AD-A049 397 WISCONSIN UNIV MADISON MATHEMATICS RESEARCH CENTER F/G 12/1
ON A MEASURE OF LACK OF FIT IN TIME SERIES MODELS.(U)
SEP 77 G M LJUNG, G E BOX DAAG29-75-C-0024
UNCLASSIFIED MRC-TSR-1790 NL

1 OF 1
AD
A049397

END
DATE
FILMED
3-78
DDC

MRC Technical Summary Report #1790

ON A MEASURE OF LACK OF FIT IN TIME SERIES MODELS

G. M. Ljung and G. E. P. Box

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

September 1977

(Received August 12, 1977)

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

Air Force Office of Scientific Research
Building 410
Bolling AFB
Washington, DC 20332

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

ON A MEASURE OF LACK OF FIT IN TIME SERIES MODELS

G. M. Ljung and G. E. P. Box

Technical Summary Report #1790
September 1977

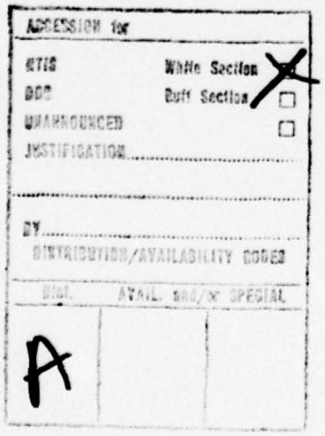
ABSTRACT

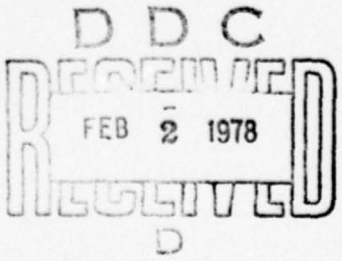
The overall test for lack of fit in autoregressive-moving average models proposed by Box and Pierce (1970) is considered. It is shown that a substantially improved approximation results from a simple modification of this test. Some consideration is given to the power of such tests and their robustness when the innovations are non-normal. Similar modifications in the overall tests used for transfer function-noise models are proposed.

AMS (MOS) Subject Classifications: Primary 62M10; Secondary 62E15, 62E25

Key Words: Autoregressive-moving average models; Residual autocorrelations; Tests for lack of fit; Transfer function-noise models

Work Unit Number 4 (Probability, Statistics and Combinatorics)





Sponsored by the United States Army under Contract No. DAAG29-75-C-0024 and by the Air Force Office of Scientific Research under Grant AFOSR 72-2363D.

SIGNIFICANCE AND EXPLANATION

Very many physical situations are described by a series of numbers occurring sequentially in time, the observed numbers arising from a deterministic underlying smoothly varying sequence perturbed by random errors. Typical examples are observations of the position of a vehicle or missile at successive intervals in time, and many other examples occur in connection with engineering production and business management.

Such time-series are normally analyzed, after allowing for the deterministic part, by assuming some underlying model for the process involved, the relevant model in the present work being an autoregressive moving average model described in the first paragraph of the paper. This is the model on which the Box-Jenkins technique is based.

If the parameters in a model are estimated from an experimental time-series, the next question is the adequacy of fit of the data by the model. In this connection it is useful to study the residuals and their autocorrelations $\hat{r}_k$. Box and Pierce (1970) propose an overall test for lack of fit based on approximating the distribution of $Q(\hat{r}) = n \sum_{k=1}^{m} \hat{r}_k^2$ by the χ^2_{m-p-q} distribution where n is the number of observations in the time series, m is the number of lags, and $p + q$ is the number of parameters in the model. Recent studies have shown that this approximation is not adequate unless n is large relative to m. The present paper illustrates that the use of the modified test statistic $\tilde{Q}(\hat{r}) = n(n+2) \sum_{k=1}^{m} (n-k)^{-1} r_k^2$ leads to a substantially improved approximation. The power of the overall test and its robustness to non-normality of the innovations (i.e. random perturbations on the data) in the model are discussed briefly. Some consideration is also given to testing for lack of fit in transfer function-noise models, where one has two time series, variations in one series being related to variations in the other.

