Advances and Applications in Statistics

Published Online: May 2016
http://dx.doi.org/10.17654/AS048030169
Volume 48, Number 3, 2016, Pages 169-183 **ISSN: 0972-3617**

THE SHQC CRITERION FOR ORDER SELECTION IN AUTOREGRESSIVE MODELS

Enas Gawdat Yehia

Department of Statistics
Faculty of Commerce
Al-Azhar University
Girl's Branch
Tafahna Al-Ashraf, Egypt

Abstract

We propose the SHQC as a new order selection criterion based on both two information criteria, Schwarz information criterion (SIC) and Hannan-Quinn information criterion (HQC) for the selection in autoregressive models. The performance of our proposed criterion was investigated through a simulation study by comparing it with other well known order selection criteria. The simulation results show that the proposed criterion, SHQC is much better than other existing criteria in identifying the correct model order for moderate to large samples. Moreover, as the sample size increases, the SHQC converges to the true order which confirms the consistency of it and the probabilities of both under and over fitting of it are always the least among other criteria for all cases. Therefore, we can use the SHQC criterion as a safe alternative to any criterion.

Received: December 23, 2015; Revised: February 13, 2016; Accepted: February 20, 2016
2010 Mathematics Subject Classification: 62M10.
Keywords and phrases: autoregressive models, order selection, information criteria, SIC, HQC, simulation.
Communicated by Guvenc Arslan

1. Introduction

Selection of the model order is an important part in time series analysis that aims to determine the true order for the model to be fitted. There are many criteria used in the literature to determine the order of autoregressive (AR) model which is one of the most common approaches for modelling of time series data.

The common selection criteria based on information criteria (IC) due to Kullback and Leibler [23] who introduced asymmetric divergence known as I-divergence or the directed divergence to measure the distance between the true model and the candidate model and the best model which loses the least information relative to other models. These criteria such as Akaike information criterion (AIC) Akaike [3], Schwarz information criterion (SIC) Schwarz [25], Hannan-Quinn information criterion (HQC) Hannan and Quinn [14], Akaike information corrected criterion (AICc) Hurvich and Tsai [16] and so on. A briefly discussion about these criteria is existed in Beveridge and Oickle [4], Brockwell and Davis [6] and Burnham and Anderson [7]. Moreover, many of other selection criteria had been proposed based on IC. Chen et al. [10] proposed a resampling method for model selection of autoregressive models. Karimi [19] derived a new order selection criterion by using new approximations for the expectations of residual variance and prediction error. Haggag [12] proposed new criteria based on AIC by taking the average of it with each of some criteria such as AICc and HQC in linear regression models.

Another divergence due to Kullback [22] called *J-divergence* which is symmetric divergence based on the sum of two directed divergences for the true and the candidate models. New criteria based on Kullback's symmetric divergence had been proposed such as Kullback information criterion (KIC) due to Cavanaugh [8] and a corrected version of it called *Kullback information corrected criterion* (KICc) which introduced by Boosarawongse and Chongcharoen [5]. The KICc criterion had been justified for linear regression model Cavanaugh [9], nonlinear regression with normal errors

Kim and Cavanaugh [21] and autoregressive models Hafidi and Mkhadri [11].

In general, selecting the model order requires minimizing the estimated I-divergence or J-divergence. All of criteria involve the summation of two terms, the first is a measure of goodness of fit for the model to the data, and the second is a penalty term for overfitting which occurs when the criterion value for the true model of order p more than its value for the candidate model of order $p + k$, where $k > 0$ is the number of overfitted order.

In other words, order selection criteria are based on minimizing the following loss function:

$$\delta(p) = h(\hat{\sigma}_a^2) + g(n,\ p), \tag{1}$$

where $\hat{\sigma}_a^2$ is an estimated variance of white noise a, n is the sample size, p is the order of the candidate model, $h(\hat{\sigma}_a^2)$ is a measure of goodness of fit for the candidate model to the data, and $g(n,\ p)$ is a penalty term for overfitting depending on n and p only.

The objective of this paper is to provide a new order selection criterion which its penalty term based on the penalty term of SIC and HQC criteria for the selection of autoregressive models, and to investigate the performance of our proposed criterion with respect to other well known criteria through a simulation study.

In Section 2, we discuss the order selection criteria for autoregressive models. In Section 3, the proposed criterion in this paper is introduced. This is followed by a simulation study to investigate the performance of this criterion with respect to other existing criteria in Section 4, and the simulation results are discussed in Section 5. Finally, the conclusions are presented in Section 6.

