KYBERNETIKA — VOLUME 33 (1997), NUMBER 2, PAGES 161 – 170

ON RESIDUAL ANALYSIS FOR TIME SERIES MODELS

JIŘÍ ANDĚL

Residuals are frequently used as a diagnostic tool for verification that a time series model fits to data. In the cases when the series is nonnormal and/or the model is nonlinear, the squared residuals and squared values of the series are taken into account. In our paper asymptotic formulas for the mean value and variance of the corresponding sample correlation functions are calculated. Small sample properties are investigated in a simulation study. The results can be used for testing linearity and normality of an autoregressive time series.

1. INTRODUCTION

In this paper we assume that $\{X_t\}$ is a strictly stationary time series with vanishing mean and a covariance function $R(t)$. Define $\mathcal{X}_t = \{X_t, X_{t-1}, \ldots\}$.

Let $\{\varepsilon_t\}$ be a strict white noise with vanishing mean and a variance $\sigma^2 > 0$. Let $b_1, \ldots, b_p$ be parameters such that the polynomial $1 - b_1 z - \ldots - b_p z^p$ has all its roots outside the unit circle. Then there exists a unique linear process $\{X_t\}$ satisfying

$$X_t = b_1 X_{t-1} + \ldots + b_p X_{t-p} + \varepsilon_t. \tag{1.1}$$

The process $\{X_t\}$ is called an autoregressive process of the order p.

Let us return to the general process $\{X_t\}$ mentioned at the beginning of this section. One says that $\{X_t\}$ has linear conditional expectation if

$$E(X_t|\mathcal{X}_{t-1}) = b_1 X_{t-1} + \ldots + b_p X_{t-p} \tag{1.2}$$

where $b_1, \ldots, b_p$ are some parameters. If the covarince function $R(t)$ satisfies linear difference equation of the form

$$R(t) = b_1 R(t-1) + \ldots + b_p R(t-p), \quad t \geq 1, \tag{1.3}$$

then $\{X_t\}$ is called Yule–Walker autoregression. Let us remark that

$$(1.1) \Rightarrow (1.2) \Rightarrow (1.3). \tag{1.4}$$

The implication (1.1) $\Rightarrow$ (1.2) is clear. If (1.2) holds then

$$E(X_s X_{s-t}|\mathcal{X}_{s-1}) = X_{s-t} E(X_s|\mathcal{X}_{s-1}) = b_1 X_{s-1} X_{s-t} + \ldots + b_p X_{s-p} X_{s-t}$$

and thus

$$R(t) = EX_sX_{s-t} = E[E(X_sX_{s-t}|\mathcal{X}_{s-1})] = b_1R(t-1) + \ldots + b_pR(t-p).$$

It means that (1.2) $\Rightarrow$ (1.3). The reverse to (1.4) is not true. For example, the product autoregression model (see [11]) is the case in which (1.3) holds, but in which (1.1) and (1.2) do not. For further details, see Lawrance and Lewis [8]).

Let $\{X_t\}$ be such a time series that (1.3) holds. Then the expression $E(X_t - \beta_1X_{t-1} - \ldots - \beta_pX_{t-p})^2$ is minimized for $\beta_1 = b_1, \ldots, \beta_p = b_p$ and the variables

$$W_t = X_t - b_1X_{t-1} - \ldots - b_pX_{t-p}$$

are called linear autoregressive residuals of order p (shortly residuals). If $\{X_t\}$ is an autoregression given by (1.1) then $W_t = \varepsilon_t$.