ON A MEASURE OF LACK OF FIT IN TIME SERIES MODELS

G. M. Ljung and G. E. P. Box

1. INTRODUCTION

Consider a discrete time series $\{w_t\}$ generated by a stationary autoregressive moving average model

$$\phi(B)w_t = \theta(B)a_t \tag{1.1}$$

where $\phi(B) = 1-\phi_1 B-\ldots-\phi_p B^p$, $\theta(B) = 1-\theta_1 B-\ldots-\theta_q B^q$, $B^k w_t = w_{t-k}$, $\{a_t\}$ is a sequence of independent and identically distributed $N(0,\sigma^2)$ random deviates, and where the w_t's can represent the d-th difference or some other suitable transformation of a nonstationary series $\{z_t\}$.

After a model of this form has been fitted to a series $w_1,\ldots,w_n$, it is useful to study the adequacy of the fit by examining in various ways the residuals $\hat{a}_1,\ldots,\hat{a}_n$ and, in particular, their autocorrelations

$$\hat{r}_k = \frac{\sum_{t=k+1}^{n} \hat{a}_t\hat{a}_{t-k}}{\sum_{t=1}^{n} \hat{a}_t^2}, \quad k = 1,2,\ldots .$$

An informal graphical analysis of these quantities combined with overfitting [see, for example, Box and Jenkins (1970)] usually proves most effective in detecting possible deficiencies in the model. In addition, however, it is often worthwhile to look at an overall criterion of adequacy of fit. Box and Pierce (1970) noted that if the model were appropriate and the parameters were known, the quantity

---

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024 and by the Air Force Office of Scientific Research under Grant AFOSR 72-2363D.

$$\tilde{Q}(r) = n(n+2) \sum_{k=1}^{m} (n-k)^{-1} r_k^2 , \qquad (1.2)$$

where

$$r_k = \frac{\sum_{t=k+1}^{n} a_t a_{t-k}}{\sum_{t=1}^{n} a_t^2} ,$$

would for large n be distributed as χ^2_m since the limiting distribution of $\underset{\sim}{r} = (r_1,\ldots,r_m)$ is multivariate normal with mean vector zero [Anderson (1942), Anderson and Walker (1964)], $\mathrm{Var}(r_k) = (n-k)/n(n+2)$ and $\mathrm{Cov}(r_k,r_\ell) = 0$, $k \neq \ell$. Using the further approximation $\mathrm{Var}(r_k) = 1/n$, Box and Pierce suggested that the distribution of

$$Q(r) = n \sum_{k=1}^{m} r_k^2 \qquad (1.3)$$

could be approximated by that of χ^2_m. Furthermore, they showed that when the $p + q$ parameters of an appropriate model are estimated and the $\hat{r}_k$'s replace the r_k's, then

$$Q(\hat{r}) = n \sum_{k=1}^{m} \hat{r}_k^2$$

would for large n be distributed as χ^2_{m-p-q} yielding an approximate test for lack of fit.

In applications of this test, suspiciously low values of $Q(\hat{r})$ have sometimes been observed and studies by Ljung and Box (1976) and Davies et al. (1977) have verified that the distribution of $Q(\hat{r})$ can deviate from χ^2_{m-p-q}. This observation was also made by Prothero and Wallis (1976, Discussion). The observed discrepancies could be accounted for by several factors among these departures from normality of the autocorrelations. It appears, however, that the main difficulty is caused by the approximation of (1.2) by (1.3).

A modified test based on the criterion

$$\tilde{Q}(\hat{r}) = n(n+2) \sum_{k=1}^{m} (n-k)^{-1} \hat{r}_k^2$$

was recommended by Ljung and Box (1976) but its usefulness was questioned by Davies et al. (1977) on the ground that the variance of $\tilde{Q}(\hat{r})$ exceeds that of the χ^2_{m-p-q} distribution. Our studies show however that the modified test provides a substantially improved approximation that should be adequate for most practical purposes.

2. MEANS AND VARIANCES OF $Q(r)$ AND $\dot{Q}(r)$

To examine the overall test, it is useful to initially consider the quantities $Q(r)$ and $\tilde{Q}(r)$ which involve the white noise autocorrelations $\underset{\sim}{r}$. Since the limiting distribution of $\underset{\sim}{r}$ is $N(\underset{\sim}{0}, n^{-1}I_m)$, $Q(r)$ and $\dot{Q}(r)$ are asymptotically distributed as χ^2_m and have expectation m and variance $2m$. For finite values of n, $\dot{Q}(r)$ has expectation m, whereas

$$EQ(r) = n \sum_{k=1}^{m} E(r_k^2) = \frac{mn}{n+2}\left\{1 - \frac{m+1}{2n}\right\} . \tag{2.1}$$

Clearly, unless n is large relative to m, $EQ(r)$ can be much smaller than m.

