both theoretically and empirically, over other tests. In particular, the test statistic we recommend for problem (1.2) is easy to compute with an asymptotic χ^2 distribution that can be corrected in finite samples to yield a close to nominal level. Moreover, as a byproduct of the procedure, diagnostic information is available that helps in understanding which aspects of the null hypothesis are not supported by the data.

Note that we concentrate here on the development of valid tests along this paradigm and do not dwell into their theoretical properties (i.e. local power and asymptotic efficiency). We also stress that the tests proposed here are valid solely for the case where $E(Y_t)$ is assumed known. The case where an unknown trend is present in (1.1) requires a special treatment and is the object of current research.

The paper is organized as follows. In Section 2, we develop the smooth goodness-of-fit test in the ARMA(p, q) context of (1.1). In Section 3, we describe the data-driven technology that allows to "fine tune" the test by choosing a good value for its order. In Section 4, a Monte-Carlo study is conducted for some values of (p, q) to study the behavior of the proposed tests under the null hypothesis and compare their power to some competitors. It emerges that, under the null hypothesis, one of our data driven smooth tests holds its level over most of the parameter space and, under the alternatives studied, is in general more powerful than other methods. It can thus be recommended as a good tool for problem (1.2). An example concludes the paper.

2 Smooth test of normality in the ARMA context

Let $\Phi(.)$ be the cumulative distribution function of the $N(0,1)$ distribution with density $\phi(.)$ and let $U_t = 2\Phi(\epsilon_t/\sigma) - 1$ with density $g(.)$. Under H_0 of (1.2), the U_t's are *i.i.d.* U[-1, 1] so that (1.2) reduces to testing $g(u) = 1/2$ on [-1, 1]. The ϵ_t's are unobserved so the test must be based on residuals. Since the process $(Y_t, t \in \mathbb{Z})$ is invertible, we have

$$\epsilon_t = -\sum_{r=0}^{\infty} \delta_r Y_{t-r} \tag{2.1}$$

where the δ_r's are functions of $\boldsymbol{\theta}$ and $\boldsymbol{\varphi}$ (see (A.2), (A.3) of Appendix A). Let $\hat{\delta}_r$ be the Gaussian maximum likelihood estimator (*m.l.e.*) of δ_r under (1.2), obtained by plugging in the *m.l.e.* $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\varphi}}$ under H_0. We define the residuals of the fitted ARMA model by

$$\hat{\epsilon}_t = -\sum_{r=0}^{\infty} \hat{\delta}_r Y_{t-r}. \tag{2.2}$$

In practice, some scheme must be used to compute these residuals, for example by taking $Y_t = 0$ if $t < 1$. Note that other residuals can be defined for ARMA models (see Brockwell and Davis (1991), Section 9.4) but the definition above is convenient for the following derivation. Consider $\hat{U}_t = 2\Phi(\hat{\epsilon}_t/\hat{\sigma}) - 1$, $t = 1, ..., T$. Let $\{L_k(.), k \geq 0\}$ be the normalized (over [-1, 1]) Legendre polynomials (Sansone (1959)) with $L_0(.) \equiv 1$ satisfying

$$\frac{1}{2}\int_{-1}^{1} L_k(x)L_j(x)dx = 1 \text{ if } k = j \text{ and } 0 \text{ otherwise.} \tag{2.3}$$

For some integer $K \geq 1$, consider the density defined on [-1, 1] by

$$g_K(u; \boldsymbol{\omega}) = c(\boldsymbol{\omega}) \exp\left\{\sum_{k=1}^{K} \omega_k L_k(u)\right\}, \tag{2.4}$$

where $c(\boldsymbol{\omega})$ is a normalizing constant such that $c(\mathbf{0}) = 1/2$. In the smooth test paradigm, (2.4) is the K-th order alternative with $g_K(.; 0)$ being the U[-1, 1] density. Thus, if $g(u)$ can be approximated by (2.4), (1.2) reduces to testing H_0: $\boldsymbol{\omega} = \mathbf{0}$. For this, we use the following route. Let $\boldsymbol{L}_t = (L_1(U_t), ..., L_K(U_t))^{\mathsf{T}}$, $\hat{\boldsymbol{L}}_t = (L_1(\hat{U}_t), ..., L_K(\hat{U}_t))^{\mathsf{T}}$ and

$$\overline{\hat{\boldsymbol{L}}} = T^{-1} \sum_{t=1}^{T} \hat{\boldsymbol{L}}_t. \tag{2.5}$$

Under H_0, $\boldsymbol{L}_t$ has mean $\mathbf{0}$ and covariance matrix $\boldsymbol{I}_K$, the K-th order identity matrix. Under (2.4), these moments will differ and (2.5) can be used to capture departures from the U[-1, 1] in the "direction" of $g_K(.; \boldsymbol{\omega})$. This suggests as a test statistic a quadratic form in $\overline{\hat{\boldsymbol{L}}}$. To complete the test, we need the null asymptotic distribution of (2.5). This is given in the following theorem.

Theoreme 2.1. *Consider the causal and invertible ARMA(p, q) process of (1.1) where we assume* $1 - \varphi_1 z - ... - \varphi_p z^p$ *and* $1 + \theta_1 z + ... + \theta_q z^q$ *have no common zeroes. Under* H_0*, we have*

$$\sqrt{T}\overline{\hat{\boldsymbol{L}}} \xrightarrow{L} N_K\left(\boldsymbol{0}, \boldsymbol{I}_K - \frac{1}{2}\boldsymbol{b}_K\boldsymbol{b}_K^{\mathsf{T}}\right) \tag{2.6}$$

where $\boldsymbol{b}_K = (b_1, ..., b_K)^{\mathsf{T}}$*, with* $b_k = \int_{\mathbb{R}} L_k(2\Phi(x) - 1)x^2\phi(x)dx$*. Hence, the smooth test statistic*

$$\mathcal{R}_K = T\overline{\hat{\boldsymbol{L}}}^{\mathsf{T}}\left(\boldsymbol{I}_K - \frac{1}{2}\boldsymbol{b}_K\boldsymbol{b}_K^{\mathsf{T}}\right)^{-1}\overline{\hat{\boldsymbol{L}}} \xrightarrow{L} \chi^2_K.$$

Proof. We present an outline of the argument. More details are given in the appendices and in Ducharme and Lafaye de Micheaux (2002). Let

$$\mathcal{I}_{\boldsymbol{\beta}} = Var\left[\frac{\partial}{\partial\boldsymbol{\beta}}Log\left(\frac{1}{\sigma}\phi\left(\frac{\epsilon_t}{\sigma}\right)\right)\right]$$

be Fisher's information matrix for $\boldsymbol{\beta}$. From standard results (see Gouriéroux and Monfort (1995), p.325), we have,

$$\sqrt{T}\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathcal{I}_{\boldsymbol{\beta}}^{-1}\frac{\partial}{\partial\boldsymbol{\beta}}\left[Log\left(\frac{1}{\sigma}\phi\left(\frac{\epsilon_t}{\sigma}\right)\right)\right] + o_P(1).$$

Since $\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) = O_P(T^{-1/2})$, a Taylor expansion yields

$$\sqrt{T}\overline{\hat{\boldsymbol{L}}} = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\boldsymbol{L}_t + \left[\frac{1}{T}\sum_{t=1}^{T}\frac{\partial}{\partial\boldsymbol{\beta}}\boldsymbol{L}_t\right]\sqrt{T}\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) + o_P(1). \tag{2.7}$$

The first term on the right hand side of (2.7) converges to a $N_K(\boldsymbol{0}, \boldsymbol{I}_K)$. Moreover, it is shown in Appendix A that

$$\left[\frac{1}{T}\sum_{t=1}^{T}\frac{\partial}{\partial\boldsymbol{\beta}}\boldsymbol{L}_t\right] \xrightarrow{P} \left[\underset{K\times(p+q)}{\boldsymbol{0}}, -\frac{1}{\sigma}\boldsymbol{b}_K\right] = -\mathcal{J}_K. \tag{2.8}$$