2. Autoregressive Order Selection Criteria

We consider the stationary AR model of order p, denoted by AR(p) for a time series $\{Y_t,\ t = 0, \pm 1, \pm 2, \ldots\}$,

$$\Phi_p(B)Y_t = a_t, \tag{2}$$

where $\Phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$ is called an *autoregressive operator* and B is the backward shift operator such that $B^k Y_t = Y_{t-k}$, all the roots of the polynomial $\Phi_p(B) = 0$ lie outside the unit circle. The $\{\phi_i, i = 1, 2, ..., p\}$ are autoregressive parameters and $\{a_t\}$ are independent identically distributed normal random variables with zero mean and finite variance σ^2 which can be estimated by $\hat{\sigma}_a^2 = (n - p - 1)^{-1} \sum_{t=1}^{n} \hat{a}_t^2$. Also, the $\{a_t\}$ satisfy that $E(a_t^4) < \infty$.

There are many criteria used in determining the order of autoregressive model. First, Akaike [1] proposed the final prediction error (FPE) criterion for AR(p) based on minimizing the one-step ahead mean square forecast error and defined as follows:

$$\text{FPE} = \hat{\sigma}_a^2 \left(\frac{n + p}{n - p} \right). \tag{3}$$

This criterion is asymptotically efficient but it is not consistent due to Akaike [2]. Therefore, Akaike [3] suggested a new criterion based on Kullback-Leibler information known as AIC, and defined as

$$\text{AIC} = \ln(\hat{\sigma}_a^2) + n^{-1}(2p). \tag{4}$$

Jones [18] showed that AIC is asymptotically unbiased and efficient, however, it has a tendency to overfit Shibata [26].

Another criterion based on Bayesian information was introduced by Schwarz [25] known as SIC. It also called *Schwarz-Rissanen information criterion* due to Rissanen [24] who arrived at the same criterion independently, it has the following form:

$$\text{SIC} = \ln(\hat{\sigma}_a^2) + n^{-1}(p \ln(n)). \tag{5}$$

According to another form of penalty term, Hannan and Quinn [14] suggested a new criterion denoted by HQC as follows:

$$\mathrm{HQC} = \ln(\hat{\sigma}_a^2) + n^{-1}(2p(\ln(\ln(n)))). \tag{6}$$

To speed up the convergence of this criterion, Hannan and Rissanen [15] replaced the term $\ln(\ln(n))$ by $\ln(n)$. However, Kavalieris [20] found that this modification led to overfit the true order.

On the other hand, Hurvich and Tsai [16] suggested more extreme penalty to avoid the overfitting of the AIC through a new criterion known as Akaike information corrected criterion (AICc) which can be expressed as

$$\mathrm{AICc} = \ln(\hat{\sigma}_a^2) + \left(\frac{1 + p/n}{1 - (p+2)/n}\right), \tag{7}$$

but it tends to overfit when the sample size increases Hurvich and Tsai [17].

Indeed, Cavanaugh [8] introduced a new order selection criterion, KIC as an asymptotically unbiased estimate of Kullback's symmetric divergence as follows:

$$\mathrm{KIC} = \ln(\hat{\sigma}_a^2) + n^{-1}(3p). \tag{8}$$

As for AIC, KIC is strongly biased in small samples, therefore, a corrected version of KIC, KICc was introduced by Boosarawongse and Chongcharoen [5], and defined as

$$\mathrm{KICc} = \ln(\hat{\sigma}_a^2) + \frac{[(n+p)(n-p) + (n-p-2)]}{(n-p-2)(n-p)}, \tag{9}$$

which is an approximately unbiased correction of KIC in small samples, but it is not consistent Karimi [19].

3. The SHQC for Autoregressive Models Selection

We propose a new order selection criterion denoted by SHQC, for AR(p) model based on SIC and HQC to combine the strengths of them, which their penalty term lead to consistent criterion where

$$g_S(n,\ p) = p\ln(n), \tag{10}$$

is the penalty term of SIC which make the minimize of SIC converges to *p* with probability one and

$$g_{HQ}(n,\ p) = 2p(\ln(\ln(n))), \tag{11}$$

is the penalty term of HQC which lead to a consistent criterion as the sample size increases.