The variances are

$$\text{Var } Q(r) = n^2 \sum_{k=1}^{m} \text{Var}(r_k^2) + 2n^2 \sum_{k=1}^{m-1} \sum_{\ell=k+1}^{m} \text{Cov}(r_k^2, r_\ell^2) \tag{2.2}$$

and

$$\text{Var } \dot{Q}(r) = n^2(n+2)^2 \sum_{k=1}^{m} (n-k)^{-2} \text{Var}(r_k^2)$$

$$+ 2n^2(n+2)^2 \sum_{k=1}^{m-1} \sum_{\ell=k+1}^{m} (n-k)^{-1}(n-\ell)^{-1} \text{Cov}(r_k^2, r_\ell^2) ,$$

where for fixed n, $\mathrm{Cov}(r_k^2, r_\ell^2)$ is non-zero. The univariate and bivariate moments of the r_k's needed to evaluate (2.2) can be obtained using the identity

$$E(r_k^i r_\ell^j) = \frac{E\{(\Sigma\, a_t a_{t-k})^i (\Sigma\, a_t a_{t-\ell})^j}{E(\Sigma\, a_t^2)^{i+j}} \tag{2.3}$$

which follows from independence of the r_k's and $\Sigma\, a_t^2$ [see, for example, Anderson (1971), p. 304]. Taking $\mathrm{Var}(a_t) = 1$ without loss of generality, $\Sigma\, a_t^2$ is distributed as χ_n^2 and $E(\Sigma\, a_t^2)^{i+j} = n(n+2) \ldots (n+2i+2j-2)$. The term in the numerator of (2.3) can be evaluated by multiplying term by term taking the expected value. Using this procedure, it can be verified that for $k \leq n/2$

$$\mathrm{Var}(r_k^2) = \frac{6(3n-5k) + 3(n-k)^2}{n(n+2)(n+4)(n+6)} - \frac{(n-k)^2}{n^2(n+2)^2}$$

and (2.4)

$$\mathrm{Cov}(r_k^2, r_\ell^2) = \frac{(n-k)(n-\ell) + 4(n-\ell) + 8(n-k-\ell)}{n(n+2)(n+4)(n+6)} - \frac{(n-k)(n-\ell)}{n^2(n+2)^2} \quad .$$

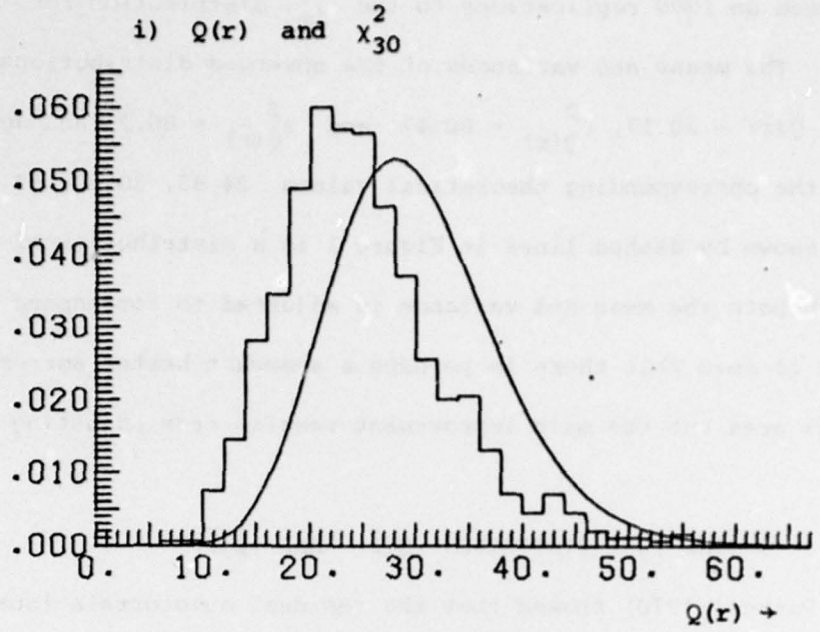
The exact variances of $Q(r)$ and $\tilde{Q}(r)$ are readily evaluated using (2.2) and (2.4). By ignoring terms of order higher than $1/n$ it may be shown that approximately, for n large relative to m,

