Krämer, Walter; Runde, Ralf

Working Paper
Diagnostic checking in linear processes with infinite variance

Technical Report, No. 1998,08

Provided in Cooperation with:
Collaborative Research Center 'Reduction of Complexity in Multivariate Data Structures' (SFB 475), University of Dortmund

Suggested Citation: Krämer, Walter; Runde, Ralf (1998) : Diagnostic checking in linear processes with infinite variance, Technical Report, No. 1998,08, Universität Dortmund, Sonderforschungsbereich 475 - Komplexitätsreduktion in Multivariaten Datenstrukturen, Dortmund

This Version is available at:
http://hdl.handle.net/10419/77139

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
Diagnostic checking in linear processes with infinite variance

by Walter Krämer and Ralf Runde
Fachbereich Statistik, Universität Dortmund

Abstract

We consider empirical autocorrelations of residuals from infinite variance autoregressive processes. Unlike the finite-variance case, it emerges that the limiting distribution, after suitable normalization, is not always more concentrated around zero when residuals rather than true innovations are employed.

1 Introduction and summary

In the context of standard ARMA-models

\[ y_t - \phi_1 y_{t-1} - \cdots - \phi_p y_{t-p} = e_t + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} \quad (t = 1, \ldots, n) \quad (1) \]

it is common practice to check the residuals \( \hat{e}_t \) from the fitted process for possible remaining autocorrelation. If the \( e_t \)'s are iid(0, \( \sigma^2 \)) (which in particular implies a finite variance) it is well known from Box and Pierce (1970) that the standardized empirical autocorrelations have a limiting normal distribution with mean zero, i.e.

\[ \sqrt{n} \hat{\rho}_i := \sqrt{n} \frac{\sum_{i=i+1}^{n} \hat{e}_t \hat{e}_{t-i}}{\sum_{i=1}^{\infty} \hat{e}_t^2} \xrightarrow{d} N(0, c_i) , \quad (2) \]

\[ ^1 \text{Research supported by Deutsche Forschungsgemeinschaft.} \]
where ‘\( \rightarrow^d \)’ denotes convergence in distribution, the only complication being that the variance \( c_i \) of this limiting distribution depends on the \( \phi \)'s and \( \theta \)'s from (1).

Given an increasing number of applications where ARMA-models are fitted to data the variance of which is very much in doubt, there appears to be some interest in generalizing such results to processes of the type (1) with infinite-variance innovations \( \epsilon_t \). While the correlation theory of the \( y_t \)'s themselves is in this context rather well developed (Davis and Resnick 1986; Krämer and Runde 1991; Brockwell and Davis 1991, chapter 13; Runde 1997), an extension of standard limit results to the empirical autocorrelations of the fitted residuals is still missing.

Similar to the standard case, it is easily seen that the limiting distributions of empirical autocorrelations for true innovations and fitted innovations are not identical; rather, one has to adjust the former by some scaling factor which again depends on the parameters in (1), and the present paper shows for some special cases how this adjustment can be done. It emerges that, at least for the cases we consider here, the residual-based limiting distribution can be both more concentrated around zero, as well as more spread out, so the application of the critical values appropriate for true disturbances does no longer induce a conservative test as is the case with finite variance innovations.
2 Residual autocorrelations in the standard case

To better appreciate the intricacies of infinite variance innovations, it is helpful to consider first the standard case. Let the \( \hat{\epsilon}_t \)'s be given by

\[
y_t - \hat{\phi}_1 y_{t-1} - \cdots - \hat{\phi}_p y_{t-p} = \hat{\epsilon}_t + \hat{\theta}_1 \hat{\epsilon}_{t-1} + \cdots + \hat{\theta}_q \hat{\epsilon}_{t-q} ,
\]

where the \( \hat{\phi}_i \) and \( \hat{\theta}_j \) are the ML-estimates or some other consistent estimates for the \( \phi \)'s and \( \theta \)'s. Although, by assumption, the \( \epsilon_t \)'s are iid \( (0, \sigma^2) \), the \( \hat{\epsilon}_t \)'s are not. Following Brockwell and Davis (1991, p. 481), let

\[
\tilde{\phi}(z) = \phi(z) \theta(z) = 1 - \tilde{\phi}_1 z - \cdots - \tilde{\phi}_{p+q} z^{p+q} ,
\]

where \( \phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \) and \( \theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q \), and let

\[
a(z) = \frac{1}{\phi(z)} = \lim_{j \to \infty} a_j z^j .
\]