Hence,

$$\begin{aligned}\sqrt{T}\overline{\hat{\boldsymbol{L}}} &= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\boldsymbol{L}_t - \frac{1}{\sqrt{T}}\mathcal{J}_K\mathcal{I}_{\boldsymbol{\beta}}^{-1}\sum_{t=1}^{T}\frac{\partial}{\partial\boldsymbol{\beta}}\left[Log\left(\frac{1}{\sigma}\phi\left(\frac{\epsilon_t}{\sigma}\right)\right)\right] + o_P(1)\\ &= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\boldsymbol{B}\boldsymbol{V}_t + o_P(1),\end{aligned}$$

where $\boldsymbol{B} = (\boldsymbol{I}_K, -\mathcal{J}_K \mathcal{I}_{\boldsymbol{\beta}}^{-1})$ and

$$\boldsymbol{V}_t = \left(\boldsymbol{L}_t^{\mathsf{T}}, \frac{\partial}{\partial \boldsymbol{\beta}^{\mathsf{T}}} \left[Log \left(\frac{1}{\sigma} \phi \left(\frac{\epsilon_t}{\sigma} \right) \right) \right] \right)^{\mathsf{T}}.$$

From Appendix B, it follows that, $E(\boldsymbol{V}_t) = \boldsymbol{0}$ and $Var(\boldsymbol{B}\boldsymbol{V}_t) = \boldsymbol{I}_K - \boldsymbol{b}_K \boldsymbol{b}_K^{\mathsf{T}}/2$. The central limit theorem yields (2.6). □

It is possible to write $\mathcal{R}_K$ in a form that makes it easy to use. A Cholesky decomposition of $\left(\boldsymbol{I}_K - \boldsymbol{b}_K \boldsymbol{b}_K^{\mathsf{T}}/2\right)$ yields $\left(\boldsymbol{I}_K - \boldsymbol{b}_K \boldsymbol{b}_K^{\mathsf{T}}/2\right)^{-1} = \boldsymbol{P}\boldsymbol{P}^{\mathsf{T}}$ with $\boldsymbol{P} = (p_{ij})$, an upper triangular matrix. Some algebra gives $p_{ij} = 0$ if $i > j$, while

$$p_{ii} = \sqrt{\frac{2 - \sum_{k=1}^{i-1} b_k^2}{2 - \sum_{k=1}^{i} b_k^2}} \text{ and } p_{ij} = \frac{b_i b_j}{\sqrt{\left(2 - \sum_{k=1}^{j-1} b_k^2\right)\left(2 - \sum_{k=1}^{j} b_k^2\right)}} \text{ if } j > i.$$

Thus

$$\mathcal{R}_K = \sum_{k=1}^{K} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} L_k^*(\hat{U}_t) \right)^2,$$

where

$$L_k^*(\hat{U}_t) = \sum_{l=1}^{k} p_{lk} L_l(\hat{U}_t). \tag{2.9}$$

Numerical integration gives $(b_2, b_4, ..., b_{10}) = (1.23281, 0.521125, 0.304514, 0.205589, 0.150771)$ with $b_k = 0$ if k is odd. This yields the first ten "modified" Legendre polynomials

$$\begin{aligned}
&L_1^*(u) = 1.73u,\\
&L_2^*(u) = 6.85u^2 - 2.28,\\
&L_3^*(u) = 6.61u^3 - 3.97u,\\
&L_4^*(u) = 19.91u^4 - 10.26u^2 - 0.56,\\
&L_5^*(u) = 26.12u^5 - 29.02u^3 + 6.22u,\\
&L_6^*(u) = 69.84u^6 - 81.84u^4 + 28.36u^2 - 3.06,\\
&L_7^*(u) = 103.84u^7 - 167.75u^5 + 76.25u^3 - 8.47u,\\
&L_8^*(u) = 260.07u^8 - 450.18u^6 + 247.18u^4 - 38.73u^2 - 1.11,\\
&L_9^*(u) = 413.92u^9 - 876.55u^7 + 613.58u^5 - 157.33u^3 + 10.73u,\\
&L_{10}^*(u) = 994.51u^{10} - 2250.43u^8 + 1782.83u^6 - 569.92u^4 + 67.54u^2 - 3.58.
\end{aligned}$$

Remark 2.1. *Theorem 2.1 shows that we can slightly extend the result of Pierce and Gray (1985) and state that neither the estimation of* $\boldsymbol{\varphi}$ *and* $\boldsymbol{\theta}$ *nor the dependence of the* Y_t*'s has any asymptotic impact on a smooth test of (1.2) in the ARMA context. In pre-asymptotic situations, these elements and the complexity of the model will affect the null distribution of* $\mathcal{R}_K$*. This will be further explored in simulations of Section 4.*

Remark 2.2. *Each term* $T^{-1}\left(\sum_{t=1}^{T} L_k^*(\hat{U}_t)\right)^2$ *is a component of the test statistic and has an asymptotic* χ_1^2 *distribution under* H_0*. When the null hypothesis is rejected, some of these components will be large. The simple structure of the first few polynomials in (2.9) helps in understanding what aspects of the normal are not supported by the data. For example, the first component detects departure from symmetry under* H_0 *in the "direction" of asymmetry. This diagnostic analysis must be undertaken with some care however; see Henze (1997) for details.*

Remark 2.3. *The above methodology can in principle be applied to other distributions than the normal. For location-scale densities, one needs to replace the normal distribution in the definition of* U_t *and follow the derivation using the new null density. The structure of* $\mathcal{R}_K$ *will be similar to what is obtained above but the modified Legendre polynomials will change. For distributions with a shape parameter, the statistic is more complex since the coefficients of these polynomials will in general depend on this unknown shape parameter that must be estimated.*

3 Choosing the order K of the alternative

Before applying the test strategy of Section 2, one must choose the value of K. Ideally, this choice should be made so that members of the embedding family $g_K(.;\boldsymbol{\omega})$ of (2.4) provide a good approximation to any plausible density $g(.)$ of U_t under the alternative. If K is too small, this approximation may be crude and the test loses power. If K is too large, power dilution can occur since $g_K(.;\boldsymbol{\omega})$ encompasses unnecessary "directions".

In practice, the user has only, at best, a qualitative idea of the plausible alternatives and no specific value of K emerges naturally. In the *i.i.d.* case, some authors (Rayner and Best (1989)) argue that, as a rule of thumb, one can use a trade-off value of K between 2 and 4.

Recently, Ledwina (1994) and Kallenberg and Ledwina (1997a,b) have proposed and explored for *i.i.d.* data a method to choose adaptively a value for K. At the first step, Schwarz (1978)'s criterion is used to choose the value $\hat{K}$ that seems best in view of the data at hand. The smooth test strategy is then applied using the statistic $\mathcal{R}_{\hat{K}}$. Extensive simulations have shown that, even for small sample sizes, this so-called "data driven smooth test" can yield power close to what could be obtained if one knew the true form of the alternative and had chosen the best value of K accordingly.

So far, this approach has been investigated for *i.i.d.* data only but it can be extended to the ARMA context. Choose two integers $1 \leq d \leq D$ and consider the set of statistics $(\mathcal{R}_d, ..., \mathcal{R}_D)$. We seek a rule that will select a good $\mathcal{R}_K$ in this set. Write

$$\hat{K} = \min \left[\underset{d \leq s \leq D}{Argmax} \{ \mathcal{R}_s - sLog(T) \} \right] \tag{3.1}$$

and denote $\mathcal{R}_{\hat{K}}(d)$, the test statistic $\mathcal{R}_{\hat{K}}$ selected by (3.1) in $(\mathcal{R}_d, ..., \mathcal{R}_D)$.

Theoreme 3.1. *Under H_0, $\hat{K} \longrightarrow d$ in probability and thus, $\mathcal{R}_{\hat{K}}(d)$ is asymptotically χ^2_d .*

Proof. Set $e_k = (k-d)LogT$. For $k \geq d$, $P(\hat{K} = k) \leq P(\mathcal{R}_k > e_k)$. Now, since each $\mathcal{R}_k$ is asymptotically χ^2_k under H_0, as T increases,

$$P(\mathcal{R}_k > e_k) \longrightarrow 0,$$

when k $>$ d. It follows that $P(\hat{K} = d) = 1 - P(\hat{K} \geq d+1) \longrightarrow 1$. □

For finite sample sizes, the asymptotic null distribution of Theorem 3.1 may not provide a good approximation to that of $\mathcal{R}_{\hat{K}}(d)$ since there is a positive probability that $\hat{K} \geq d+1$. A simple correction has been developed by Janic-Wroblewska and Ledwina (2000) when $d = 1$ (*i.i.d.* data). Because of the asymptotic independence between the components of $\mathcal{R}_k$, this correction can easily be extended to $d > 1$ and to the present ARMA context. A direct application of the argument in their Section 4 leads to the following approximation, which can be solved for x by numerical integration

$$P(\mathcal{R}_{\hat{K}}(d) \leq x) \approx P(\chi_d^2 \leq x) P(\chi_1^2 \leq Log(T)) + \int_{Log(T)}^{x} P(\chi_d^2 < x - z) \frac{1}{\sqrt{2\pi z}} e^{-z/2} dz. \tag{3.2}$$

Some quantiles corrected through (3.2) are listed in Table 3.1.