The penalty term of the proposed criterion, SHQC has been chosen more extreme penalty term for extra parameters to avoid overfitting by taking the sum of two penalty terms of SIC and HQC. Consequently, the SHQC criterion can be expressed as follows:

$$\text{SHQC} = \ln(\hat{\sigma}_a^2) + n^{-1}p[\ln(n) + 2\ln(\ln(n))], \tag{12}$$

where the minimum of the above function gives the estimate of the optimum model order. As we see from equations (10) and (11), SIC is always greater than HQC, thus the difference between SHQC and HQC is always greater than the difference between SHQC and SIC. Also, for all *p*, as *n* increases, the SHQC has the most rapid increase between SIC and HQC. Hence, the probability of overfitting of SHQC is always the least between these criteria.

Theorem. *For SHQC, the estimated order* $\hat{p}$, *is weakly consistent if the last term from equation* (12) *is replaced by* $n^{-1}pc_n$, *where* c_n *increases to infinity*.

Proof. Under general conditions of Section 2, and from Hannan [13] we can see that c_n converges to infinity as $n \to \infty$ and $\lim_{n\to\infty} \frac{c_n}{n} = 0$, where $c_n = [\ln(n) + 2\ln(\ln(n))]$. Hence, the resulting $\hat{p}$ converges to *p* as $n \to \infty$.

4. Simulation Study

To investigate the performance of SHQC as a new criterion for the selection of AR models with respect to various criteria under study including

FPE, AIC, AICc, SIC, HQC, KIC and KICc, we generated 1000 data series from each AR(1), AR(2) and AR(3) models according to equation (2) at $p = 1, 2, 3,$ respectively, for sample sizes 25, 50, 100 and 500 in each model with specific parameters were chosen to satisfy the stationary conditions through the following models:

$$\mathrm{AR}(1) : Y_t = 0.9Y_{t-1} + a_t, \tag{13}$$

$$\mathrm{AR}(2) : Y_t = -0.1Y_{t-1} + 0.8Y_{t-2} + a_t, \tag{14}$$

$$\mathrm{AR}(3) : Y_t = 1.3Y_{t-1} - 1.2Y_{t-2} + 0.7Y_{t-3} + a_t, \tag{15}$$

where $\{a_t\}$ are $I.I.D.N(0, 1)$ random variables. In all our simulations, we estimated the parameters and the residual variance of the candidate models by the least squares method. After that, we used each of above selection criteria to determine the estimated order $\hat{p},$ in each sample. Also, for every sample in these models, the maximum order was cut-off of ten and the frequency distribution of selected model orders for the existing criteria was obtained for each case. Thus, the probability of estimating the true order denoted by $P(\hat{p} = p),$ and defined as

$$P(\hat{p} = p) = \frac{\text{number of times } (\hat{p} = p)}{1000}. \tag{16}$$

The probability of overfitting the true order denoted by $P(\hat{p} > p),$ which is given by

$$P(\hat{p} > p) = \frac{\text{number of times } (\hat{p} > p)}{1000}, \tag{17}$$

and the probability of underfitting the true order denoted by $P(\hat{p} < p),$ and defined as follows:

$$P(\hat{p} < p) = \frac{\text{number of times } (\hat{p} < p)}{1000}. \tag{18}$$

All these simulation experiments are carried out by using MATHCAD program.

5. Simulation Results

The frequency of the selected orders by the existing criteria at different sample sizes for AR(1), AR(2) and AR(3) models was shown in Tables 1-3, respectively. Moreover, the relative frequency curves of the correct model order selections of various criteria for each model as the sample size increases were plotted in Figures 1, 3 and 6, respectively. Also, the probabilities of overfitting the true order for AR(1), AR(2) and AR(3) models were presented in Figures 2, 4 and 7, respectively. As well as, the probabilities of underfitting the true order for AR(2) and AR(3) models were presented in Figures 5 and 8, respectively.

Table 1. Frequencies of the estimated order for an AR(1)

		Criteria							
n	$\hat{p}$	FPE	AIC	AICc	SIC	HQC	SHQC	KIC	KICc
25	1	796	779	786	868	798	961	868	874
	2	117	117	143	87	106	36	89	101
	3	35	36	44	17	31	2	19	18
	4	23	24	20	10	21	0	10	5
	5-10	29	44	7	18	44	1	14	2
50	1	808	797	717	920	862	978	877	842
	2	111	111	138	65	91	21	85	105
	3	39	41	71	11	27	1	24	33
	4	22	26	37	4	16	0	11	16
	5-10	20	25	37	0	4	0	3	4
100	1	800	791	650	962	906	986	904	829
	2	101	105	125	36	67	13	68	101
	3	46	49	81	1	16	1	16	39
	4	20	19	51	1	6	0	7	15
	5-10	33	36	93	0	5	0	5	16
500	1	809	808	610	982	914	999	891	810
	2	96	97	127	14	62	1	71	96
	3	43	43	76	3	17	0	25	44
	4	23	23	51	0	4	0	8	22
	5-10	29	29	136	1	3	0	5	28