Set \( a_j = 0 \) for \( j < 0 \), and set

\[
T_h = [a_{i-j}]_{1 \leq i \leq h, 1 \leq j \leq p+q} ,
\]

\[
\Gamma = \sum_{i=0}^{\infty} \sum_{j=0}^{p+q} a_k a_{k+j-i} \quad \text{and} \quad \Gamma^{-1} = \sum_{i=0}^{p+q} \sum_{j=0}^{p+q} a_k a_{k+j-i}
\]

\[
Q = T_h \Gamma^{-1} T_h = [q_{i,j}]_{i,j=1}^h .
\]

Then the \( h \times 1 \)-vector \( \sqrt{n} \hat{\rho} := \sqrt{n} (\hat{\rho}_1, \ldots, \hat{\rho}_h) \)' is asymptotically multivariate normal if model (1) is correct (Box and Pierce 1970):

\[
\sqrt{n} \hat{\rho} \overset{d}{\to} \mathcal{N}(0, I_h - Q) .
\]
This compares with \( \sqrt{n} \hat{\rho} \overset{d}{\rightarrow} \mathcal{N}(0, I_h) \) for the case where true innovations are used in (2). Since \( Q \) is positive semidefinite, one therefore obtains a conservative test when ignoring the fact that fitted rather than true innovations have been used.

For the important special case of a stationary AR(1) process where \( p = 1 \) and \( q = 0 \), i.e. \( y_t = \phi y_{t-1} + \epsilon_t \), where we are in addition only interested in the first-order serial correlation of the \( \epsilon_t \)'s, we have under \( H_0 \) that \( a(z) = \sum_{j=0}^{\infty} \phi^j z^j \), \( T_h = 1 \), \( \Gamma = \frac{1}{1-\phi^2} \), \( Q = 1 - \phi^2 \), and

\[
\sqrt{n} \hat{\rho}_1 \overset{d}{\rightarrow} \mathcal{N}(0, \phi^2), \quad (10)
\]

so the variance of the limiting distribution of \( \sqrt{n} \hat{\rho}_1 \) becomes rather small as \( |\phi| \to 0 \). (Note that the statement to the opposite in Brockwell and Davis (1991, p. 219, figure 9.17) is wrong.)

As we will focus mainly on this AR(1)-model below, it is instructive to derive the limiting relationship in (10) more directly, to highlight the pivotal steps in the proof where the conventional reasoning breaks down with infinite variance innovations. Estimating \( \phi \) by \( \hat{\phi} = (\frac{1}{n} \sum_{i=2}^{n} y_i y_{i-1})/(\frac{1}{n} \sum_{i=2}^{n} y_i^2) \) we have

\[
\hat{\epsilon}_t = y_t - \hat{\phi} y_{t-1} = y_t - \phi y_{t-1} - (\hat{\phi} - \phi) y_{t-1} = \epsilon_t - (\hat{\phi} - \phi) y_{t-1}, \quad (11)
\]

so

\[
\hat{\epsilon}_t \hat{\epsilon}_{t-1} = \epsilon_t \epsilon_{t-1} - (\hat{\phi} - \phi) y_{t-1} \epsilon_{t-1} - (\hat{\phi} - \phi) \epsilon_t y_{t-2} + (\hat{\phi} - \phi)^2 y_{t-1} y_{t-2} \quad (12)
\]