	T	a = 0.10	a = 0.05	a = 0.01
d = 1	50	3.692	5.410	8.805
	100	3.275	5.201	8.703
	200	3.057	4.751	8.590
d = 2	50	5.466	7.137	10.807
	100	5.262	6.972	10.684
	200	5.043	6.796	10.558

Table 3.1: Some quantiles obtained from approximation (3.2)

One may have the feeling that this data driven approach replaces the problem of selecting K with that of selecting d and D. To answer this, Kallenberg and Ledwina (1997a,b) have studied a version of the above procedure where D is allowed to increase with T. In the *i.i.d.* case, they obtain rates connecting these quantities. These rates are theoretically interesting but do not help in practice in selecting a value for D. To get more insight, they have conducted extensive simulations. It turns out that the power levels off rapidly as D increases and there is little to be gained by choosing D much greater than 10. As for the choice of d, again Kallenberg and Ledwina (1997a) briefly discuss this problem where it emerges that in their context $d = 1$ or 2 appears reasonable. In the simulation study of the next section we use both these values of d and take $D = 10$.

In closing this section, note that, by plotting $g_{\hat{K}}(.; \hat{\boldsymbol{\omega}})$ where $\hat{\boldsymbol{\omega}}$ is an estimate of $\boldsymbol{\omega}$, one can get an idea of the true shape of the density when the null hypothesis has been rejected. This can be helpful in finding a more appropriate distribution for the innovations.

4 Simulation Results

To get an idea of the behavior of our test statistics as compared to some competitors, a simulation study was conducted. Samples $\{Y_t, t = 1, ..., T\}$ from various ARMA(p, q) models were generated with the innovations arising, in the first part of the simulation, from the normal distribution and, in the second, from various alternatives. For each sample, we estimated the parameters of the model and computed test statistics. From there, we obtained approximations to their level and power. All programs are written in Fortran 77. The subroutines listed below are from the Numerical Algorithms Group (NAG) MARK 16 Fortran library.

4.1 Levels

The first part of the simulation study was designed to see if the critical values obtained from the asymptotic χ^2 or from (3.2) can be relied upon in finite samples. We took $T = 50$, 100 and 200 and restricted attention to the models MA(2), AR(2), ARMA(1,2), ARMA(2,1) and ARMA(2, 2). To generate ARMA(p,q) samples with Gaussian innovations, we used subroutine G05EGF and G05EWF. These samples were submitted to subroutine G13DCF that returns estimates of the parameters of the model as well as residuals. The definition of these residuals, given at equation (9.4.1) in Brockett et al. (1988), differs from (2.2) but their numerical values are almost identical. These residuals were then submitted to the various tests. The actual level of each test was computed for nominal level $\alpha = 0.10$ and 0.05.

Regarding the parameter $\boldsymbol{\beta}$, note that our test statistics are in theory invariant to the choice of σ and we took $\sigma = 1$. Numerically, this invariance holds approximately because of the stopping rule in G13DCF. But the finite distribution of our test statistics depends on the values of $\boldsymbol{\theta}$ and $\boldsymbol{\varphi}$. To explore this, we have proceeded as follows. First, causality requires that, if $p = 1$, $\varphi_1 \in]-1, 1[$ while if $p = 2$, $\boldsymbol{\varphi}$ must be in the region $\triangle_{\boldsymbol{\varphi}} = \{(\varphi_1, \varphi_2) | \varphi_1 + \varphi_2 < 1, \varphi_2 - \varphi_1 < 1, |\varphi_2| < 1\}$ (Brockett et al. (1988), p. 110, ex.3.2). Similarly, invertibility implies that if $q = 1$, $\theta_1 \in]-1, 1[$ while if $q = 2$, $\boldsymbol{\theta}$ must be in $\triangledown_{\boldsymbol{\theta}} = \{(-\theta_1, -\theta_2) | \theta_1 + \theta_2 < 1, \theta_2 - \theta_1 < 1 \text{ and } |\theta_2| < 1\}$. In addition, the polynomials $1 - \varphi_1 z$ when $p = 1$ and $1 - \varphi_1 z - \varphi_2 z^2$ when $p = 2$ must have no common zeroes with $1 + \theta_1 z$ when q = 1 and $1 + \theta_1 z + \theta_2 z^2$ when q = 2.

For the AR(2) model, we have taken the values of $\boldsymbol{\varphi}$ in the grid of 64 points $\{(-2.0+0.25j, -0.9+0.25k) \in \triangle_{\boldsymbol{\varphi}} | j, k \geq 0\}$. A similar grid was used for the MA(2). This makes it possible to see whether the tests maintain the proper critical level over a large section of the parameter space. For the ARMA(1, 2), the grid over $\bigtriangledown_{\boldsymbol{\theta}}$ was reduced to $\{(-2.0+0.40j, -0.9+0.40k) \in \bigtriangledown_{\boldsymbol{\theta}} | j, k \geq 0\}$ while $\varphi_1 = -0.9 + 0.2j, j = 0, ..., 9$. This gives a set of 250 points on the parameter space of $(\varphi_1, \boldsymbol{\theta})$. For the ARMA(2, 1) model, the same was done with $\boldsymbol{\varphi}$ and θ_1 instead. Finally, for the ARMA (2, 2) model, points $(\boldsymbol{\varphi}, \boldsymbol{\theta})$ satisfying the "no common zeroes" condition were taken in $\{(-1.95+0.45j, -0.85+0.45k) \in \triangle_{\boldsymbol{\varphi}} | j, k \geq 0\} \cup \{-(-1.95+0.45j, -0.95+0.45k) \in \bigtriangledown_{\boldsymbol{\theta}} | j, k \geq 0\}$. This yields 294 $(\boldsymbol{\varphi}, \boldsymbol{\theta})$ parameter points. For each of these parameter points, 10000 samples of size T were generated as described above.

To summarize the results, the following approach was adopted. A 95% confidence interval for the true level when $\alpha = 0.10$ is (0.094, 0.106). Similarly, for $\alpha = 0.05$, 95% of the p-values are expected in the interval (0.046, 0.054). Thus the range of possible p-values was divided in 5 sub-intervals. For $\alpha = 10\%$, these are $I_1 = (0, 0.085)$, $I_2 = [0.085, 0.094)$, $I_3 = [0.094, 0.106)$, $I_4 = [0.106, 0.115)$ and $I_5 = [0.115, 1]$. For $\alpha = 0.05$, $I_1 = (0, 0.035)$, $I_2 = [0.035, 0.046)$, $I_3 = [0.046, 0.054)$, $I_4 = [0.054, 0.065)$ and $I_5 = [0.065, 1]$. For each model, the percentage of p-values in each interval was recorded. Table 4.1 reports the results for statistics $\mathcal{R}_3$ and $\mathcal{R}_{\hat{K}}(2)$ which, as discussed in Section 3, are representative of the two schools of thought for the choice of K. The results for the AR(2) and ARMA(2,1) models being similar to those of the MA(2) and ARMA(1,2) respectively, are omitted for brevity (see Ducharme and Lafaye de Micheaux (2002) for more complete results).

Insert Table 4.1 about here

The actual levels for $\mathcal{R}_3$ are concentrated on I_1, I_2 and I_3. The mode of the distribution is generally located on I_2 for $T = 50$ and is shifted to I_3 as T increases. This lead, at worst, to slightly conservative tests. To appreciate this, the last column of Table 4.1 gives the smallest p-value recorded over the parameter points. For $\mathcal{R}_{\hat{K}}(2)$, the distribution is concentrated on I_2, I_3 and I_4 with, in all cases, a mode centered on I_3. For this statistic, the minimal p-values are also closer to the nominal level (no maximal p-value was very far from the upper bound of I_4). Thus correction (3.2) works nicely, at least for the cases considered here.