$$\mathrm{Var}Q(r) = 2m\{1 + \frac{m - 10}{n}\}$$

and

$$\mathrm{Var}\tilde{Q}(r) = 2m\{1 + \frac{2m - 5}{n}\} \quad .$$

The variance of $\tilde{Q}(r)$ exceeds $2m$ but the absence of a location bias makes its distribution much closer to χ_m^2 than that of $Q(r)$. This is illustrated in Figure 1 which compares Monte Carlo distributions of $Q(r)$



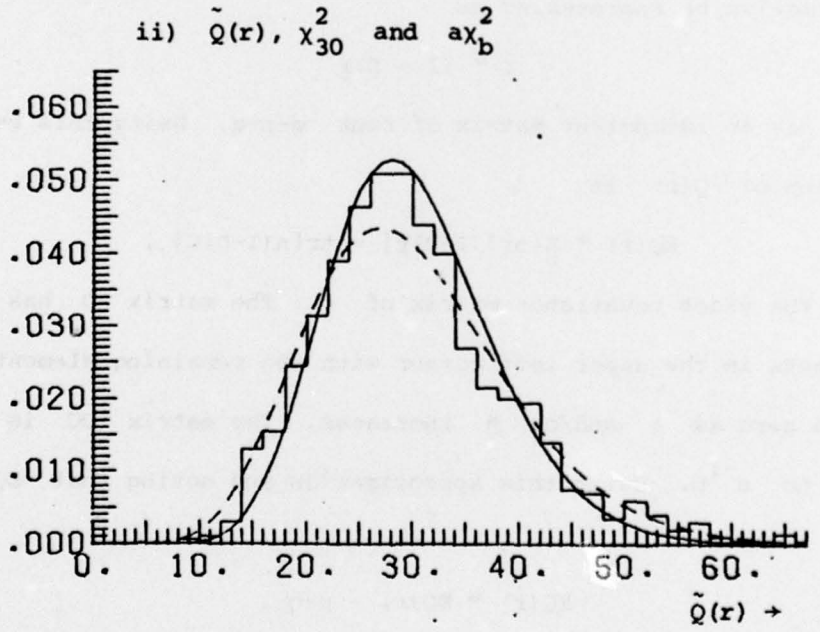


Fig. 1. Monte Carlo distributions of $Q(r)$ and $\tilde{Q}(r)$ and approximations; 1000 replications, $n = 100$ and $m = 30$; _____ χ^2_{30}, ---- $a\chi^2_b$ ($a = 1.52$, $b = 19.68$).

and $\tilde{Q}(r)$ based on 1000 replications to the χ^2_m distribution for $m = 30$ and $n = 100$. The means and variances of the observed distributions are $\bar{Q}(\hat{r}) = 24.97$, $\bar{\tilde{Q}}(\hat{r}) = 30.17$, $s^2_{Q(\hat{r})} = 60.47$ and $s^2_{\tilde{Q}(\hat{r})} = 88.25$ and agree quite closely with the corresponding theoretical values 24.85, 30.00, 63.15 and 91.48. Also shown by dashed lines in Figure 1 is a distribution of the form $a\chi^2_b$ for which both the mean and variance is adjusted to correspond with that of $\tilde{Q}(r)$. It is seen that there is perhaps a somewhat better agreement in the upper tail area but the main improvement results from adjusting the mean.

3. THE TEST STATISTICS $Q(\hat{r})$ AND $\tilde{Q}(\hat{r})$

Box and Pierce (1970) showed that the residual autocorrelations $\hat{r} = (\hat{r}_1,\ldots,\hat{r}_m)'$ from a correctly identified and fitted model can to a close approximation be represented as

$$\hat{r} \simeq (I - D)r \tag{3.1}$$

where $I - D$ is an idempotent matrix of rank $m-p-q$. Using this relationship, the expectation of $Q(\hat{r})$ is