and

\[
\sqrt{n} \hat{\rho}_1 = \sqrt{n} \sum_{t=2}^{n} \frac{\hat{\epsilon}_t \hat{\epsilon}_{t-1}}{\sum_{t=1}^{n} \hat{\epsilon}_t^2}
\]
Since $\frac{1}{n} \sum_{t=1}^{n} \epsilon_t^2 \overset{p}{\to} \sigma^2$, we focus on the numerator from now on. Ignoring terms that are $o_p(1)$ and using

$$\hat{\phi} - \phi = \frac{\sum_{t=2}^{n} y_{t-1} \epsilon_t}{\sum_{t=2}^{n} y_{t-1}^2},$$

this numerator can be rewritten as

$$\frac{1}{\sqrt{n}} \sum_{t=2}^{n} \epsilon_t \epsilon_{t-1} - \frac{1}{\sqrt{n}} \sum_{t=2}^{n} \epsilon_t y_{t-1} \frac{1}{n} \sum_{j=2}^{n} \epsilon_j^2 - \frac{1}{n} \sum_{j=2}^{n} y_j^2,$$

where the ratio of the sums of squares tends to $1 - \phi^2$ in probability. We can thus further simplify the numerator to

$$\frac{1}{\sqrt{n}} \sum_{t=2}^{n} \epsilon_t (\epsilon_{t-1} - (1 - \phi^2) y_{t-1}) + o_p(1),$$

which is a normalized sum of a martingale difference sequence with variance $\sigma^4 \phi^2$, so by standard limit results from e.g. Hall and Heyde (1982, chap. 3.2)

$$\sqrt{n} \hat{\beta}_1 = \frac{1}{\sigma^2 \sqrt{n}} \sum_{t=2}^{n} \epsilon_t (\epsilon_{t-1} - (1 - \phi^2) y_{t-1}) + o_p(1) \overset{d}{\to} \mathcal{N}(0, \phi^2).$$

When the $\epsilon_i$'s have infinite variance this line of reasoning breaks down. It is easily seen that, apart from a different scaling factor, eq. (15) still obtains, but there the analogies end: The terms in the sum, though still a martingale difference sequence, have no finite variances due to the infinite variances of the $\epsilon_i$'s, and conventional limit theory does not apply.
3 First order residual autocorrelation with infinite variance innovations

Consider now the case where the $\epsilon$'s are iid with

\[
x^\alpha P(\epsilon_i > x) \rightarrow p e \quad \text{as} \quad x \rightarrow \infty
\]

\[
x^\alpha P(\epsilon_i \leq -x) \rightarrow q e \quad \text{as} \quad x \rightarrow \infty ,
\]

where $1 < \alpha < 2$ and $0 \leq p = 1 - q \leq 1$. This class of random variables includes the stable distributions as a special case, which in the wake of Mandelbrot (1963) and Fama (1965) are often proposed as models for returns of speculative assets (see Mittnik and Rachev 1994 for a survey). It is easily checked that $\alpha < 2$ excludes a finite variance; in addition, we confine ourselves to the empirically most important case $\alpha > 1$ where the expectation does exist.

Given iid innovations of the type (17), the unique stationary solution of the ARMA-equation (1) is given by

\[
y_t := \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i} ,
\]

where the $\psi$'s are from

\[
\frac{\theta(z)}{\phi(z)} = \sum_{j=0}^{\infty} \psi_j z^j .
\]

Under the usual stationarity conditions, we have $|\psi_j| < \infty$, so (18) exists and gives a well defined strictly stationary process which solves equation (1).