Theorem 1.1. Let $\{X_t\}$ satisfy (1.3). Then the residuals W_t are uncorrelated (although not necessarily independent) and

$$cov(W_t, X_{t-k}) = 0 \quad \text{for} \quad k \geq 1.$$

Proof. See Lawrance and Lewis [8]. □

If the process $\{X_t\}$ satisfying (1.3) is normal then Theorem 1.1 implies that the residuals $\{W_t\}$ are independent and the variables W_t and X_{t-k} are also independent for $k \geq 1$. To detect their eventual dependence for the case that the distribution is not normal and/or the relation (1.3) does not hold, the following cross-correlation and autocorrelation functions were proposed:

$$\rho_1(k) = corr(W_t, W_{t-k}^2), \qquad \rho_2(k) = corr(W_t^2, W_{t-k}^2),$$
$$\rho_3(k) = corr(X_t, W_{t-k}^2), \qquad \rho_4(k) = corr(X_t^2, W_{t-k}).$$

For a few simple models these functions can be computed explicitly (see [8,9]). The simplest case is the AR(1) model $X_t = bX_{t-1} + \varepsilon_t$ with $|b| < 1$. If α_3 is the skewness of X_t then

$$\rho_3(k) = \alpha_3 \frac{\sigma^2}{\sqrt{\text{var}\,\varepsilon_t^2}} \frac{1-b^3}{1-b^2} b^k, \quad k \geq 0,$$

and $\rho_3(k) = 0$ for $k < 0$.

Because of rather complicated structure we do not investigate $\rho_4(k)$ in our paper.

2. ANALYSIS OF SAMPLE CORRELATION FUNCTIONS

In this section we assume that $\{X_t\}$ is a normal AR(p) process defined by (1.1). Define

$$\eta_1(k) = \frac{1}{n}\sum_{t=1}^{n} \varepsilon_t\varepsilon_{t-k}^2, \quad \eta_2 = \frac{1}{n}\sum_{t=1}^{n} \varepsilon_t^2, \quad \eta_3(k) = \frac{1}{n}\sum_{t=1}^{n} \varepsilon_{t-k}^4 - \left(\frac{1}{n}\sum_{t=1}^{n} \varepsilon_{t-k}^2\right)^2,$$

$$\eta_4(k) = \frac{1}{n}\sum_{t=1}^{n}\left(\varepsilon_t^2 - \frac{1}{n}\sum_{s=1}^{n}\varepsilon_s^2\right)\varepsilon_{t-k}^2, \quad \eta_5 = \frac{1}{n}\sum_{t=1}^{n}\varepsilon_t^4 - \left(\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t^2\right)^2,$$

$$\xi_1(k) = \frac{1}{n}\sum_{t=1}^{n}X_t\varepsilon_{t-k}^2, \quad \xi_2 = \frac{1}{n}\sum_{t=1}^{n}X_t^2.$$

Introduce the functions

$$r_1(k) = \frac{\eta_1(k)}{\sqrt{\eta_2\eta_3(k)}}, \quad r_2(k) = \frac{\eta_4(k)}{\sqrt{\eta_5\eta_3(k)}}, \quad r_3(k) = \frac{\xi_1(k)}{\sqrt{\xi_2\eta_3(k)}}.$$

The functions $r_1(k)-r_3(k)$ can be called sample correlation functions of the corresponding processes and considered as some estimates of $\rho_1(k)-\rho_3(k)$, respectively.

Note, however, that in practical applications the variables ε_t are not known and they are replaced by their estimates. In some cases such a substitution can change the asymptotic properties considerably. We discuss this problem in Section 3 below.

Consider random variables $Z_i = Z_i(X_1, \ldots, X_n)$, $i = 1, \ldots, m$. In some cases expectation and variance of a function $g(Z_1, \ldots, Z_m)$ can be calculated using the following theorem.

Theorem 2.1. Let $Z_1, \ldots, Z_m$ be random variables with finite second moments. Assume that

$$EZ_i = \theta_i + O(n^{-1}), \quad cov(Z_i, Z_j) = \frac{1}{n}c_{ij} + O(n^{-2}),$$

where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)'$ is a given vector and $\boldsymbol{C} = (c_{ij})$ is a given matrix.