$$EQ(\hat{r}) \simeq E\{nr'(I-D)r\} = tr\{n(I-D)C\} ,$$

where C is the exact covariance matrix of r. The matrix D has its largest elements in the upper left corner with the remaining elements d_{ij} decreasing to zero as i and/or j increases. The matrix DC is therefore nearly equal to $n^{-1}D$. Using this approximation and noting that $EQ(r) = tr(nC)$, we have

$$EQ(\hat{r}) \simeq EQ(r) - p-q . \tag{3.2}$$

Combining (2.1) and (3.2), the expected value of $Q(r)$ is approximately

$$EQ(\hat{r}) \simeq \frac{mn}{n+2}\left\{1 - \frac{m+1}{2n}\right\} - p-q , \tag{3.3}$$

which indicates that the distribution of $Q(\hat{r})$ can deviate markedly from χ^2_{m-p-q} unless n is large relative to m. However, using the same approximations it can be shown that

$$E\tilde{Q}(\hat{r}) \simeq E\tilde{Q}(r) - p-q$$

$$= m-p-q \ .$$

It may be expected therefore that the distribution of $\tilde{Q}(\hat{r})$ might be approximated by the χ^2_{m-p-q} distribution.

The adequacy of this approximation was questioned by Davies et al. (1977) on the ground that the variance of $\tilde{Q}(\hat{r})$ exceeds $2(m-p-q)$. However, results from a simulation study reported in the next section suggest that the reduction in the location bias results as before in a markedly improved approximation that should be adequate for most practical purposes. It also appears that the expression for the variance given by Davies et al., which is not exact, overestimates the variance of $\tilde{Q}(\hat{r})$. For example, for fitting a first order autoregressive model to white noise, Davies et al. obtain for $m = 20$, $n = 50$, 100 and 200, $\mathrm{Var}\tilde{Q}(\hat{r}) = 58.80$, 50.08 and 44.20, respectively, while our study gives $\mathrm{Var}\tilde{Q}(\hat{r}) = 46.84$, 43.20 and 41.97, respectively.

4. SOME NUMERICAL RESULTS

Comparison of the Overall Tests

A Monte Carlo study was conducted by generating 4000 sets of observations $\{w_1,\ldots,w_n\}$ from the first order autoregressive model $w_t - \phi w_{t-1} = a_t$, estimating ϕ by the approximate maximum likelihood estimator

$$\hat{\phi} = \frac{n-2}{n-1} \times \frac{\sum_{t=2}^{n} w_t w_{t-1}}{\sum_{t=2}^{n-1} w_t^2}$$

[Box and Jenkins (1970), p. 279], and calculating autocorrelations of the residuals $\hat{a}_1 = (1-\hat{\phi}^2)w_1$, $\hat{a}_t = w_t - \hat{\phi}w_{t-1}$, $t = 2,\ldots,n$. The statistics $Q(\hat{r})$ and $\tilde{Q}(\hat{r})$ were then calculated.

Table 1 shows the proportion of $Q(\hat{r})$ and $\tilde{Q}(\hat{r})$ values exceeding the upper i) 5, ii) 10 and iii) 25 percentage points of the χ^2_{m-1} distribution for a few combinations of n and m and for $\phi = .5$. The table also gives the means and variances of the observed distributions. It seems clear that although the variance of $\tilde{Q}(\hat{r})$ exceeds $2(m-1)$ a test based on this statistic would for smaller sample sizes provide a considerable improvement over the previously used $Q(\hat{r})$ test.

An Alternative Test Based on $Q(\hat{r})$

The above results suggest that a closer approximation to the distribution of $Q(\hat{r})$ should be obtainable by appropriately adjusting the mean of the approximating distribution. *Furthermore*, Table 1 shows values of $\mathrm{Var}Q(\hat{r})$ which are nearly twice the mean, suggesting the approximation

$$Q(\hat{r}) \sim \chi^2_{EQ(\hat{r})}$$

with $EQ(\hat{r})$ given by (3.3). Empirical significance levels obtained using this approximation and the criterion $\tilde{Q}(\hat{r})$ are compared in Table 2. The agreement is quite close.

A Power Calculation

The two criteria $Q(\hat{r})$ and $\tilde{Q}(\hat{r})$ differ essentially in the weighting which is applied to the autocorrelations $\hat{r}_k$, $\tilde{Q}(\hat{r})$ giving more emphasis to later autocorrelations. This would perhaps be an advantage if serial correlation occurs at high lags k. However, for large n this difference should

Table 1. Empirical means, variances and significance levels of the statistics $Q(\hat{r})$ and $\tilde{Q}(\hat{r})$; data generated from the model $w_t - .5w_{t-1} = a_t$.