While autocorrelations of this process do not exist, it is still possible to define an analogue to the standard autocorrelation function, i.e.

\[
\gamma_i := \sum_{j=0}^{\infty} \psi_j \psi_{j+i} \quad (i = 1, 2, \ldots) ,
\]
and to estimate these pseudo-autocorrelations by their empirical counterparts

$$\hat{\gamma}_i := \frac{\sum_{j=i+1}^{n} y_j y_{j-i}}{\sum_{j=1}^{n} y_j^2}.$$  \hspace{1cm} (21)

With finite variance innovations, $\sqrt{n} (\hat{\gamma}_i - \gamma_i) \xrightarrow{d} \mathcal{N}(0,1)$, but with innovations as in (17), both the scaling factor and the limiting distribution differ (Davis and Resnick 1986; Krämer and Runde 1991). For symmetric innovations, we have

$$\frac{n}{\ln(n)} \left(\frac{1}{\alpha}\right) (\hat{\gamma}_i - \gamma_i) \xrightarrow{\mathcal{D}} \gamma_{i+j} + \gamma_{i-j} - 2\gamma_i \gamma_j \left(\frac{1}{\alpha}\right) \frac{S(\alpha)}{S(\alpha/2)},$$  \hspace{1cm} (22)

where $S(\alpha)$ and $S(\alpha/2)$ are independent stable random variables with characteristic functions

$$E(e^{i u S(\alpha)}) = e^{-\Gamma(1-\alpha) \cos(\pi \alpha/2)|u|^\alpha}$$  \hspace{1cm} (23)$$

and

$$E(e^{i u S(\alpha/2)}) = e^{-\Gamma(1-\alpha/2) \cos(\pi \alpha/4)|u|^{\alpha/2}(1-i \text{sign}(u) \tan(\pi \alpha/4))},$$  \hspace{1cm} (24)

respectively (see also Davis and Resnick 1992, p. 539). This limiting distribution is not normal, and depends on $\alpha$ (but not on other parameters of the disturbance distribution, as we have assumed symmetry, and the dispersion $c$ from (17) cancels out. We therefore set without loss of generality $c = 1$ below).

Also, from the form of the scaling factor, we see that $\hat{\gamma}_i$ converges to the true $\gamma$ faster than in the standard case.

The present paper is concerned with the $\epsilon$'s rather than the $y$'s. Using true innovations, and defining $\hat{\epsilon}_i = (\sum_{j=i+1}^{n} \epsilon_j \epsilon_{j-i})/(\sum_{j=1}^{n} \epsilon_j^2)$ similar to (21), the limiting relationship (22) simplifies to

$$\frac{n}{\ln(n)} \left(\frac{1}{\alpha}\right) \hat{\epsilon}_i \xrightarrow{\mathcal{D}} \frac{S(\alpha)}{S(\alpha/2)},$$  \hspace{1cm} (25)
where $S(\alpha)$ and $S(\alpha/2)$ have characteristic functions as in (23) and (24). However, using fitted residuals $\hat{\epsilon}_t$ in place of true innovations $\epsilon_t$, this need not and in general will not hold. We consider here the special case where the $y$’s are stationary AR(1) and where only the first order empirical autocorrelation is considered. As above, we estimate $\phi$ by $\hat{\phi} = (n \sum_{i=2}^{n} y_t y_{t-1})/(n \sum_{i=2}^{n} y_{t-1}^2)$. Replicating (11) and (12), it is easily checked that

$$
\frac{n}{\ln(n)} \frac{1/\alpha}{2/\alpha} \hat{\rho}_1 = \frac{n}{\ln(n)} \frac{1/\alpha}{2/\alpha} \frac{\sum_{i=1}^{n} \hat{\epsilon}_t \hat{\epsilon}_{t-1}}{\sum_{i=1}^{n} \hat{\epsilon}_t^2} = \frac{1}{\ln(n)} \frac{1/\alpha}{2/\alpha} \frac{\sum_{i=2}^{n} \epsilon_t \epsilon_{t-1} - (\hat{\phi} - \phi) \sum_{i=2}^{n} \epsilon_{t-1}^2}{\sum_{i=1}^{n} \epsilon_t^2} + o_p(1).
$$

(26)

Using $\hat{\phi} - \phi = (n \sum_{i=2}^{n} \epsilon_t y_{t-1})/(n \sum_{i=2}^{n} y_{t-1}^2)$ and

$$
\frac{n}{\sum_{i=2}^{n} y_t^2} \to 1 - \phi^2
$$

(27)