Let $g(\boldsymbol{z})$ be a function of $\boldsymbol{z} = (z_1, \ldots, z_m)$ which has continuous first and second partial derivatives in a neighbourhood of the point $\boldsymbol{\theta}$. Let $\frac{\partial g}{\partial \theta_i}$ denote the value of the partial derivative $\frac{\partial g}{\partial z_i}$ at the point $\boldsymbol{z} = \boldsymbol{\theta}$. Then we have

$$\begin{aligned} Eg(Z_1, \ldots, Z_m) &= g(\theta_1, \ldots, \theta_m) + O(n^{-1}), \\ \operatorname{var} g(Z_1, \ldots, Z_m) &= \frac{1}{n}\sum_{i=1}^{m}\sum_{j=1}^{m}c_{ij}\frac{\partial g}{\partial \theta_i}\frac{\partial g}{\partial \theta_j} + O(n^{-2}). \end{aligned}$$

Proof. The assertion can be proved similarly as that in section 10.6 in Kendall and Stuart [7]. □

Theorem 2.2. Let $\{\varepsilon_t\}$ be i.i.d. $N(0, \sigma^2)$ variables with $\sigma^2 > 0$. Let $k \neq 0$ be fixed. Then

$$Er_1(k) = O(n^{-1}), \quad \operatorname{var} r_1(k) = \frac{3}{2n} + O(n^{-2}).$$

Proof. We have

$$E\eta_1(k) = 0, \quad E\eta_2 = \sigma^2, \quad E\eta_3(k) = 2\sigma^4 + O(n^{-1}), \quad \operatorname{var}\eta_1(k) = \frac{3\sigma^6}{n}.$$

Applying Theorem 2.1 we get our assertion. □

Theorem 2.3. Let $\{\varepsilon_t\}$ be i.i.d. $N(0, \sigma^2)$ variables with $\sigma^2 > 0$. Let $k \neq 0$ be fixed. Then

$$Er_2(k) = O(n^{-1}), \quad \operatorname{var} r_2(k) = \frac{1}{n} + O(n^{-2}).$$

Proof. We derive

$$E\eta_5 = E\eta_3(k) = 2\sigma^4 + O(n^{-1}), \quad E\eta_4(k) = O(n^{-1}), \quad \operatorname{var}\eta_4(k) = \frac{4\sigma^8}{n} + O(n^{-2})$$

and then the result follows from Theorem 2.1. □

Results for $r_3(k)$ given in the following theorem were derived only in a special case.

Theorem 2.4. Let $\{\varepsilon_t\}$ be i.i.d. $N(0, \sigma^2)$ variables with $\sigma^2 > 0$ and let $\{X_t\}$ be an AR(1) process defined by $X_t = bX_{t-1} + \varepsilon_t$, $|b| < 1$. Then for any fixed $k \geq 1$ we have

$$\begin{aligned} Er_3(k) &= O(n^{-1}), \\ \operatorname{var} r_3(k) &= \frac{1}{n}\left[6b^{2k}(1-b^2) + b^k(2-b^k)(1+b) + \frac{3-b}{2(1-b)}\right] + O(n^{-2}). \end{aligned}$$

Proof. It is easy to see that

$$E\xi_1(k) = 0, \quad E\xi_2 = \frac{\sigma^2}{1-b^2}, \quad E\eta_3(k) = 2\sigma^4 + O(n^{-1}),$$
$$E\xi_1(k)\,\xi_2 = 0, \quad E\xi_1(k)\,\eta_3(k) = 0,$$
$$cov(\xi_1(k), \xi_2) = 0, \quad cov(\xi_1(k), \eta_3(k)) = O(n^{-2})$$

and after some computation we get

$$\operatorname{var}\xi_1(k) = \frac{2\sigma^6}{n}\left[6b^{2k} + \frac{b^k(2-b^k)(1+b)}{1-b^2} + \frac{3-b}{2(1-b)(1-b^2)}\right] + O(n^{-2}).$$