We see from Table 1 and from Figures 1 and 2, that for sample sizes, the proposed criterion, SHQC performs better than other criteria. It has the least probability of overfitting, whereas the other criteria have a tendency to overfit the true model. In addition, as the sample size increases, the performance of SHQC followed by SIC and HQC in identifying the correct model order improves, whereas AIC shows the worst performance at small size of $n = 25,$ as well as, AICc has the worst performance at moderate to large sample sizes of $n = 50, 100, 500,$ respectively.

Table 2. Frequencies of the estimated order for an AR(2)

		Criteria							
n	$\hat{p}$	FPE	AIC	AICc	SIC	HQC	SHQC	KIC	KICc
25	1	39	39	33	55	40	107	54	52
	2	792	740	810	827	762	862	837	863
	3	96	104	110	74	95	24	75	70
	4	36	45	39	24	40	5	23	13
	5-10	37	72	8	20	63	2	11	2
50	1	1	1	1	2	1	3	1	1
	2	817	798	724	918	864	977	883	855
	3	110	116	148	62	93	18	85	102
	4	42	48	68	15	32	2	22	33
	5-10	30	37	59	3	10	0	9	9
100	1	0	0	0	0	0	0	0	0
	2	803	794	659	950	891	992	890	824
	3	107	107	148	46	75	6	76	104
	4	42	46	68	2	19	1	19	38
	5-10	48	53	125	2	15	1	15	34
500	1	0	0	0	0	0	0	0	0
	2	780	779	612	977	914	997	873	785
	3	122	125	141	19	63	3	87	121
	4	49	46	80	3	18	0	22	48
	5-10	49	50	167	1	5	0	18	46

From Table 2 and from Figures 3, 4 and 5, we note that for small sample, $n = 25$ the KICc and SHQC have the best performance which performs similarly at selecting the correct model order. On the other hand, for moderate to large sample sizes, the SHQC has the best performance followed by SIC, then KIC at $n = 50$ and followed by SIC, then HQC at $n = 100$ and 500. Indeed the AIC criterion gives the worst selection of *p* at $n = 25,$ while the AICc is the worst at $n = 50, 100, 500.$ Moreover, the tendency to overfit the true model for SHQC, SIC and HQC decreases as *n* increases. Also, for all criteria, the probability of underfitting decreases as *n* increases.

The results of Table 3 demonstrate that for small sample, at $n = 25,$ all criteria have a tendency to overfit the true model, but the KICc criterion gives the best selection of *p*. While for moderate to large sample sizes $n = 50, 100, 500$ respectively, the SHQC criterion outperforms other criteria in selecting the true order (see Figure 6), as well as, the overfit probability of SHQC is always the least and decreases when the sample size increases as shown in Figure 7. On the other hand, the probability of underfitting decreases as *n* increases for all criteria (see Figure 8). Also, we note that AIC gives the worst selection of *p* at small sample, but for moderate to large samples, the AICc is the worst performance among other criteria.

Table 3. Frequencies of the estimated order for an AR(3)

		Criteria							
n	$\hat{p}$	FPE	AIC	AICc	SIC	HQC	SHQC	KIC	KICc
25	1	18	16	13	32	17	99	31	32
	2	81	65	75	96	69	142	97	103
	3	724	683	795	727	685	721	740	811
	4	90	101	89	69	96	29	70	45
	5	23	31	19	17	28	1	17	8
	6-10	64	104	9	59	105	8	45	1

50	1	0	0	0	1	0	4	0	0
	2	0	0	0	4	1	14	3	1
	3	836	825	782	941	873	963	904	883
	4	95	95	120	40	72	18	60	74
	5	33	39	49	11	28	1	20	26
	6-10	36	41	49	3	26	0	13	16
100	1	0	0	0	0	0	0	0	0
	2	0	0	0	0	0	0	0	0
	3	810	793	696	959	898	992	897	841
	4	105	107	133	32	69	8	70	95
	5	36	39	58	5	15	0	15	29
	6-10	49	61	113	4	18	0	18	35
500	1	0	0	0	0	0	0	0	0
	2	0	0	0	0	0	0	0	0
	3	798	796	641	985	925	997	881	807
	4	112	112	135	13	59	2	87	111
	5	39	39	73	1	8	1	18	40
	6-10	51	53	151	1	8	0	14	42