(this latter relationship carries over to stationary infinite variance AR(1)-processes; see Davis and Resnick 1986), we obtain

$$
\frac{n}{\ln(n)} \frac{1/\alpha}{2/\alpha} \hat{\rho}_1 = \frac{1}{\ln(n)} \frac{1/\alpha}{2/\alpha} \left[ \sum_{i=2}^{n} \epsilon_t \epsilon_{t-1} - (1 - \phi^2) \sum_{i=2}^{n} \epsilon_{t-1}^2 \right] + o_p(1)
$$

(28)

where the denominator tends in distribution to $S(\alpha/2)$. As to the numerator, we have from

$$
y_t = \epsilon_t + \phi \epsilon_{t-1} + \phi^2 \epsilon_{t-2} + \cdots
$$

that

$$
\frac{1}{n \ln(n)} \frac{1/\alpha}{n} \epsilon_t \epsilon_{t-1} - (1 - \phi^2) \epsilon_{t-1} y_{t-1}
$$

8
where \( \phi^2 \sum_{t=2}^{n} \epsilon_t \epsilon_{t-1} \) and the individual terms in the infinite sum tend in distribution to independent symmetric stable random variables (independent among each other and from \( S(\alpha/2) \)), with distributions \( \phi^2 S_0(\alpha) \), \( (1 - \phi^2) \phi S_1(\alpha) \), \( (1 - \phi^2) \phi^2 S_2(\alpha) \), \ldots. Therefore, the numerator tends in distribution to

\[
\phi^2 S_0(\alpha) + (1 - \phi^2) \sum_{j=1}^{\infty} \phi^j S_j(\alpha)
\]

\[
\overset{d}{= \phi^{2\alpha} + (1 - \phi^2)^\alpha (\phi^\alpha)^j S(\alpha)}
\]

\[
\overset{d}{= \phi^{2\alpha} + (1 - \phi^2)^\alpha \frac{\phi^\alpha}{1 - \phi^\alpha} S(\alpha)}
\]

where \( \overset{d}{=} \) denotes equality in distribution. Since the denominator tends in distribution to \( S(\alpha/2) \), we therefore have

\[
\frac{n}{\ln(n)} \overset{d}{\sim} \phi^{2\alpha} + (1 - \phi^2)^\alpha \frac{\phi^\alpha}{1 - \phi^\alpha} S(\alpha) / S(\alpha/2).
\]

The scaling factor

\[
\kappa(\alpha, \phi) := \phi^{2\alpha} + (1 - \phi^2)^\alpha \frac{\phi^\alpha}{1 - \phi^\alpha}
\]
in front of \(S(\alpha)/S(\alpha/2)\) reduces to \(\phi\) when \(\alpha = 2\) and is thus always smaller than 1 when the disturbance variance is finite. For \(\alpha < 2\), however, and \(\phi \to 1\), this factor can be larger than 1 as shown in figure 1. Therefore, the limiting distribution of the empirical first order autocorrelation of the residuals need no longer be more concentrated around zero when estimated rather than true residuals are used.

**Figure 1:** The scaling factor \(\kappa(\alpha, \phi)\)

Table 1, adapted from Krämer and Runde (1996) gives selected quantiles of the limiting distribution (31). It is seen that these critical values are monotonely increasing in both \(\phi\) and \(\alpha\), so the limiting distribution (31) is spreading out as \(\alpha\) and \(\phi\) are increasing.
### Table 1: Selected critical values of the limiting distribution (31)