We also investigated what areas of the parameter space give p-values in I_1. Intuitively, one expects these points to be near the boundary. However, the pattern that emerges, which is very similar for both $\mathcal{R}_3$, and $\mathcal{R}_{\hat{K}}(2)$, is more precise. For AR(2) models, these points correspond mainly to positive (φ_1, φ_2) close to the right boundary of $\triangle_{\boldsymbol{\varphi}}$ and, to a lesser degree, to those with positive φ_1 and negative φ_2 but again close to that boundary. For MA(2) models, the situation is reversed, which is not surprising since $\bigtriangledown_{\boldsymbol{\theta}} = -\triangle_{\boldsymbol{\varphi}}$. For ARMA(2, 1), the points giving small p-values correspond to positive (φ_1, φ_2) combined with values of θ_1 close to -1. Again, for ARMA(1, 2) the situation is reversed and small p-values are associated with negative values of (θ_1, θ_2) with a value of φ_1 close to 1. Finally, for the ARMA(2, 2), the points that yield p-values in I_1 are mainly those with positive (φ_1, φ_2) and negative (θ_1, θ_2).

We have also investigated the behavior under H_0 of some other tests that have been recommended in the time series literature for (1.2). We first considered the Anderson-Darling ($\mathcal{AD}$) test (Pierce and Gray (1985)) for case 2 (known mean) used in conjunction with the quantiles given in D'Agostino and Stephens (1986) p. 122. Our simulations show that, for large T this yields valid critical levels. We also studied a variant of the Shapiro-Wilk test known as the Weisberg and Bingham (1975) ($\mathcal{WB}$) test. To adapt this test to our context where the mean is known, the denominator of equation (9.68) of D'Agostino and Stephens (1986) was replaced by $T\hat{\sigma}^2$, where $\hat{\sigma}^2$ is the estimate of σ^2 returned by subroutine G13DCF. Up to the numerical accuracy of procedure G13DCF, this corresponds to the sum of squares of the residuals. Our simulations show that the quantiles for this test can be approximated by Monte Carlo using *i.i.d.* data, although we found no theoretical result supporting this. Thus, we simulated 100000 samples from an ARMA(0,0) model and computed the empirical quantiles. For $T = 50$, 100 and 200, we got, for $\alpha = 10\%$, 0.920, 0.958 and 0.978. For 5%, we found 0.899, 0.947 and 0.973. A third approach, the Jarque and Bera (1987) eq. (5) ($\mathcal{JB}$) test was also investigated. Although developed in the linear regression context, this test has been recommended in the time series literature (see Cromwell et al. (1994); Frances (1998)). A summary of the results for these tests in the ARMA(1, 2) model is given in Table 4.2. Also appearing in this table are the levels of the test based on $\mathcal{R}_{\hat{K}}(1)$ using quantiles derived from (refequation3.2).

Overall, the best tests, according to the criterion of maintaining the proper level throughout the parameter space, are $\mathcal{R}_{\hat{K}}(2)$ followed by $\mathcal{R}_{\hat{K}}(1)$ and then $\mathcal{R}_3$, $\mathcal{AD}$ and $\mathcal{WB}$. In general, the AD test yields distributions of p-values in between those of $\mathcal{R}_3$ and $R_{\hat{K}}(1)$. More troublesome is the fact that this test, as well as the $\mathcal{WB}$ test, may vastly underestimate the intended level, as can be seen by the minimal p-values (last column of Table 4.2) encountered on the grids. Also, there appears to be a problem with the $\mathcal{JB}$ test as the quantiles, obtained from the χ^2_2 approximation, lead to gross error. Further simulations indicate that the convergence to the χ^2_2 is very slow. The $\mathcal{JB}$ statistic is a version of the Bowman and Shenton test statistic that, for *i.i.d.* data, has a notoriously slow convergence. The simulation results in Lutkepohl and Schneider (1989) tend to show that this is also the case for AR(1) and AR(2) models. In view of this problem, we choose to drop from further investigations the $\mathcal{JB}$ test.

Insert Table 4.2 about here

4.2 Power

The second part of the simulation was designed to study the power of our tests and allow comparison with the competitors mentioned above. We restricted attention to *i.i.d.* innovations. We generated samples $\{Y_t, t = 1, ..., T\}$ according to model (1.1) from various alternatives to the normal distribution. These alternatives were taken as the centered version of the densities listed in Table V of Kallenberg and Ledwina (1997b). They comprise a large range of departure from the normal distribution both in skewness, kurtosis and shape.

To generate ARMA(p, q) samples $\{Y_t, t = 1, ..., T\}$ according to model (1.1) with non-Gaussian innovations, we used the random shock method (algorithms IA 1 with $m = 50$ and SA 1 with $M = 200$) of Burn (1987). To allow a proper comparison of the various tests, we used for each model a set of parameters for which the p-values computed in the first part of the simulation were in I_3 for all tests. More precisely we took: ARMA(2, 1): $(\boldsymbol{\varphi}, \theta_1) = (-0.8, -0.1, 0.7)$, ARMA(1, 2): $(\varphi_1, \boldsymbol{\theta}) = (-0.7, 0.4, 0.5)$ and

ARMA(2, 2): $(\boldsymbol{\varphi}, \boldsymbol{\theta}) = (-1.05, -0.4, 0.15, 0.85)$. Also we took $T = 50$ (more complete simulations appear in Ducharme and Lafaye de Micheaux (2002)). For each combination of model and alternative distribution, we generated 10000 samples and performed the various tests. From there, empirical powers were computed.

Table 4.3 presents these empirical powers for the tests $\mathcal{R}_3$, $\mathcal{R}_{\hat{K}}(2)$ and $\mathcal{WB}$ when $\alpha = 10\%$. Similar results were obtained for $\alpha = 5\%$. The tests $\mathcal{R}_3$ and $\mathcal{R}_{\hat{K}}(2)$ behave similarly with, overall, $\mathcal{R}_{\hat{K}}(2)$ being slightly better. Both these tests generally dominate the others. The $\mathcal{AD}$ approach, not shown here, often yields a power that is much lower than these two tests whereas $\mathcal{WB}$ generally lies somewhere in between. For *i.i.d.* data, the $\mathcal{WB}$ test, as a variant of the Shapiro-Wilk test, is considered among the best omnibus tests of normality. In ARMA situations, this does not seem to hold at the same degree.

We have also computed the power of the test based on $\mathcal{R}_{\hat{K}}(1)$. The tabulated results are not presented here for brevity. We found that, for $T = 50$ and symmetric alternatives, the test based on $\mathcal{R}_{\hat{K}}(1)$ yields slightly better power that $\mathcal{R}_{\hat{K}}(2)$. For asymmetric alternatives, the situation is reversed. But for $T = 100$, $\mathcal{R}_{\hat{K}}(2)$ is more powerful almost everywhere. This behavior of $\mathcal{R}_{\hat{K}}(1)$ is explained by the fact that for asymmetric alternatives, $\mathcal{R}_1$ yields little, sometimes trivial, power. Moreover, power as a function of K usually levels off at $K = 3$, and not infrequently at $K = 2$. This empirical observation is behind the rule of thumb stated in Section 3. Thus to have good power, the selection rule with $d = 1$ must give $\hat{K} \geq 3$, which may be difficult. Starting at $d = 2$ gives a better chance that $\hat{K} \geq 3$ when necessary.

In view of the results of these simulations, we recommend the use of $\mathcal{R}_{\hat{K}}(2)$ for testing (1.2) when $E(Y_t)$ is known. The levels are stable over most of the parameter points and close to nominal for moderate samples. Moreover, the power is generally better than that of other tests that have been recommended in the time series literature. Finally, the test is very easy to apply.

Insert Table 4.3 about here

5 An example

In the course of a study to forecast the amount of daily gas required, Shea (1987) has studied a bivariate time series of $T = 366$ points. The first component of this time series pertains to differences in daily temperature between successive days ($\bigtriangledown\tau_t$) and he found, after an iteration process of fitting and diagnostic checking, that the following MA(4) model could be entertained:

$$\bigtriangledown\tau_t = \epsilon_t + 0.07\epsilon_{t-1} - 0.30\epsilon_{t-2} - 0.15\epsilon_{t-3} - 0.20\epsilon_{t-4}.$$

The residual variance is 2.475. All these parameters are obtained by maximizing the Gaussian likelihood so that problem (1.2) is of some importance. Shea does not discuss the normality of the innovations in assessing the fit of this model but rather goes on to find a good model for the bivariate series based on an analysis of the residuals' cross correlation matrix.

An application of our tests yields $\mathcal{R}_3 = 22.85$, with a p-value of 0.00004 while $\mathcal{R}_{\hat{K}}(2) = 22.77$ ($\hat{K} = 2$) yielding a p-value of 0.00003 according to (3.2). Thus, both tests strongly reject the null hypothesis (1.2). A complementary analysis helps understanding what aspect of the Gaussian is not supported by the data. We found $\mathcal{R}_1 = 0.15$ ($p = 0.69$) with a skewness coefficient of 0.13. Thus there is no reason to suspect an asymmetrical distribution for the innovations. On the other hand, we can notice that 9.3% of the absolute standardized residuals are greater than 2.5 and the kurtosis is 4.33. Thus, if the model entertained above is correct, the conclusion that emerges from the present analysis is that the $\bigtriangledown\tau_t$ series could have been generated from innovations with a symmetric distribution having fatter tails than the Gaussian.