				Level					Level		
n	m	mean	variance	5.0	10.0	25.0	mean	variance	5.0	10.0	25.0
50	10	7.48	13.79	2.3	4.7	13.4	8.82	19.11	5.3	9.5	23.0
	20	13.96	27.50	1.3	2.3	6.4	18.58	47.76	6.1	10.4	23.2
100	10	8.14	16.04	3.4	7.0	18.2	8.83	18.88	5.0	9.9	23.1
	20	16.26	35.45	2.5	5.0	13.1	18.63	46.46	5.8	10.2	22.8
	30	23.53	55.74	1.7	3.6	9.1	28.58	81.71	7.2	11.6	23.4
200	10	8.57	16.76	4.2	8.3	21.5	8.92	18.16	5.0	9.8	23.9
	20	17.46	36.36	3.5	6.9	17.6	18.66	41.51	5.4	10.0	22.7
	30	26.11	56.01	2.9	5.6	14.2	28.66	67.37	5.9	10.5	23.8

Table 2. Empirical significance levels based on the approximations $Q(\hat{r}) \sim \chi^2_{EQ(\hat{r})}$ and $\tilde{Q}(\hat{r}) \sim \chi^2_{m-1}$; data generated from the model $w_t - \phi w_{t-1} = a_t$.

		$Q(\hat{r}) \sim \chi^2_{EQ(\hat{r})}$						$\tilde{Q}(\hat{r}) \sim \chi^2_{m-1}$					
		m = 10			m = 20			m = 10			m = 20		
	Level / ϕ	5.0	10.0	25.0	5.0	10.0	25.0	5.0	10.0	25.0	5.0	10.0	25.0
n=50	.1	4.1	8.3	21.2	4.6	8.1	20.9	4.7	9.3	21.4	5.9	10.1	22.5
	.3	4.2	8.2	21.9	4.5	8.5	21.4	4.9	9.0	22.5	5.9	10.2	22.5
	.5	4.4	8.7	22.2	4.6	8.6	21.7	5.3	9.5	23.0	6.1	10.4	23.2
	.7	4.7	9.5	23.3	5.1	9.6	22.6	5.4	10.1	23.6	6.7	11.3	24.0
	.9	5.6	10.6	25.6	6.3	10.9	24.9	6.3	11.4	25.7	7.9	12.8	25.7
n=100	.1	4.3	8.8	23.4	5.1	9.3	22.2	4.7	9.3	23.5	5.9	10.0	22.7
	.3	4.4	8.5	23.4	5.2	9.0	22.7	4.8	9.1	23.5	6.0	9.8	23.1
	.5	4.4	8.5	23.5	5.3	9.1	22.6	4.7	9.3	23.4	6.0	10.1	22.9
	.7	4.7	9.0	24.1	5.6	9.6	22.7	4.9	9.4	24.0	6.2	10.3	23.2
	.9	5.3	9.6	25.4	6.2	10.2	23.7	5.5	9.9	25.4	7.0	11.2	24.1
n=200	.1	5.0	9.6	24.1	5.2	9.8	22.7	5.2	9.9	24.2	5.5	10.2	23.2
	.3	4.9	9.3	23.8	5.1	9.7	22.6	5.2	9.6	24.0	5.4	10.1	22.8
	.5	4.8	9.6	23.8	5.1	9.4	22.4	5.0	9.8	23.9	5.4	10.0	22.7
	.7	4.8	9.9	24.1	4.9	10.0	22.5	5.0	10.1	24.2	5.3	10.5	22.8
	.9	5.4	10.2	25.8	5.4	10.7	23.3	5.6	10.3	25.8	5.7	11.1	23.5

be rather small. If the type of discrepancies to be expected is known, tests specifically aimed at detecting these discrepancies should be used. Such specific tests will of course be much more powerful. This point is illustrated in Table 3 which empirically compares the power of the overall tests and the method of "overfitting" [Box and Jenkins (1970)]. The results are based on data generated from a second order autoregressive model, with a first order model being fitted to obtain $Q(\hat{r})$ and $\tilde{Q}(\hat{r})$. As might be expected, the overall tests are much less powerful than overfitting which tests the hypothesis that the second order autoregressive coefficient is zero. A smaller value of m improves the power of the overall tests for this particular alternative.