Then we use again Theorem 2.1. □

3. REMARKS

Let w_t be estimated residuals which are calculated in autoregressive models AR(p) in a similar way as W_t with the exception that the parameters b_i are substituted by their least-squares estimates. Define

$$r^*(k) = \frac{\sum_{t=k+1}^{n} w_t w_{t-k}}{\sum_{t=1}^{n} w_t^2}.$$

Box and Pierce [2] obtained the asymptotic distribution of $r^*(k)$. They pointed out that it is different from that of

$$r(k) = \frac{\sum_{t=k+1}^{n} \varepsilon_t \varepsilon_{t-k}}{\sum_{t=1}^{n} \varepsilon_t^2}.$$

They proposed a test of fit based on the portmanteau statistic

$$Q_{BP} = n \sum_{k=1}^{M} r^{*2}(k).$$

Under the assumption of the model adequacy, Q_{BP} is distributed approximately χ^2_{M-p}. Davies et al [4] and Ljung and Box [10] showed that the modified statistic

$$Q_{LB} = n(n+2) \sum_{k=1}^{M} \frac{r^{*2}(k)}{n-k}$$

is distributed closer to χ^2_{M-p} in small samples.

The situation concerning squared residuals is quite different. Define

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{t=1}^{n} w_t^2, \quad r_2^*(k) = \frac{\sum_{t=k+1}^{n} (w_t^2 - \hat{\sigma}^2)(w_{t-k}^2 - \hat{\sigma}^2)}{\sum_{t=1}^{n} (w_t^2 - \hat{\sigma}^2)^2}.$$

McLeod and Li [13] proved that the vector $(r_2^*(1), \ldots, r_2^*(M))'$ is asymptotically normal $N(\mathbf{0}, \boldsymbol{I})$. They introduced the portmanteau statistic

$$Q_{LL} = n(n+2) \sum_{k=1}^{M} \frac{r_2^{*2}(k)}{n-k}$$

which is asymptotically distributed as χ^2_M when the model is correct.

Some authors prefer to use tests based on individual values of $r^*(k)$ instead on the portmanteau statistic. The vector $(r^*(1), \ldots, r^*(M))'$ has asymptotically $N(\mathbf{0}, \frac{1}{n}\boldsymbol{V})$ distribution and a formula for $\boldsymbol{V} = (v_{ij})$ can be found in several papers (e. g. Box and Pierce [2], McLeod [12], Hosking and Ravishanker [6]. Hosking and Ravishanker [6] recommend for general use the classical Bonferroni approximation to the critical region for simultaneous testing of $r^*(1), \ldots, r^*(M)$. The model is not rejected if

$$|r^*(k)| \leq u\left(\frac{\alpha}{2M}\right) \sqrt{\frac{n}{v_{kk}^*}} \quad \text{for} \quad k = 1, \ldots, M$$

where α is the level of significance, $u(\beta)$ is the critical value of the $N(0,1)$ distribution on level β and v_{kk}^* is an estimate of v_{kk}.

4. A SIMULATION STUDY

A simulation study was performed for small and moderate sample sizes in order to get answers to the following questions:

(i) Are the approximations for $Er_i(k)$ and $\mathrm{var}\, r_i(k)$ based on Theorems 2.2–2.4 precise enough?

(ii) Are the distributions of $r_i(k)$ approximately normal?

(iii) How much are the moments and the distributions of $r_i(k)$ influenced when their definitions are modified in such a way that the variables ε_t are substituted by their estimates w_t?

Now, we formulate our problems more rigorously. From Theorem 2.2 we have approximations $Er_1(k) \doteq 0$, $\mathrm{var}\, r_1(k) \doteq 3/(2n)$; similar approximations for $Er_2(k)$, $\mathrm{var}\, r_2(k)$ and $Er_3(k)$, $\mathrm{var}\, r_3(k)$ are based on Theorems 2.3 and 2.4, respectively.

