CENTRAL CIRCULATION BOOKSTACKS

The person charging this material is responsible for its renewal or its return to the library from which it was borrowed on or before the Latest Date stamped below. **You may be charged a minimum fee of $75.00 for each lost book.**

Theft, mutilation, and underlining of books are reasons for disciplinary action and may result in dismissal from the University.

TO RENEW CALL TELEPHONE CENTER, 333-8400

UNIVERSITY OF ILLINOIS LIBRARY AT URBANA-CHAMPAIGN

When renewing by phone, write new due date below previous due date.

AUG 01 1997
Preliminary Test Estimation for the Second Order Autoregression

Robert Bobrer  
Department of Statistics  
University of Illinois

Kunjung Lai  
Department of Statistics  
University of Illinois

Thomas A. Yancey  
Department of Economics  
University of Illinois

Bureau of Economic and Business Research  
College of Commerce and Business Administration  
University of Illinois at Urbana-Champaign
Preliminary Test Estimation for the
Second Order Autoregression

Robert Bohrer
Kunjung Lai
Department of Statistics

Thomas A. Yancey
Department of Economics
PRELIMINARY TEST ESTIMATION FOR THE SECOND ORDER AUTOREGRESSION

Robert Bohrer         Kunjing Lai         Thomas A. Yancey
Dept. of Statistics  Dept. of Statistics  Dept. of Economics
Univ. of Illinois    Univ. of Illinois    Univ. of Illinois
Champaign 61820      Champaign 61820      Champaign 61820

Key Words and Phrases: preliminary test; autoregressive model; asymptotic squared error risk; Monte Carlo.

ABSTRACT

Consequences of preliminary test model identification procedures in time series analysis are examined in the context of a squared error risk function. Asymptotic risk comparisons and Monte Carlo studies are used in comparing these procedures.

1. INTRODUCTION

In order to obtain parameter estimators for statistical models with desirable statistical properties, assumptions are made about the properties of the random elements and the functional form of the model generating the data. Since the investigator is unsure of the applicability of these assumptions, null hypotheses reflecting the assumptions are tested, and the model is retained or altered and reestimated depending on the outcomes of these tests. This
procedure leads to preliminary test estimators whose properties include the outcomes of these preliminary tests. Most of the work on preliminary test estimators to date has concentrated on the standard linear statistical model and some extensions of it. A series of papers and books, Miyazaki, Judge and Yancey (1986), Judge and Yancey (1986), Yancey, Judge and Bohrer (1989), Giles and Clarke (1989), Judge, Bohrer and Yancey (1990), Giles (1991) and Yancey and Bohrer (1992) concentrate on inequality constraint conjectures for the parameters of the linear model and extensions with various assumptions about the generation of the error term.

In this paper we are interested in determining the order of an autoregressive model based on testing hypotheses to decide which parameters in the model are zero, and the model is altered if necessary. Since pretesting is commonly done to determine the correct autoregressive model in time series analysis in econometrics, we are interested in comparing the risk function of the pretest estimator with that of the maximum likelihood estimator, mle. We use a second order autoregressive model, AR(2), which may be an AR(1) or an AR(0) model depending on where we are in the parameter space.

The paper is organized as follows: Section 2 discusses the model and the estimation procedures. Section 3 presents the risk of the mle procedure for the initial model. Section 4 discusses the risk of a pseudo pretest estimator which honors the outcome of the hypothesis test, but uses only the initial mle. Section 5 presents the true pretest estimator risk where the model is altered and reestimated if the hypothesis is accepted. Section 6 is the summary and conclusions.
2. DEFINITION OF FRAMEWORK AND PROCEDURES

For \( t = 1, \ldots, n \), \( X_t \) is a stationary normal second order autoregression, that is,

\[
X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + Z_t,
\]

where \( Z_t \) are identically distributed, independent normal \( N(0, \sigma^2) \) random variables. Stationarity requires that the zeros of \( s^2 - \phi_1 s - \phi_2 = 0 \) be both within the unit circle. Denote maximum likelihood estimators for the autoregressive parameters as \( \hat{\phi}_1 \) and \( \hat{\phi}_2 \). Brockwell and Davis (1987) derived the asymptotic distribution for these maximum likelihood estimators as unbiased with \( \text{Var}(\hat{\phi}_1) = \text{Var}(\hat{\phi}_2) = (1 - \phi_2^2)/n \) and \( \text{Cov}(\hat{\phi}_1, \hat{\phi}_2) = -\phi_1(1 + \phi_2)/n \).