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$\phi = 0.3$</th>
<th>$\phi = 0.6$</th>
<th>$\phi = 0.9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>0.643 0.664 0.754</td>
<td>1.671 1.726 1.961</td>
<td>2.924 3.020 3.431</td>
</tr>
<tr>
<td>1.3</td>
<td>0.799 0.865 1.097</td>
<td>1.906 2.063 2.617</td>
<td>2.997 3.245 4.116</td>
</tr>
<tr>
<td>1.5</td>
<td>1.057 1.158 1.361</td>
<td>2.364 2.589 3.043</td>
<td>3.544 3.882 4.563</td>
</tr>
<tr>
<td>1.7</td>
<td>1.441 1.602 1.952</td>
<td>3.062 3.403 4.146</td>
<td>4.530 5.035 6.135</td>
</tr>
<tr>
<td>1.9</td>
<td>2.044 2.295 2.880</td>
<td>4.164 4.676 5.866</td>
<td>6.201 6.962 8.735</td>
</tr>
</tbody>
</table>

### 4 Finite sample behaviour

Figure 2 compares finite sample densities of $\left( \frac{n}{\ln(n)} \right)^{1/\alpha} \hat{\rho}_1$ to the density of the limiting distribution derived in (31) when the innovations of the AR(1)-process are independent symmetric stable random variables with location parameter 0, scale parameter 1 and characteristic exponent $\alpha$. The stable variates were generated along the lines of Chambers et al. 1976, and finite sample densities were estimated by Monte Carlo, using 2000 replications, and by then applying a non-parametric kernel estimate with a biweight-kernel and a bandwidth of 0.5 (for details see e.g. Silverman 1986).

The figure shows that even for samples as large as $n = 1000$, the asymptotic distribution is far away from the finite sample distribution. This is confirmed by table 2, which gives empirical rejection rates for various $\alpha$, $\phi$, $n$ and significance levels $\theta$, again computed from 2000 runs. It shows that upper tail probabilities in finite samples are not well approximated by the asymptotic distribution: for $\alpha$ close to 1, the asymptotic distribution understates the true finite sample rejection probabilities, for $\alpha \to 2$, the asymptotic distribution
overstates the true finite sample rejection probabilities, sometimes by wide margins, and these discrepancies persist for samples as large as $n = 1000$.

**Figure 2:** Finite sample and limiting densities of first order residual autocorrelations

a) $\alpha = 1.1, \phi = 0.3$

b) $\alpha = 1.9, \phi = 0.3$
c) $\alpha = 1.1, \phi = 0.9$

d) $\alpha = 1.9, \phi = 0.9$
### Table 2: Empirical rejection probabilities for finite samples

<table>
<thead>
<tr>
<th></th>
<th>$\phi = 0.3$</th>
<th>$\phi = 0.6$</th>
<th>$\phi = 0.9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>$\hat{\varrho}$</td>
<td>$\hat{\varrho}$</td>
<td>$\hat{\varrho}$</td>
</tr>
<tr>
<td>1.1</td>
<td>0.155 0.121 0.070</td>
<td>0.131 0.110 0.087</td>
<td>0.124 0.098 0.072</td>
</tr>
<tr>
<td>1.3</td>
<td>0.132 0.097 0.042</td>
<td>0.124 0.083 0.055</td>
<td>0.119 0.079 0.048</td>
</tr>
<tr>
<td>1.5</td>
<td>0.021 0.012 0.002</td>
<td>0.028 0.019 0.003</td>
<td>0.041 0.027 0.004</td>
</tr>
<tr>
<td>1.7</td>
<td>0.001 0.000 0.000</td>
<td>0.002 0.000 0.000</td>
<td>0.003 0.002 0.001</td>
</tr>
<tr>
<td>1.9</td>
<td>0.000 0.000 0.000</td>
<td>0.000 0.000 0.000</td>
<td>0.000 0.000 0.000</td>
</tr>
</tbody>
</table>

#### 5 Conclusion

Using estimated rather than true residuals in diagnostic checking is not an innocuous procedure; the distribution of empirical residual autocorrelations is markedly affected by this substitution. While the effect of this substitution is always conservative with finite variance innovations, it cuts both ways when the disturbance variance does not exist. In addition, the asymptotic distribution of the first order empirical autocorrelation coefficient is a poor guide to in behaviour in finite samples.
References