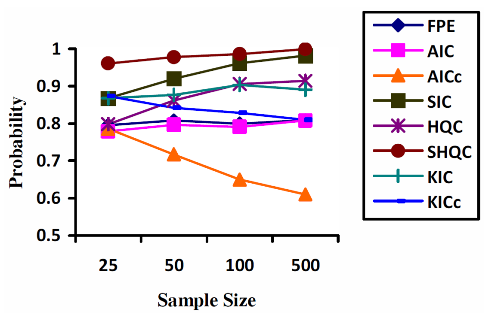

Figure 1. Probability of correctly estimating the true order for AR(1).

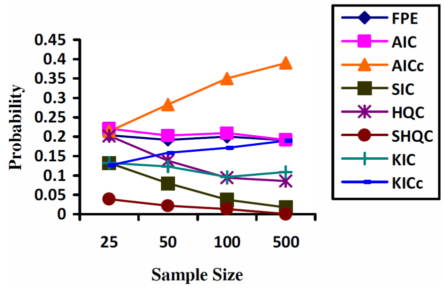

Figure 2. Probability of overfitting for AR(1).

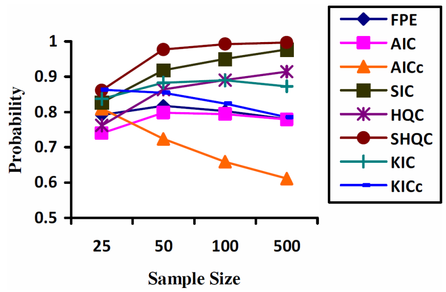

Figure 3. Probability of correctly estimating the true order for AR(2).

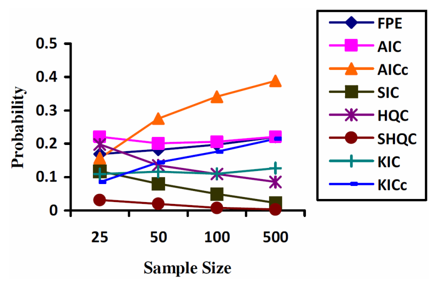

Figure 4. Probability of overfitting for AR(2).

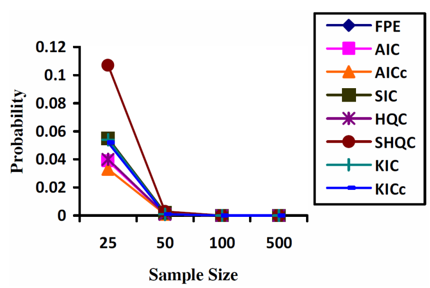

Figure 5. Probability of underfitting for AR(2).

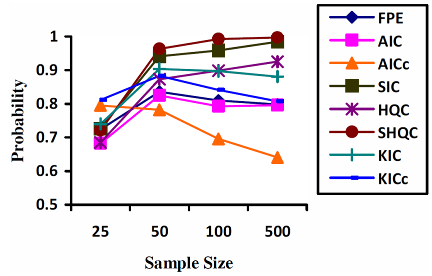

Figure 6. Probability of correctly estimating the true order for AR(3).

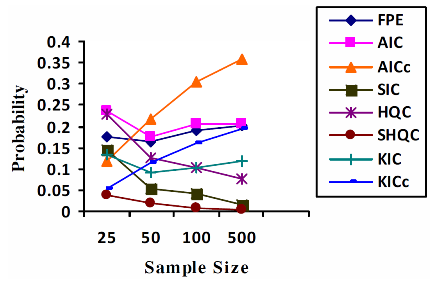

Figure 7. Probability of overfitting for AR(3).

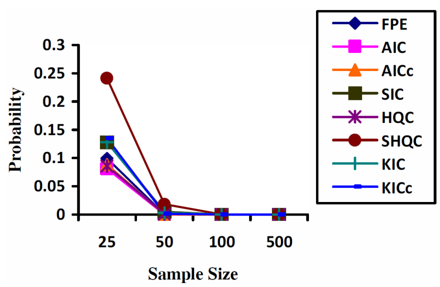

Figure 8. Probability of underfitting for AR(3).