Denote \( \phi = (\phi_1, \phi_2) \) and \( p = (p_1, p_2) \). The goal is to estimate \( \phi \) to minimize the square error risk of the estimator \( p \), that is, the risk of the estimator \( p \) at \( \phi \) is \( R(p, \phi) = E(p_1 - \phi_1)^2 + E(p_2 - \phi_2)^2 \). Three procedures are considered. Procedure zero is not a preliminary test estimator; rather it merely estimates \( \phi \) by the maximum likelihood estimator; rather it merely estimates \( \phi \) by the maximum likelihood estimators \( \hat{\phi} = (\hat{\phi}_1, \hat{\phi}_2) \).

Procedure one gives preliminary test estimator \( \hat{\phi} = (\hat{\phi}_1, \hat{\phi}_2) \). First, it does a test, \( T_0 \), of size \( \alpha \) for the null hypothesis that \( \phi_1 = \phi_2 = 0 \). If \( T_0 \) accepts the null hypothesis, then set \( \hat{\phi}_1 = \hat{\phi}_2 = 0 \). If \( T_0 \) rejects the null hypothesis, then the test \( T_1 \) of size \( \alpha \) is done for the first order, i.e., that \( \phi_2 = 0 \). If \( T_1 \) accepts, then set \( \hat{\phi}_1 = \hat{\phi}_1 \) and \( \hat{\phi}_2 = 0 \). If \( T_1 \) rejects, then set \( \hat{\phi}_1 = \hat{\phi}_1 \) and \( \hat{\phi}_2 = \hat{\phi}_2 \).

The tests \( T_0 \) and \( T_1 \) are chi-square tests based on the asymptotic distribution for \( \hat{\phi}_1 \) and \( \hat{\phi}_2 \), namely, if the null hypothesis that \( \phi_1 = \phi_2 = 0 \) is true, then the large sample distribution of \( \hat{\phi}_1 \) and \( \hat{\phi}_2 \) is that of independent normal
Thus the acceptance region for \( T_0 \) is \( C = [\hat{\phi}_1^2 + \hat{\phi}_2^2 \leq \chi^2_0(2)/n] \).

Also under the hypothesis of first order, the large sample distribution of \( \hat{\phi}_2 \) is normal \( N(0,1/n) \). Thus the acceptance region for \( T_1 \) is the strip \( S = [\hat{\phi}_2^2 \leq \chi^2_0(1)/n] \). Let \( C' \) be the complement of \( C \) and \( S' \) be the complement of \( S \). Thus \((\hat{\phi}_1, \hat{\phi}_2) = (0,0)\) on \( C \); \((\hat{\phi}_1, \hat{\phi}_2) = (\hat{\phi}_1, 0)\) on \( C' \cap S \); and \((\hat{\phi}_1, \hat{\phi}_2) = (\hat{\phi}_1, \hat{\phi}_2)\) on \( C' \cap S' \).

Procedure two is very similar to procedure one, but leads to two pretest estimators. The only difference is in the case \( T_0 \) rejects and \( T_1 \) accepts. In procedure two \( \phi_1 \) is replaced by \( \hat{\phi}_1 \) which is the maximum likelihood estimator for \( \phi_1 \) assuming a first order model rather than using \( \hat{\phi}_1 \). Thus, procedure two builds the estimate \((\hat{\phi}_1, \hat{\phi}_2) = (0,0)\) on \( C \); \((\hat{\phi}_1, \hat{\phi}_2) = (\hat{\phi}_1, 0)\) on \( C' \cap S \); and \((\hat{\phi}_1, \hat{\phi}_2) = (\hat{\phi}_1, \hat{\phi}_2)\) on \( C' \cap S' \).

3. ASYMPTOTIC RISK OF PROCEDURE ZERO

The asymptotic risk of procedure zero is

\[
R(\hat{\phi}, \phi) = \text{Var}(\hat{\phi}_1) + \text{Var}(\hat{\phi}_2) = 2(1 - \phi_2^2)/n.
\]

4. ASYMPTOTIC RISK OF PROCEDURE ONE

For \( A \) any two-dimensional set, denote the indicator function, \( I_A(p) \), which takes the value one when \( p \) falls in \( A \) and is zero otherwise. The risk \( R(\hat{\phi}, \phi) \) of procedure one is the expectation with respect to the asymptotic distribution for \( \hat{\phi} \) of the loss

\[
L(\hat{\phi}, \phi) = (\phi_1^2 + \phi_2^2)I_{C}(\phi) + ((\phi_1 - \phi_1)^2 + \phi_2^2)I_{C'\cap S}(\phi) + ((\phi_1 - \phi_1)^2 + (\phi_2 - \phi_2)^2)I_{C'\cap S'}(\phi)
\]
And this risk can be calculated by numerical integration.