ACKNOWLEDGMENTS
The authors would like to thank Dr. B.L. Shea for some insight on subroutine G13DCF of the NAG library and for providing them with the data set used in Section 5.

Appendix A

We show that(2.8) holds under H_0. Assume p and $q > 0$. It suffices to show that

$$\frac{1}{T}\sum_{t=1}^{T}\frac{\partial}{\partial\sigma}L_k(U_t) \xrightarrow{P} E\left[\frac{\partial}{\partial\sigma}L_k(U_t)\right] = -\frac{1}{\sigma}b_k, \tag{A.1.a}$$

$$\frac{1}{T}\sum_{t=1}^{T}\frac{\partial}{\partial\varphi_1}L_k(U_t) \xrightarrow{P} E\left[\frac{\partial}{\partial\varphi_1}L_k(U_t)\right] = 0, \tag{A.1.b}$$

$$\frac{1}{T}\sum_{t=1}^{T}\frac{\partial}{\partial\theta_1}L_k(U_t) \xrightarrow{P} E\left[\frac{\partial}{\partial\theta_1}L_k(U_t)\right] = 0. \tag{A.1.c}$$

First,

$$\frac{\partial}{\partial\sigma}L_k(U_t) = -\frac{2\epsilon_t}{\sigma^2}\phi\left(\frac{\epsilon_t}{\sigma}\right)L_k'(x)|_{x=2\Phi\left(\frac{\epsilon_t}{\sigma}\right)-1} = -\frac{\epsilon_t}{\sigma^2}w\left(\frac{\epsilon_t}{\sigma}\right) \text{ say.}$$

The law of large numbers yields (A.1.a). For (A.1.b), define for $r \geq 0$,

$$B_{r-1} = \frac{\partial}{\partial\varphi_1}\delta_r(\boldsymbol{\theta},\boldsymbol{\varphi}),$$

where, setting $\varphi_0 = -1$, $\gamma_0 = \theta_0 = 1$, we have

$$\delta_r(\boldsymbol{\theta},\boldsymbol{\varphi}) = \delta_r = \sum_{i=0}^{\min(r,p)}\varphi_i\gamma_{r-i} \quad r \geq 0, \tag{A.2}$$

$$\gamma_r = -\sum_{i=1}^{\min(r,q)}\gamma_{r-i}\theta_i \quad r \geq 1. \tag{A.3}$$

Obviously $B_{r-1} = \gamma_{r-1}$ when $r \geq 1$. For $r \geq q$, from Brockwell and Davis (1991), p.107,

$$\gamma_r = \sum_{i=1}^{j}\sum_{n=0}^{r_i-1}c_{in}r^n\alpha_i^{-r}$$

for some constants c_{in} and where the α_i's are the j distinct roots of $1+\theta_1 z + \ldots + \theta_q z^q$ and r_i is the multiplicity of α_i, $i = 1,\ldots,j$. Thus, when $r \geq q+1$,

$$B_{r-1} = \sum_{i=1}^{j}\sum_{n=0}^{r_i-1}c_{in}(r-1)^n\alpha_i^{-r+1}. \tag{A.4}$$

If $(X_t, t \in \mathbb{Z})$ is a weak stationary process such that $Cov(X_t, X_{t+h}) \to 0$ as $h \to \infty$, then $\overline{X}_T \xrightarrow{P} E(X_t)$. We apply this result with $X_t = \partial L_k(U_t)/\partial\varphi_1$. From (1.1) and (2.1), we have

$$X_t = \frac{1}{\sigma} w\left(\frac{\epsilon_t}{\sigma}\right) \frac{\partial}{\partial\varphi_1}\epsilon_t = -\frac{1}{\sigma} w\left(\frac{\epsilon_t}{\sigma}\right) \left(Y_{t-1} - \sum_{r=0}^{\infty}\left(\frac{\partial}{\partial\varphi_1}\delta_r\right)\boldsymbol{\theta}^{\mathsf{T}}\boldsymbol{Y}^{(q)}_{t-1-r}\right). \tag{A.5}$$

Thus $E(X_t) = 0$. Moreover, $Var(X_t) < \infty$ as shown in Appendix C and it is seen by Lemma C.1 that $Cov(X_t, X_{t+h})$ depends only on h. Thus $(X_t, t \in \mathbb{Z})$ is stationary. We show that $Cov(X_t, X_{t+h}) \to 0$ as $h \to \infty$. From (A.5), for h large, $|Cov(X_t, X_{t+h})| = |d_1| E|w(\epsilon_{t+h}/\sigma)|/\sigma$, where

$$d_1 = E\left[\frac{1}{\sigma} w\left(\frac{\epsilon_t}{\sigma}\right)\left\{Y_{t-1} - \sum_{r=0}^{\infty} B_{r-1}\boldsymbol{\theta}^{\mathsf{T}}\boldsymbol{Y}^{(q)}_{t-1-r}\right\}\left\{Y_{t+h-1} - \sum_{r=0}^{\infty} B_{r-1}\boldsymbol{\theta}^{\mathsf{T}}\boldsymbol{Y}^{(q)}_{t+h-1-r}\right\}\right]. \tag{A.6}$$

But, $|d_1| \le d_2 + \sum_{j=1}^{q}|\theta_j|(d_{3j} + d_{4j}) + \sum_{i=1}^{q}\sum_{j=1}^{q}|\theta_i\theta_j| d_{5ij}$ where

$$d_2 = \left|E\left[\frac{1}{\sigma} w\left(\frac{\epsilon_t}{\sigma}\right) Y_{t-1}Y_{t+h-1}\right]\right|,\ d_{3j} = \sum_{r=0}^{\infty}\left|B_{r-1}E\left[\frac{1}{\sigma} w\left(\frac{\epsilon_t}{\sigma}\right) Y_{t-r-j}Y_{t+h-1}\right]\right|,$$

$$d_{4j} = \sum_{r=0}^{\infty}\left|B_{r-1}E\left[\frac{1}{\sigma} w\left(\frac{\epsilon_t}{\sigma}\right) Y_{t-1}Y_{t+h-r-j}\right]\right|$$

and

$$d_{5ij} = \sum_{r=0}^{\infty}\sum_{r'=1}^{\infty}\left|B_{r-1}B_{r'-1}E\left[\frac{1}{\sigma} w\left(\frac{\epsilon_t}{\sigma}\right) Y_{t-r-i}Y_{t+h-r'-j}\right]\right|.$$

It can be shown that d_2, d_{3j}, d_{4j} and $d_{5ij} \to 0$ when $h \to \infty$. Proof for d_{4j}, which is typical, is sketched in Appendix D. This yields (A.1.b).

As for (A.1.c), let $A_r = \frac{\partial}{\partial\theta_1}\delta_r(\boldsymbol{\theta}, \boldsymbol{\varphi})$. From (A.3), we obtain, for $r \ge q$, the system

$$\begin{cases} \gamma'_r + \theta_1\gamma'_{r-1} + \dots + \theta_q\gamma'_{r-q} = -\gamma_{r-1} \\ \gamma_r + \theta_1\gamma_{r-1} + \dots + \theta_q\gamma_{r-q} = 0 \end{cases}$$

from which we find

$$0 = \sum_{j=0}^{q}\theta_j\left(-\sum_{i=0}^{q}\theta_i\gamma'_{r-j-i+1}\right) = \sum_{h=0}^{2q} a_h\gamma'_{r-h+1} \text{ where } a_h = \sum_{\substack{i+j=h \\ 0\le i,j\le q}}\theta_i\theta_j, \text{ for all } r \ge 2q-1.$$

Again from Brockwell and Davis (1991), p.107, we have, for some constants d_{in}

$$\gamma'_r = \sum_{i=1}^{j} \sum_{n=0}^{s_i-1} d_{in} r^n \beta_i^{-r}$$

where the β_i's are the j distinct roots (with multiplicity s_i) of $1 + a_1 z + a_2 z^2 + ... + a_{2q} z^{2q}$. Now

$$\left(\sum_{i=0}^{q} \theta_i z \right)^2 = \sum_{h=0}^{2q} \left(\sum_{\substack{i+j=h \\ 0 \le i,j \le q}} \theta_i \theta_j \right) z^h = \sum_{h=0}^{2q} a_h z^h$$

where $a_0 = \theta_0^2 = 1$. This shows that the roots of $1 + a_1 z + a_2 z^2 + ... + a_{2q} z^{2q}$ are exactly the same than that of $1 + \theta_1 z + \theta_2 z^2 + ... + \theta_q z^q$, apart from the multiplicity. Thus, we obtain

$$A_r = \sum_{l=0}^{p} \sum_{i=1}^{j} \sum_{n=0}^{s_i-1} d_{in} (r-l)^n \alpha_i^{-(r-l)} \varphi_l, \text{ for all } r \ge \max(2q, p). \tag{A.7}$$

By the same argument, using A_r of (A.7) instead of B_{r-1} of (A.4), we get (A.1.c).