6. Conclusions

In this paper, we have proposed the SHQC criterion as a new order selection criterion based on both SIC and HQC criteria for the selection of autoregressive models. Moreover, a simulation investigation was carried out to compare the performance of our proposed criterion with the existing order selection criteria. The results showed that SHQC has the best performance in identifying the correct model order for moderate to large samples in all models and for small sample in AR(1) model. Indeed, for all cases, the tendency to underfit and overfit the true order for SHQC decreases as the sample size increases. Therefore, the SHQC can be used as a safe alternative to any criterion.

Acknowledgement

The author thanks the anonymous referees for their valuable suggestions which led to the improvement of the manuscript.

References

[1] H. Akaike, Fitting autoregressive models for prediction, Ann. Inst. Statist. Math. 21 (1969), 243-247.

[2] H. Akaike, Statistical predictor identification, Ann. Inst. Statist. Math. 22 (1970), 203-217.

[3] H. Akaike, Information theory and extention of the maximum likelihood principle, B. N. Petrov and F. Csaki, eds., Proceedings of the 2nd International Symposium on Information Theory, Akadeniai Kiado, Budapest, 1973, pp. 267-281.

[4] S. Beveridge and C. Oickle, A comparison Box-Jenkins and objective methods for determining the order of a non-seasonal ARMA model, J. Forecasting 13 (1994), 419-434.

[5] R. Boosarawongse and S. Chongcharoen, Model selection criteria for autoregressive model, The Proceeding of the 8th APDSI, China, 2003.

[6] P. J. Brockwell and R. A. Davis, Introduction to Time Series and Forecasting, Springer Verlag, New York, 1996.

[7] K. P. Burnham and D. R. Anderson, Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, 2nd ed., Springer Verlag, New York, 2002.

[8] J. E. Cavanaugh, A large-sample model selection criterion based on Kullback's symmetric divergence, Statist. Probab. Lett. 44 (1999), 333-344.

[9] J. E. Cavanaugh, Criteria for linear model selection based on Kullback's symmetric divergence, New Zealand J. Statist. 46 (2004), 257-274.

[10] C. Chen, R. A. Davis, P. J. Brockwell and Z. D. Bai, Order determination for autoregressive processes using resampling methods, Stat. Sinica 3 (1993), 481-500.

[11] B. Hafidi and A. Mkhadri, A corrected Akaike criterion based on Kullback's symmetric divergence: applications in time series, multiple and multivariate regression, Comput. Statist. Data Anal. 50 (2006), 1524-1550.

[12] M. M. Haggag, New criteria of model selection and model averaging linear regression models, Amer. J. Theoret. Appl. Statist. 3 (2014), 148-166.

[13] E. J. Hannan, The estimation of the order of an ARMA process, Ann. Statist. 8 (1980), 1071-1081.

[14] E. J. Hannan and B. G. Quinn, The determination of the order of an autoregression, J. Roy. Statist. Soc. Ser. B 41 (1979), 190-195.

[15] E. J. Hannan and J. Rissanen, Recursive estimation of mixed autoregressive moving average order, Biometrika 69 (1982), 81-94.

[16] C. M. Hurvich and C. L. Tsai, Regression and time series model selection in small samples, Biometrika 76 (1989), 297-307.

[17] C. M. Hurvich and C. L. Tsai, A corrected Akaike information criterion for vector autoregressive model selection, J. Time Ser. Anal. 14 (1993), 271-279.

[18] R. H. Jones, Fitting autoregressions, J. Amer. Statist. Assoc. 70 (1975), 590-592.

[19] M. Karimi, Finite sample criteria for autoregressive model order selection, Iran. J. Sci. Technol. B. 31 (2007), 329-344.

[20] L. Kavalieris, A note on estimating the dimension of a model, Ann. Stat. 6 (1991), 461-464.

[21] H.-J. Kim and J. E. Cavanaugh, Model selection criteria based on Kullback information measures for nonlinear regression, J. Stat. Plan. Inference 134 (2005), 332-349.

[22] S. Kullback, Information Theory and Statistics, Dover, New York, 1968.

[23] S. Kullback and R. A. Leibler, On information and sufficiency, Ann. Math. Stat. 22 (1951), 79-86.

[24] J. Rissanen, Modelling by shortest data description, Automatica 14 (1978), 467-471.

[25] G. Schwarz, Estimating the dimension of a model, Ann. Statist. 6 (1978), 461-464.

[26] R. Shibata, Selection of the order of an autoregressive model by Akaike's information criterion, Biometrika 63 (1976), 117-126.