4.1 NOTES ON COMPUTATION

Calculations are accomplished using an algorithm for bivariate integration. The outer integral is over $\phi_2$, and the plane is partitioned into nine sets including $C$, $C' \cap S \cap [\phi_1 > 0]$, $C' \cap S \cap [\phi_1 < 0]$, and six sets convenient for partitioning $C' \cap S'$. One such bivariate integration algorithm is that of Tavernini as described by Milton (1972). Tavernini’s algorithm uses the Newton-Cotes method and has an approximate bound for the error. Subdivision is terminated when successive Simpson rule approximations are sufficiently close to each other. This leaves the possibility of premature termination if the integrand is minuscule over most of the region of the integration. For successful use, integration must be confined to subsets of the desired integration region on which the integrand is not too close to zero.

4.2 EXAMPLE

Let $\alpha=0.05$ and the sample size be $n=100$. The asymptotic risk $R(\hat{\phi}, \phi)$ is evaluated for stationary pairs $(\phi_1, \phi_2)$ such that $\phi_i = -0.9, -0.8, \ldots, 0.8, 0.9$. Stationary pairs are those for which $\pm \phi_1 + \phi_2 < 1$. Of more interest than the asymptotic risk of procedure one is the comparison of this risk with the asymptotic risk of procedure zero. This is measured by relative regret $(R(\hat{\phi}, \phi) - R(\hat{\phi}, \phi))/R(\hat{\phi}, \phi)$ for procedure one and shown in Figure 1.
The preliminary test procedure one is preferred to the non-preliminary test procedure zero, where the relative regret is negative. Figure 1 shows a contour plot of the relative regret in the \((\phi_1, \phi_2)\) plane. The fact that procedures zero and one are nearly identical at \(\phi\) values away from \(C \cup S\) is shown by the small values of regret there. For other \(\phi\) values, preliminary test procedure one is preferred to procedure zero near the origin and along the \(\phi_1 = 0\) axis. This follows because procedure one takes advantage of preliminary testing and making the correct decision in the hypothesis test. In an annulus around the boundary of region \(C = [|| \hat{\phi}_1 || \leq 0.245]\), the regret is positive because of the probability of using \(\phi_1 = 0\) when it is not.

FIG. 1: Contour plot of the regret of procedure one.
4.3 MONTE CARLO VERIFICATION OF THE ASYMPTOTIC RISK OF PROCEDURE ONE

It is of interest to assess whether the asymptotic risk $R(\dot{\phi}, \phi)$ is a good approximation to the risk with finite sample sizes. And to this end, 100 samples of $n$ observations are generated for each of the 55 systematically sampled $\phi$ points shown in Figure 2.

For each sample of $n$ points the loss is calculated and the risk is estimated by the average $\hat{R}(\dot{\phi}, \phi)$ of these losses. In addition, a sample variance $V_1$ of these 100 losses is calculated. The measure of the fit for the data to the asymptotic theory is given by the $p$-level of the $\chi^2(1)$ statistic 

$$100(\hat{R}(\dot{\phi}, \phi) - R(\dot{\phi}, \phi))^2 / V_1.$$ 

The larger the $p$-level, the better the fit. If the asymptotic theory were exact, these $p$-levels would be uniformly distributed.
on the unit interval. Using the maximum likelihood fitting procedure using the *arimamle* command for autoregressions from Splus on the Sun 3/50 system, the 55 *p*-levels are calculated for *n*=100 and *n*=400. The empirical cumulative distribution for these *p*-levels is shown in Figure 3 along with the ramp shape cumulative distribution for the uniform on the unit interval. For both *n*=100 and *n*=400, the Kolmogorov-Smirnov test of size 0.05 would accept the hypothesis of the uniform distribution. It is somewhat surprising that the Kolmogorov-Smirnov statistic is somewhat larger for the larger sample size.