Effect of nonnormality of the a_t's

In developing the overall test, the assumption is made that the innovations a_t in the model are normally distributed. Circumstances occur where this assumption is not true. For example, it is known that stock price innovations often have highly leptokurtic distributions. While Anderson and Walker (1964) show that the asymptotic normality of the r_k's does not require normality of the a_t's, only that their variance exists, results for finite sample sizes are lacking. An empirical investigation was therefore conducted into the behavior of the statistic $\tilde{Q}(\hat{r})$ when the a_t's have i) a double exponential and ii) a uniform distribution. The results, which are given in Table 4, agree closely with those obtained under the normality assumption for the a_t's in Table 1.

5. EXTENSION TO TRANSFER FUNCTION-NOISE MODELS

To check the adequacy of a fitted model of the form

$$w_t = \frac{\omega(B)}{\delta(B)} \alpha_{t-b-1} + \frac{\theta(B)}{\phi(B)} a_t$$

Table 3. Empirical power of the overall tests and the method of overfitting for $n = 100$. Assumed model: $w_t - \phi w_{t-1} = a_t$; true model: $(1 - .7B)(1 - G_2B)w_t = a_t$. Nominal significance level: 5 percent.

Test	G_2 .0	.1	.3	.5	.7	.9
Overfitting	5.3	12.0	59.7	93.8	99.7	99.1
$Q(\hat{r}) \sim \chi^2_{EQ(\hat{r})}$						
m = 10	4.7	6.7	28.6	72.0	96.6	99.9
m = 20	5.6	7.3	24.4	62.8	93.7	99.7
m = 30	6.0	7.7	22.9	58.1	91.7	99.5
$\tilde{Q}(\hat{r}) \sim \chi^2_{m-1}$						
m = 10	4.9	7.0	28.9	71.6	96.2	99.9
m = 20	6.2	8.0	24.7	61.7	93.2	99.6
m = 30	7.0	9.0	23.7	57.0	90.5	99.3

Table 4. Empirical means, variances and significance levels of $\tilde{Q}(\hat{r})$ when the innovations a_t have i) a double exponential and ii) a uniform distribution; data generated from the model $w_t - .5w_{t-1} = a_t$.

i) $a_t \sim$ double exponential ii) $a_t \sim$ uniform

n	m	i) mean	variance	Level 5.0	10.0	25.0	ii) mean	variance	Level 5.0	10.0	25.0
50	10	8.50	18.59	4.7	8.6	20.7	9.01	19.35	5.6	10.0	24.4
	20	17.77	47.00	5.4	8.8	19.6	18.95	52.39	7.3	12.1	24.3
100	10	8.80	18.70	5.0	9.1	22.4	9.11	19.41	5.5	10.8	25.3
	20	18.37	43.62	4.8	9.2	22.0	19.00	47.52	6.4	11.5	25.7
	30	27.94	76.60	6.3	10.1	21.9	28.98	81.72	7.5	12.4	25.3
200	10	8.86	18.78	4.9	10.3	23.9	9.00	19.24	5.6	10.1	25.7
	20	18.60	43.33	5.6	9.7	23.6	18.93	43.99	6.2	11.2	25.2
	30	28.46	69.67	6.0	10.5	23.2	28.94	72.29	6.6	11.6	25.3

where

$$\omega(B) = \omega_0 - \omega_1 B - \cdots - \omega_u B^u$$

$$\delta(B) = 1 - \delta_1 B - \cdots - \delta_v B^v$$

and $\phi(B)$ and $\theta(B)$ are as in (1.1), and where the input series $\{\alpha_t\}$ is assumed to be white noise, it is useful to examine, in addition to the residual autocorrelations $\hat{r}_k$, the crosscorrelations between the residuals and the input series

$$\hat{r}_k^* = \frac{\sum_{t=k+1}^{n} \alpha_{t-k}\hat{a}_t}{\left(\sum_{t=1}^{n} \alpha_t^2 \sum_{t=1}^{n} \hat{a}_t^2\right)^{1/2}} .$$