Appendix B

We show that $E(\boldsymbol{V}_t) = 0$ and $Var(\boldsymbol{B}\boldsymbol{V}_t) = \boldsymbol{I}_K - \boldsymbol{b}_K \boldsymbol{b}_K^{\mathsf{T}}/2$. In view of (1.1) and (2.1),

$$\frac{\partial}{\partial \boldsymbol{\varphi}} Log \left(\frac{1}{\sigma} \phi \left(\frac{\epsilon_t}{\sigma} \right) \right) = \frac{\epsilon_t}{\sigma^2} \left[\boldsymbol{Y}_{t-1}^{(p)} - \sum_{r=0}^{\infty} \left(\frac{\partial}{\partial \boldsymbol{\varphi}} \delta_r \right) \boldsymbol{\theta}^{\mathsf{T}} \boldsymbol{Y}_{t-1-r}^{(q)} \right],$$

$$\frac{\partial}{\partial \boldsymbol{\theta}} Log \left(\frac{1}{\sigma} \phi \left(\frac{\epsilon_t}{\sigma} \right) \right) = \frac{\epsilon_t}{\sigma^2} \left[\boldsymbol{\epsilon}_{t-1}^{(q)} - \sum_{r=0}^{\infty} \left(\frac{\partial}{\partial \boldsymbol{\theta}} \delta_r \right) \boldsymbol{\theta}^{\mathsf{T}} \boldsymbol{Y}_{t-1-r}^{(q)} \right]$$

and

$$\frac{\partial}{\partial \sigma} Log \left(\frac{1}{\sigma} \phi \left(\frac{\epsilon_t}{\sigma} \right) \right) = \frac{1}{\sigma} \left(\left(\frac{\epsilon_t}{\sigma} \right)^2 - 1 \right).$$

It follows that $E(\boldsymbol{V}_t) = 0$ under H_0. Moreover, under H_0, $Var(\boldsymbol{L}_t) = \boldsymbol{I}_K$. Thus,

$$Cov\left(\boldsymbol{L}_t, \frac{\partial}{\partial \boldsymbol{\beta}} Log\left(\frac{1}{\sigma}\phi\left(\frac{\epsilon_t}{\sigma}\right)\right)\right)^{\mathsf{T}} = \left[\begin{array}{c} \mathbf{0} \\ \mathbf{0} \\ \frac{1}{\sigma}\boldsymbol{b}_K^{\mathsf{T}} \end{array}\right] = \mathcal{J}_K^{\mathsf{T}}.$$

Finally, $Var\left(\frac{\partial}{\partial \boldsymbol{\beta}} Log\left(\frac{1}{\sigma}\phi\left(\frac{\epsilon_t}{\sigma}\right)\right)\right) = \mathcal{I}_\beta = \left[\begin{array}{cc} \boldsymbol{C} & \mathbf{0} \\ \mathbf{0} & \frac{2}{\sigma^2} \end{array}\right]$, for some matrix $\boldsymbol{C}$ whose exact expression is not needed. Thus

$$Var(\boldsymbol{V}_t) = \left[\begin{array}{cc} \boldsymbol{I}_K & \mathcal{I}_K \\ \mathcal{I}_K^{\mathsf{T}} & \mathcal{I}_\beta \end{array}\right].$$

Appendix C

We show that $Var(X_t) < \infty$. Without loss of generality, set $\sigma = 1$. This will be assumed here and in the next appendix. Since Y_t is causal, we can write $Y_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$ and from (A.5)

$$Var(X_t) = E(w(\epsilon_t))^2 E\left(Y_{t-1} - \sum_{r=0}^{\infty} B_{r-1}\boldsymbol{\theta}^{\mathsf{T}}\boldsymbol{Y}_{t-1-r}^{(q)}\right)^2 = E(w(\epsilon_t))^2 E\left(\sum_{h=1}^{\infty} d_h \epsilon_{t-h}\right)^2$$

where $d_h = \psi_{h-1} - \sum_{\substack{r+j+l=h \\ 1 \le j \le q \\ 0 \le r,l \le h-1}} \psi_l \gamma_{r-1}\theta_j$. We now need the following lemma.

Lemma C.1. *If the ARMA process (1.1) is causal and invertible, then* $\sum_{h=1}^{\infty} |d_h| < \infty$.

Proof. From (A.3), $\sum_{\substack{r+j+l=h \\ 1 \le j \le q \\ 0 \le r,l \le h-1}} \psi_l \gamma_{r-1}\theta_j = \sum_{k=1}^{h} \psi_{h-k} \left(\sum_{\substack{r+j=k \\ 1 \le j \le q \\ 0 \le r \le h-1}} \gamma_{r-1}\theta_j \right) = -\sum_{k=1}^{h} \psi_{h-k}\gamma_{k-1}$.
We have also $\sum_{h=1}^{\infty}\sum_{k=1}^{h} |\psi_{h-k}\gamma_{k-1}| = \sum_{h=0}^{\infty}\sum_{k=0}^{h} |\psi_{h-k}\gamma_k| = \sum_{k=0}^{\infty}|\gamma_k| \sum_{h=0}^{\infty}|\psi_h|$. Thus, $\sum_{h=1}^{\infty}|d_h| \le \sum_{h=1}^{\infty}|\psi_{h-1}| + \sum_{h=1}^{\infty}\sum_{k=1}^{h}|\psi_{h-k}\gamma_{k-1}| = \sum_{h=0}^{\infty}|\psi_h| \left(\sum_{k=0}^{\infty}|\gamma_k| + 1\right)$. But from Brockwell and Davis (1991), p.87, $\sum_{k=0}^{\infty}|\gamma_k|$ is finite. Since under the assumption of Theorem 2.1, $\sum_{j=0}^{\infty}|\psi_j| < \infty$ the lemma follows. $\square$

From this lemma, we conclude that $E\left[\sum_{h=1}^{\infty} d_h \epsilon_{t-h}\right]^2 = \sum_{h=1}^{\infty} d_h^2 < \infty$. Since

$$E(w(\epsilon_t))^2 = 4\int (L_k'(2\Phi(x) - 1))^2 \phi^3(x) dx < \infty,$$

the result follows.

Appendix D

Here we sketch the proof that the typical element d_{4j} of inequality (A.6) vanishes. From $Y_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$ and the fact that the remainder of a convergent series converges toward 0, we have

$$\lim_{h\to\infty} d_{4j} = \lim_{h\to\infty} \sum_{r=0}^{\infty} |B_{r-1} E\left(w(\epsilon_t) Y_{t-1} Y_{t+h-r-j}\right)| \leq \lim_{h\to\infty} |E(w(\epsilon_t))| \sum_{r=0}^{h-j} \left| B_{r-1} \sum_{a=0}^{\infty} \psi_a \psi_{a+h-j-r+1} \right|$$

$$\leq |E(w(\epsilon_t))| \lim_{h\to\infty} \left[\sum_{a=0}^{m-1} |\psi_a| \sum_{r=0}^{a+h-j-m} |B_{r-1}\psi_{a+h-j-r+1}| + \sum_{r=a+h-j-m+1}^{h-j} |B_{r-1}\psi_{a+h-j-r+1}| + \sum_{r=0}^{h-j} |B_{r-1}| \sum_{a=m}^{\infty} |\psi_a \psi_{a+h-j-r+1}| \right] \quad \text{(D.1)}$$

where $m = \max\{p, q+1\} - p$. For the first term in the limit of (D.1), using the expression for B_{r-1} in (A.4) and that of $\psi_{a+h-j-r+1}$ given in Brockwell and Davis (1991) eq. (3.3.6), we have

$$\sum_{r=q+1}^{a+h-j-m} |B_{r-1}\psi_{a+h-j-r+1}| =$$

$$\sum_{r=q+1}^{a+h-j-m} \left| \sum_{b=1}^{k} \sum_{l=0}^{r_b-1} c_{bl} r^l \alpha_a^{-r} \sum_{b'=1}^{k'} \sum_{l'=0}^{r_{b'}-1} \alpha_{b'l'} (a+h-j-r+1)^{l'} \xi_{b'}^{-(a+h-j-r+1)} \right|$$

$$\leq \sum_{b=1}^{k} \sum_{l=0}^{r_b-1} \sum_{b'=1}^{k'} \sum_{l'=0}^{r_{b'}-1} \sum_{d=0}^{l'} \binom{l'}{d} \left\{ |c_{bl}\alpha_{b'l'}| |\xi_{b'}^{-(a+h-j+1)} (a+h-j+1)^{l'-d}| \sum_{r=q+1}^{a+h-j-m} r^{l+d} |\alpha_a|^{-r} |\xi_{b'}|^r \right\}.$$