![Figure 3: The Empirical cumulative distributions of *p*-levels.](image)

Two computing comments are note-worthy. First, for each *φ* value and for both *n*=100 and *n*=400, the 100 runs of procedure one requires about 3 hours of computing time. Second, a similar Monte Carlo study was conducted with the older Splus version of maximum likelihood fitting by
using arima.mle command, and the empirical distribution of p-levels was found to reject the hypothesis of uniform distribution at significance level 0.01.

5. SOME MONTE CARLO EVIDENCE ABOUT PROCEDURE TWO

Procedure two differs from procedure one only for \( \hat{\phi} \) values which lie on \( C' \cap S \). Thus one would not expect much difference in \( R(\hat{\phi}, \phi) - R(\hat{\phi}, \phi) \) for \( \phi \) values away from \( C' \cap S \). With \( \alpha=0.05 \) and \( n=100 \), a Monte Carlo study for the difference of risks was conducted over the 55 points in Figure 2. One hundred runs were observed. A statistic for testing of no difference between the risks of procedures one and two is

\[
100(R(\hat{\phi}, \phi) - R(\hat{\phi}, \phi))^2/(V_1 + V_2),
\]

where \( V_i \) is the sample variance of the 100 samples of procedure \( i \). Under the null hypothesis, this statistic has the \( \chi^2(1) \) distribution, and \( p \)-levels are uniformly distributed on the unit interval. Under the null hypothesis, one expects 5.5 \( p \)-levels less than 0.10. Among the 55 points, 9 have \( p \)-levels less than 0.10. Most of these were near the strip \( S \) and had \( R(\hat{\phi}, \phi) \) less than \( \hat{R}(\hat{\phi}, \phi) \). For further evidence on comparisons of procedures one and two on the region \( C' \cap S \), 100 Monte Carlo runs were done for each of 10 points with \( \phi_2 = 0 \) and \( \phi_1 \) in \( C' \cap S \). For all 10 points, the risk of procedure two was less than that of procedure one by from 5 to 43 percent. In summary, it appears that procedures one and two are not significantly different except for \( \phi \) values near the strip \( S \) where procedure two does somewhat better than procedure one.
6. SUMMARY AND CONCLUSIONS

We have considered the time series problem of determining the order of an AR process by testing the significance of coefficients of lagged variables as is commonly done to identify the model. Starting with a stationary AR(2) model, procedures zero, one and two were used and the risk functions under squared error loss were obtained. It was found that although there was little difference between the risks of procedures one and two over most of the parameter space, procedure two did lead to a smaller risk in regions where they differed. For both procedures, the properties of the estimators are recognized explicitly which is not done with the usual diagnostic and identification procedure used in time series analysis. Further analysis of the pretesting done in the identification of time series models should consider the moving average, MA, and ARMA models. One difficulty with this extension is that the asymptotic theory used here for the AR(2) model breaks down for the ARMA(1,1) model.
BIBLIOGRAPHY

Brockwell, P. J. and Davis, R. A. (1986). *Time Series: Theory and Methods*, Springer-Verlag, New York.

Giles, D. E. A. and Clarke, J. A. (1989). Preliminary-test estimation of the scale parameter in a mis-specified regression model. *Economics Letters*, 30, 201-205.

Giles, J. A. (1990). Pretesting in the misspecified regression model. *Communications in Statistics: Theory and Methods*, 20 (10), 3221-3238.

Judge, G. G., Bohrer, R. and Yancey, T. A. (1990). Some statistical implications of multivariate inequality constrained testing. *Communications in Statistics: Theory and Methods*, 19 (2), 413-430.

Judge, G. G. and Yancey, T. A. (1986). *Improved Methods of Inference in Econometrics*, North-Holland Publishing, Amsterdam.

Milton, R. C. (1972). Computer evaluation of the multivariate normal integral. *Technometrics*, 14, 881-889.

Miyazaki, S., Judge, G. G. and Yancey, T. A. (1989). Estimation of location parameters under nonnormal errors and quadratic loss. *Journal of Business and Economic Statistics*, 4(2), 263-268.

Yancey, T. A., Judge, G. G. and Bohrer, R. (1989). Sampling performance of some joint one-sided preliminary test estimators under squared error loss. *Econometrica*, 51(5), 1221-1228.

Yancey, T. A. and Bohrer, R. (1992). "Risk and power for inequality pretest estimators: general case in two dimensions" in William E. Griffiths, Helmut Lutkepohl and Mary Ellen Bock (eds) *Reading in Econometric Theory and Practice: A Volume in Honor of George Judge*. North-Holland Publishing, Amsterdam.