Pierce (1968, 1972) and Box and Jenkins (1970) propose an overall test for lack of fit in the transfer function $\omega(B)/\delta(B)$ based on approximating the distribution of

$$S(\hat{r}^*) = n \sum_{k=0}^{m} (\hat{r}_k^*)^2$$

by a χ^2_{m-v-u} distribution. However, arguing as above it appears that a criterion of the form

$$\tilde{S}(\hat{r}^*) = n^2 \sum_{k=0}^{m} (n-k)^{-1} (\hat{r}_k^*)^2$$

might be more appropriate. This is suggested by the fact that the distribution of

$$\tilde{S}(r^*) = n^2 \sum_{k=0}^{m} (n-k)^{-1} (r_k^*)^2 ,$$

where

$$r_k^* = \frac{\sum_{t=k+1}^{n} \alpha_{t-k} a_t}{\left(\sum_{t=1}^{n} \alpha_t^2 \sum_{t=1}^{n} a_t^2 \right)^{1/2}},$$

is for large n approximately χ^2_{m+1} since the r_k^*'s are asymptotically normal with mean zero, $\mathrm{Var}(r_k^*) = (n-k)/n^2$ and $\mathrm{Cov}(r_k, r_\ell) = 0$, $\ell \neq k$ [Pierce (1968)].

REFERENCES

Anderson, R. L. (1942). Distribution of the serial correlation coefficients. Ann. Math. Statist. 13, 1-13.

Anderson, T. W. (1971). The Statistical Analysis of Time Series. New York: John Wiley.

Anderson, T. W. & Walker, A. M. (1964). On the asymptotic distribution of the autocorrelations of a sample from a linear stochastic process. Ann. Math. Statist. 35, 1296-1303.

Box, G. E. P. & Jenkins, G. M. (1970). Time Series Analysis, Forecasting and Control. San Francisco: Holden Day.

Box, G. E. P. & Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. J. Am. Statist. Assoc. 65, 1509-26.

Davies, N., Triggs, C. M. & Newbold, P. (1977). Significance levels of the Box-Pierce portmanteau statistic in finite samples. Biometrika (to appear).

Ljung, G. M. & Box, G. E. P. (1976). Studies in the modelling of discrete time series III: A modification of the overall χ^2 test for lack of fit in time series models. Tech. Rep. 477, Department of Statistics, University of Wisconsin.

Pierce, D. A. (1968). Distribution of residual correlations in dynamic/stochastic time series models. Tech. Rep. 173, Department of Statistics, University of Wisconsin.

Pierce, D. A. (1972). Residual correlations and diagnostic checking in dynamic-disturbance time series models. J. Am. Statist. Assoc. 67, 636-40.

Prothero, D. L. & Wallis, K. F. (1976). Modelling macroeconomic time series (with discussion). J. R. Statist. Soc. A 139, 468-500.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER (14) MRC-TSR-1790 (1790) | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle) (6) (9) Technical
ON A MEASURE OF LACK OF FIT IN TIME SERIES MODELS

5. TYPE OF REPORT & PERIOD COVERED
Summary Report, no specific reporting period

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s) (10)
G. M. Ljung and G. E. P. Box

8. CONTRACT OR GRANT NUMBER(s) (15)
DAAG29-75-C-0024,
AFOSR 72-2363D

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Mathematics Research Center, University of Wisconsin
610 Walnut Street
Madison, Wisconsin 53706

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
4 (Probability, Statistics and Combinatorics)

11. CONTROLLING OFFICE NAME AND ADDRESS
See Item 18 below.

12. REPORT DATE (11)
September 1977

13. NUMBER OF PAGES
15 (12) 19p.

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)

15. SECURITY CLASS. (of this report)
UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)
Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

U. S. Army Research Office	Air Force Office of Scientific Research
P. O. Box 12211	Building 410
Research Triangle Park	Bolling AFB
North Carolina 27709	Washington, DC 20332

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
Autoregressive-moving average models
Residual autocorrelations
Tests for lack of fit
Transfer function-noise models

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)
The overall test for lack of fit in autoregressive-moving average models proposed by Box and Pierce (1970) is considered. It is shown that a substantially improved approximation results from a simple modification of this test. Some consideration is given to the power of such tests and their robustness when the innovations are non-normal. Similar modifications in the overall tests used for transfer function-noise models are proposed.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

JOB

221 200