For different n (and for different parameters of our models) $r_1(k)$, $r_2(k)$ and $r_3(k)$ were simulated $N = 1\,000$ times in each case. Thus we have $N = 1\,000$ realizations of $r_i(k)$ in each case. We calculated empirical skewness a_3 and empirical kurtosis a_4 from these values. It is well known (see Cramér, § 29.3) that in normal samples

$$Ea_3 = 0, \quad \mathrm{var}\, a_3 = \frac{6(N-2)}{(N+1)(N+3)},$$

$$Ea_4 = -\frac{6}{N+1}, \quad \mathrm{var}\, a_4 = \frac{21N(N-2)(N-3)}{(N+1)^2(N+3)(N+5)}$$

and that a_3, a_4 are asymptotically normal and asymptotically independent. Define the normed skewness u_3 and the normed kurtosis u_4 by

$$u_3 = \frac{a_3}{\sqrt{\mathrm{var}\, a_3}}, \quad u_4 = \frac{a_4 - Ea_4}{\sqrt{\mathrm{var}\, a_4}}.$$

A test statistic is $\chi^2 = u_3^2 + u_4^2$ and it has asymptotically χ^2_2-distribution. It is known that the sample size $N = 1\,000$ is large enough to justify the use of the asymptotic results.

Some results concerning $r_1(k)$ are summarized in Table 4.1.

Table 4.1. Results of 1 000 simulations of $r_1(k)$ for $n = 100$.

lag	average	empirical variance	a_3	a_4	u_3	u_4	χ^2
1	0.00	0.014	0.10	−0.07	0.19	−0.60	0.39
2	0.00	0.015	−0.10	−0.05	−0.18	−0.39	0.18
3	0.00	0.015	0.00	−0.08	0.01	−0.72	0.52
4	0.00	0.013	−0.01	−0.12	−0.02	−1.11	1.23
5	0.00	0.014	0.11	0.11	0.20	1.10	1.26

As mentioned above, for $n = 100$ our approximations are $Er_1(k) \doteq 0$, $\mathrm{var}\, r_1(k) \doteq 3/(2n) = 0.015$. They agree well with results of the simulations (see columns average and empirical variance). The values of a_3, a_4, u_3, u_4, and χ^2 correspond to values

expected under normality of $r_1(k)$. Similar results were obtained also for $n = 50$ and for $n = 200$.

When the same procedure was applied to residuals w_t from different AR(1) models it was found that the averages and empirical variances of $r_1(k)$ are practically the same as in simulations using the true white noise ε_t. Although a_3 and a_4 were also near to 0, a few significant values of χ^2 appeared on the five-per cent level.

It seems that $r_1(k)$ has its expectation and variance near to our approximations and its distribution is nearly normal for $n \geq 50$. Another simulation study showed that the same conclusions also hold for the case when ε_t are substituted by their estimates w_t.

Some results of a simulation concerning $r_2(k)$ can be found in Table 4.2.

Table 4.2. Results of 1 000 simulations of $r_2(k)$ for $n = 100$.

lag	average	empirical variance	a_3	a_4	u_3	u_4	χ^2
1	−0.02	0.0088	0.55	−0.20	1.03	−1.85	4.48
2	−0.02	0.0087	0.65	0.09	1.21	0.94	2.36
3	−0.02	0.0090	0.42	−0.03	0.79	−0.24	0.67
4	−0.03	0.0081	0.55	−0.36	1.04	−3.32	12.09
5	−0.02	0.0088	0.52	0.37	0.96	3.50	13.16

Our approximations based on Theorem 2.3 are $Er_2(k) \doteq 0$, $var\, r_2(k) \doteq 1/n = 0.01$. General conclusions about $r_2(k)$ are: The approximation for $var\, r_2(k)$ is not as good as that for $var\, r_1(k)$ and the distribution of $r_2(k)$ can be closely approximated by normal distribution only for $k = 1, 2, 3$. The same results were obtained also for $n = 200$. The situation is very similar also in the case when empirical residuals w_t are used instead of ε_t.

Finally, Table 4.3 contains some results about $r_3(k)$.