If $|\xi_{b'}| < |\alpha_a|$, the term in braces $\to 0$ as $h \to \infty$. Let $|\alpha_a| = 1 + \epsilon_1 < |\xi_{b'}| = 1 + \epsilon_2$ with ϵ_1, $\epsilon_2 > 0$.

$$\left| \frac{(a+h-j+1)^{l'-d}}{\xi_{b'}^{(a+h-j+1)}} \right| \sum_{r=q+1}^{a+h-j-m} r^{l+d} |\alpha_a|^{-r} |\xi_{b'}|^r \leq \frac{|a+h-j+1|^{l'-d}}{|\xi_{b'}|^{a+h-j+1}} \sum_{r=0}^{a+h-j+1} r^{l+d} \left(\frac{|\xi_{b'}|}{|\alpha_a|} \right)^r . \quad \text{(D.2)}$$

For all $\epsilon > 0$, there exist a C, C' such that the left-hand side of (D.2) is bounded above by

$$C\frac{|a+h-j+1|^{l'-d}}{|\xi_{b'}|^{a+h-j+1}}\frac{\left(\frac{|\xi_{b'}|}{|\alpha_a|}+\epsilon\right)^{h+a-j+2}-1}{\frac{|\xi_{b'}|}{|\alpha_a|}+\epsilon-1} \leq C'|a+h-j+1|^{l'-d}\left(\frac{\frac{|\xi_{b'}|}{|\alpha_a|}+\epsilon}{|\xi_{b'}|}\right)^{a+h-j+2}. \tag{D.3}$$

In (D.3), take $\epsilon > 0$ smaller than $\epsilon_1(1+\epsilon_2)/(1+\epsilon_1)$. Then the right hand side of (D.3) converges to 0 as $h \to \infty$. This shows that the first term in the limit of (D.1) converges to 0. It follows that the second term also converges toward 0. As for the last term in the limit, a similar argument yields that all terms on the right hand side of (D.1) converge to 0 so that $d_{4j} \to 0$.

References

Anděl, J., 1997. On residual analysis for time series models. Kybernetika (Prague) 33 (2), 161–170.

Beiser, A., 1985. Distributions of $\sqrt{b_1}$ and b_2 for autoregressive errors. Ph.D. thesis, Boston University.

Brockett, P. L., Hinich, M. J., Patterson, D., 1988. Bispectral-based tests for the detection of gaussianity and linearity in time series. Journal of the American Statistical Association 83, 657–664.

Brockwell, P. J., Davis, R. A., 1991. Time series: Theory and Methods, 2nd Edition. Springer-Verlag New York.

Burn, D., 1987. Simulation of stationary time series. Proceedings of the 1987 Winter Simulation Conference , 289–294.

Cromwell, J. B., Labys, W. C., Terraza, M., 1994. Univariate tests for time-series models. Sage Publications Inc, Thousand Oaks, California.

D'Agostino, R. B., Stephens, M. A., 1986. Goodness-of-fit techniques. Statistics: TextBOOKs and Monographs, 68, New York: Marcel Dekker.

Ducharme, G., Lafaye de Micheaux, P., 2002. Goodness-of-fit tests of normality for the innovations in arma models. Tech. rep., Technical report #02-02, Université Montpellier II.

Epps, T. W., 1987. Testing that a stationnary time series is gaussian. The Annals of Statistics 15 (4), 1683–1698.

Frances, P., 1998. Time series models for business and economic forecasting. Cambridge University Press, Cambridge.

Gasser, T., 1975. Goodness-of-fit tests for correlated data. Biometrika 62, 563–570.

Gouriéroux, C., Monfort, A., 1995. Séries temporelles et modèles dynamiques, 2nd Edition. Economica.

Granger, C. W. J., 1976. Tendency towards normality of linear combinations of random variables. Metrika 23 (4), 237–248.

Henze, N., 1997. Do components of smooth tests of fit have diagnostic properties? Metrika , 121–130.

Heuts, R., Rens, S., 1986. Testing normality when observations satisfy a certain low order arma-scheme. Computational Statistics Quarterly 1, 49–60.

Hinich, M. J., 1982. Testing for gaussianity and linearity of a stationary time series. Journal of Time Series Analysis 3 (3), 169–176.

Hipel, K. W., McLeod, A. I., 1994. Time series modelling of water resources and environmental systems. [Elsevier Science Publishing Co., New York; North-Holland Publishing Co., Amsterdam] (New York; Amsterdam).

Janic-Wroblewska, A., Ledwina, T., 2000. Data driven rank test for two-sample problem. Scandinavian Journal of Statistics 27, 281–298.

Jarque, C., Bera, A., 1987. A test for normality of observations and regression residuals. International Statistical Review 55 (2), 163–172.

Kallenberg, W., Ledwina, T., 1997a. Data driven smooth tests for composite hypotheses: comparison of powers. Journal of Statistical Computational Simulation 59 (2), 101–121.

Kallenberg, W., Ledwina, T., 1997b. Data-driven smooth tests when the hypothesis is composite. Journal of The American Statistical Association 92 (439), 1094–1104.

Koul, H. L., Stute, W., 1999. Nonparametric model checks for time series. Ann. Statist. 27 (1), 204–236.

Ledwina, T., 1994. Data-driven version of neyman's smooth test of fit. Journal of The American Statistical association 89 (427), 1000–1005.

Lee, S., Na, S., 2002. On the Bickel-Rosenblatt test for first-order autoregressive models. Statist. Probab. Lett. 56 (1), 23–35.

Lomnicki, Z., 1961. Tests for departure from normality in the case of linear stochastic processes. Metrika 4, 37–62.

Lutkepohl, H., Schneider, W., 1989. Testing for normality of autoregressive time series. Comput.Statistics Quaterly 2, 151–168.

Moore, D. S., 1982. The effect of dependence on chi squared tests of fit. Ann. Statist. 10 (4), 1163–1171.

Neyman, J., 1937. Smooth test for goodness of fit. Skandinavisk Aktuarietidskrift 20, 149–199.

Ojeda, R., Cardoso, J., Moulines, E., 1997. Asymptotically invariant gaussianity test for causal invertible time series. Proc. of IEEE international conference on Acoustics, Speech and Signal Processing 5, 3713–3716.

Pierce, D. A., Gray, R. J., 1985. Goodness-of-fit tests for censored survival data. The Annals of Statistics 13 (2), 552–563.

Rayner, J., Best, D., 1989. Smooth Tests of Goodness-of-Fit. Oxford: Oxford University Press.

Sansone, G., 1959. Orthogonal functions. New York: Interscience.

Schwarz, G., 1978. Estimating the dimension of a model. The Annals of Statistics 6 (2), 461–464.

Shea, B. L., 1987. Estimation of multivariate time series. J. Time Ser. Anal. 8 (1), 95–109.

Weisberg, S., Bingham, C., 1975. An approximate analysis of variance test for non-normality suitable for machine calculation. Technometrics 17, 133–134.

Table 4.1: Distribution (in % of the number of parameter points) of the empirical p-values (based on 10000 replications) for the tests based on $\mathcal{R}_3$ and $\mathcal{R}_{\hat{K}}(2)$ among 5 sub-intervals.