Table 4.3. Results of 1 000 simulations of $r_3(k)$ for $n = 100$ and $b = 0.5$.

lag	average	empirical variance	theoretical variance	a_3	a_4	u_3	u_4	χ^2
1	0.00	0.043	0.047	0.02	−0.33	0.05	−2.98	8.89
2	0.01	0.032	0.034	−0.05	−0.17	−0.10	−1.50	2.26
3	0.00	0.027	0.029	−0.05	−0.24	−0.09	−2.22	4.93
4	0.00	0.026	0.027	−0.02	−0.30	−0.04	−2.77	7.68
5	0.00	0.025	0.026	0.03	−0.30	0.05	−2.81	7.89

Our approximation to $Er_3(k)$ is 0. The approximation to $var\, r_3(k)$ based on Theorem 2.4 is introduced in Table 4.3 as "theoretical variance". The approximations agree well with the corresponding averages and empirical variances also for other

$b \in (-1, 1)$ when $n \geq 50$. But the distribution of $r_3(k)$ can be approximated by normal distribution only if $|b|$ is small, say $|b| \leq 0.1$ (even for $n = 200$). When b is estimated and $r_3(k)$ is based on the empirical residuals w_t then the results are the same.

5. USE OF RESIDUALS

Many methods used in time series analysis for detecting departures from a model are based on residuals. In this section we assume that $X_1, \ldots, X_n$ is a realization of the stationary AR(1) model

$$X_t = bX_{t-1} + e_t \tag{5.1}$$

where e_t are i.i.d. $N(0, \sigma^2)$ variables. This assumption can be violated in several ways: the white noise may not be normal, the autoregression may not be of the first order, the model may not be linear etc. In some examples given below we use models for which $EX_t \neq 0$. Thus the average $\bar{X}$ is subtracted from all X_i in all cases before other computations. Then the least squares estimate $\hat{b}$ of b is calculated. The empirical residuals

$$w_t = (X_t - \bar{X}) - \hat{b}(X_{t-1} - \bar{X}), \quad t = 2, \ldots, n$$

are used for decision if our model (5.1) is the true one or not. Define

$$r_0(k) = \frac{\sum w_t w_{t-k}}{\sqrt{\sum w_t^2 \sum w_{t-k}^2}}, \quad k = 1, 2, \ldots$$

the classical residual autocorrelation function, which is frequently used as a standard tool for model identification and checking (see Box and Jenkins [1]). For simplicity we follow the recommendation of Box and Jenkins and we approximate the distribution of $r_0(k)$ by $N(0, \frac{1}{n_0})$ where n_0 is the number of terms in the summation $\sum w_t w_{t-k}$. A more precise result can be found in Box and Pierce [2] as it was mentioned in Section 3.

We use the following abbreviations. Let $Ex(\lambda)$ denote the exponential distribution with the parameter λ. If $Y \sim Ex(\lambda)$ and a is a number then the distribution of $Y - a$ is denoted by $Ex(\lambda) - a$. Further, $R(a, b)$ is the rectangular distribution on $[a, b]$. We consider the following models:

(A)	$X_t = 0.5X_{t-1} + e_t,$	$e_t \sim Ex(1) - 1;$
(B)	$X_t = 0.5\sqrt{X_{t-1}} + e_t,$	$e_t \sim R(0, 1);$
(C)	$X_t = 0.5\lvert X_{t-1}\rvert + e_t,$	$e_t \sim N(0, 1);$
(D)	$X_t = 1.4X_{t-1} - 0.45X_{t-2} + e_t,$	$e_t \sim N(0, 1);$
(E)	$X_t = e_t + 0.9e_{t-1}e_{t-2},$	$e_t \sim N(0, 1).$

As for the model (E), it is known that the variables X_t given by $X_t = e_t + \beta e_{t-1}e_{t-2}$ are uncorrelated but dependent if e_t are i.i.d. $N(0, \sigma^2)$ (see Harvey [5], p. 267).

Realizations of the length $n = 1\,000$ of the models (A) – (E) were simulated. The values of $r_0(k) - r_3(k)$ were calculated for $|k| \leq 5$. For a rough analysis we used our approximations for $Er_i(k)$ and for $var\, r_i(k)$ as well as the normal approximation for the distribution of $r_i(k)$. The cases when $|r_i(k)|$ exceeded the Bonferroni bound $u\left(\frac{\alpha}{2M}\right)\sqrt{var\, r_i(k)}$ for $\alpha = 0.05$ and $M = 5$ are introduced in Table 5.1.

Table 5.1.

Model	$\bar{X}$	$\hat{b}$	significant values of $r_i(k)$
(A)	−0.02	0.53	$r_3(1)$, $r_3(2)$
(B)	−1.00	0.25	—
(C)	−0.50	0.16	$r_1(1)$, $r_2(1)$, $r_3(1)$
(D)	−0.17	0.96	$r_0(1)$, $r_0(2)$, $r_2(1)$
(E)	−0.03	0.00	$r_2(1)$, $r_2(2)$

The nonnormality of the model (A) was detected only by $r_3(k)$. The serious nonlinearity of the model (B) was not detected at all. The nonlinearity of the model (C) was detected by $r_1(k)$, $r_2(k)$ and $r_3(k)$ but only for $k = 1$. The misspecified order in the case (D) was detected by $r_0(k)$ (as it should be) and then only by $r_2(1)$. The case of uncorrelated but dependent variables in the model (E) was detected only by $r_2(k)$.

Our experience from other simulations and from other models is similar. No single function $r_i(k)$ is universal tool for detecting that the assumptions are not fulfilled. The model (B) is extraordinarily alarming. Although $\{X_t\}$ given by (B) is neither linear AR(1) nor normal, none of the functions $r_i(k)$ has been able to detect it.

ACKNOWLEDGEMENT

This research was supported by Grant 188/96 from the Grant Agency of Charles University.

(Received May 14, 1996.)

REFERENCES

[1] G. E. P. Box and G. M. Jenkins: Models for Prediction and Control. Holden Day, New York 1970.

[2] G. E. P. Box and D. A. Pierce: Distribution of residual autocorrelations in autoregressive–integrated moving average time series models. J. Amer. Statist. Assoc. *65* (1970), 1509–1526.

[3] H. Cramér: Mathematical Methods of Statistics. Princeton Univ. Press, Princeton 1946.

[4] N. Davies, C. M. Triggs and P. Newbold: Significance of the Box–Pierce portmanteau statistics in finite samples. Biometrika *64* (1977), 517–522.

[5] A. C. Harvey: Time Series Models. Second edition. Harvester Wheatsheaf, New York 1993.

[6] J. R. M. Hosking and N. Ravishanker: Approximate simultaneous significance intervals for residual autocorrelations of autoregressive moving–average time series models. J. Time Ser. Anal. *14* (1993), 19–26.

[7] M. G. Kendall and A. Stuart: The Advanced Theory of Statistics. Vol. I: Distribution Theory. Griffin, London 1969.
[8] A. J. Lawrance and P. A. W. Lewis: Higher-order residual analysis for nonlinear time series with autoregressive correlation structure. Internat. Statist. Rev. *55* (1987), 21–35.
[9] A. J. Lawrance and P. A. W. Lewis: Reversed residuals in autoregressive time series analysis. J. Time Ser. Anal. *13* (1992), 253–266.
[10] G. M. Ljung and G. E. P. Box: On a measure of lack of fit in time series models. Biometrika *65* (1978), 297–303.
[11] E. McKenzie: Product autoregression: a time series characterization of the gamma distribution. J. Appl. Probab. *19* (1982), 463–468.
[12] A. I. McLeod: On the distribution of residual autocorrelations in Box–Jenkins models. J. Roy. Statist. Soc. Ser. B *40* (1978), 296–302.
[13] A. I. McLeod and W. K. Li: Diagnostic checking ARMA time series models using squared-residual autocorrelations. J. Time Ser. Anal. *4* (1983), 269–273.

Prof. RNDr. Jiří Anděl, DrSc., Univerzita Karlova, Matematicko-fyzikální fakulta (Charles University — Faculty of Mathematics and Physics), Sokolovská 83, 186 00 Praha 8. Czech Republic.