$\mathcal{R}_3$			Observed level					Min
Model	T	α	I_1	I_2	I_3	I_4	I_5	p-level
MA(2) (64 points)	50	5%	18.8	68.8	12.5	0	0	2.76
	100	5%	1.6	50.0	48.4	0	0	3.41
	200	5%	0	9.4	89.1	1.6	0	4.04
	50	10%	23.4	53.1	23.4	0	0	6.49
	100	10%	6.3	20.3	73.4	0	0	7.96
	200	10%	0	7.8	90.6	1.6	0	8.92
ARMA(1,2) (250 points)	50	5%	47.2	46.4	6.4	0	0	2.43
	100	5%	8.0	71.6	20.4	0	0	2.98
	200	5%	0.8	32.4	66.4	0.4	0	3.32
	50	10%	65.6	24.0	10.4	0	0	6.20
	100	10%	21.6	35.2	42.8	0.4	0	6.80
	200	10%	4.0	19.6	75.6	0.8	0	7.42
ARMA(2,2) (294 points)	50	5%	41.2	57.1	1.7	0	0	2.56
	100	5%	5.1	74.1	20.8	0	0	3.09
	200	5%	0.3	27.9	71.8	0	0	3.47
	50	10%	57.8	37.4	4.8	0	0	6.24
	100	10%	21.1	33.7	45.2	0	0	6.88
	200	10%	3.1	18.0	78.6	0.3	0	7.86
$\mathcal{R}_{\hat{K}}(2)$			Observed level					Min
Model	T	α	I_1	I_2	I_3	I_4	I_5	p-level
MA(2) (64 points)	50	5%	0	9.4	46.9	43.7	0	4.12
	100	5%	0	14.1	68.8	17.2	0	4.17
	200	5%	0	7.8	87.5	4.7	0	4.23
	50	10%	6.3	14.1	62.5	17.2	0	7.78
	100	10%	6.3	6.3	81.3	6.3	0	8.19
	200	10%	0	6.3	89.1	4.7	0	8.83
ARMA(1,2) (250 points)	50	5%	0	38.8	46.8	14.4	0	3.53
	100	5%	0	34.8	59.2	6.0	0	3.74
	200	5%	0	23.2	75.2	1.6	0	3.80
	50	10%	24.8	31.6	35.2	8.4	0	7.06
	100	10%	13.2	27.2	57.6	2.0	0	7.33
	200	10%	4.4	18.0	76.0	1.6	0	7.61
ARMA(2,2) (294 points)	50	5%	0	32.0	55.4	12.6	0	3.65
	100	5%	0	31.0	62.9	6.1	0	3.75
	200	5%	0	23.8	75.9	0.3	0	3.89
	50	10%	21.4	30.3	47.0	1.4	0	7.14
	100	10%	11.2	24.8	62.6	1.4	0	7.51
	200	10%	2.7	16.7	80.6	0	0	7.62

Table 4.2: Distribution (in % of the number of parameter points) of the empirical p-values (based on 10000 replications) of various tests for the ARMA(1,2) model. $\mathcal{AD}$=Anderson-Darling, $\mathcal{WB}$=Weisberg-Bingham, $\mathcal{JB}$=Jarque-Bera and $\mathcal{R}_{\hat{K}}(1) = \mathcal{R}_{\hat{K}}$ with $d = 1$.

Test			Observed level					Min
Model	T	α	I_1	I_2	I_3	I_4	I_5	p-level
$\mathcal{AD}$	50	5%	43.8	24.4	20.0	6.8	0	0.54
	100	5%	32.0	34.8	33.2	0	0	0.92
	200	5%	11.6	38.0	49.6	0.8	0	1.50
	50	10%	41.2	16.0	22.8	18.4	1.6	3.38
	100	10%	23.2	24.4	48.0	3.6	0.8	3.93
	200	10%	9.6	13.2	70.0	6.8	0.4	4.65
$\mathcal{WB}$	50	5%	61.2	23.6	15.2	0	0	0.57
	100	5%	39.2	46.0	14.4	0.4	0	0.93
	200	5%	10.8	30.4	56.0	2.8	0	1.60
	50	10%	56.8	16.4	25.6	1.2	0	2.96
	100	10%	37.2	47.2	15.6	0	0	3.55
	200	10%	15.6	36.0	46.0	2.4	0	4.71
$\mathcal{JB}$	50	5%	71.2	28.8	0	0	0	3.13
	100	5%	0.4	99.2	0.4	0	0	3.13
	200	5%	0	85.2	14.4	0.4	0	4.18
	50	10%	100	0	0	0	0	4.96
	100	10%	100	0	0	0	0	5.88
	200	10%	98.8	1.2	0	0	0	7.23
$\mathcal{R}_{\hat{K}}(1)$	50	5%	0	58.0	27.6	14.4	0	3.52
	100	5%	10.0	48.8	39.2	2.0	0	3.32
	200	5%	8.4	33.6	56.4	1.6	0	3.39
	50	10%	43.6	20.8	26.0	9.6	0	4.69
	100	10%	26.0	24.8	48.0	1.2	0	4.81
	200	10%	8.8	16.4	70.4	4.4	0	5.79

Table 4.3: Empirical power (based on 10000 replications with $\alpha = 10\%$) of various tests when $T = 50$. The part above the line in the middle of the table corresponds to symmetric alternatives while those below are skewed. The distributions are ordered according to increasing kurtosis. The ARMA(2,1) model has parameter $(\boldsymbol{\varphi}, \theta_1) = (-0.8, -0.1, 0.7)$, the ARMA(1,2) model has parameter $(\varphi_1, \boldsymbol{\theta}) = (-0.7, 0.4, 0.5)$ while the ARMA(2,2) model has parameter $(\boldsymbol{\varphi}, \boldsymbol{\theta}) = (-1.05, -0.4, 0.15, 0.85)$.

$T = 50$	ARMA(2,1)			ARMA(1,2)			ARMA(2,2)		
Alternatives	$\mathcal{R}_3$	$\mathcal{R}_{\hat{K}}(2)$	$\mathcal{WB}$	$\mathcal{R}_3$	$\mathcal{R}_{\hat{K}}(2)$	$\mathcal{WB}$	$\mathcal{R}_3$	$\mathcal{R}_{\hat{K}}(2)$	$\mathcal{WB}$
SB(0;0.5)	83.19	84.76	28.93	73.16	74.90	22.47	63.85	65.44	17.57
TU(1.5)	66.94	67.96	18.27	57.86	59.43	15.07	49.17	50.50	13.62
TU(0.7)	44.47	45.64	12.42	38.69	39.44	11.02	32.58	33.88	10.76
Logistic(1)	20.74	22.75	19.02	18.97	20.87	17.10	18.59	20.36	17.60
TU(10)	94.64	96.64	83.60	89.57	91.62	74.31	85.14	87.26	65.78
SC(0.05;3)	33.65	37.38	35.98	32.82	35.65	34.40	30.81	34.69	32.37
SC(0.2;5)	96.36	96.77	92.84	94.21	94.72	89.12	92.28	92.96	85.62
SC(0.05;5)	62.33	65.22	63.63	61.43	63.86	61.81	58.90	62.20	59.77
SC(0.05;7)	74.05	76.12	75.32	73.00	75.22	73.89	72.28	74.04	72.50
SU(0;1)	75.96	76.49	66.57	71.73	72.44	62.20	68.60	69.79	59.99
SB(0.533;0.5)	91.09	89.76	59.41	83.62	82.09	48.17	76.29	74.04	36.87
SB(1;1)	53.75	56.94	32.60	45.94	48.41	26.18	42.03	44.17	23.84
LC(0.2;3)	55.58	57.79	29.72	49.58	52.11	26.19	43.88	46.16	22.62
Weibull(2)	28.10	30.72	21.25	25.63	28.66	18.52	24.32	26.85	18.68
LC(0.1;3)	44.10	43.85	35.21	40.62	40.54	31.38	37.00	37.25	27.97
χ^2 (df.=10)	41.41	45.84	34.72	37.80	41.98	31.09	35.04	38.91	29.66
LC(0.05;3)	29.50	31.25	28.28	28.05	29.86	26.18	25.91	28.32	24.17
LC(0.1;5)	96.10	96.00	95.07	93.40	93.03	90.44	91.04	89.90	86.06
SU(-1;2)	37.88	38.92	33.93	34.38	36.29	31.12	33.54	65.40	29.87
χ^2 (df.=4)	76.13	80.54	69.78	71.19	75.76	63.30	65.96	70.30	57.70
LC(0.05;5)	81.48	83.98	84.41	78.07	80.38	80.80	75.48	77.71	77.58
LC(0.05;7)	94.26	94.74	96.28	94.05	94.73	96.39	93.44	94.24	95.88
SU(1;1)	96.26	96.15	93.98	94.18	94.17	91.39	93.14	93.17	90.14
LN(0;1)	99.52	99.68	99.24	98.74	98.85	98.02	97.89	98.16	96.83