Thailand Statistician
January 2023; 21(1): 48-67
http://statassoc.or.th
Contributed paper

New Criterion for Selection in Regression Model

Warangkhana Riansut*
Department of Mathematics and Statistics, Faculty of Science, Thaksin University, Phatthalung, Thailand.
*Corresponding author; e-mail: warang27@gmail.com

Received: 27 June 2021
Revised: 14 February 2022
Accepted: 4 April 2022

Abstract

The aim of this study is to propose the new criterion for selection in regression model, called NIC, and then compare the effectiveness of NIC with ten model selection criteria, namely, AIC, BIC, HQIC, AICc, AICu, HQICc, KIC, $KICc_{C}$, $KICc_{SB}$, and $KICc_{HM}$. The conditions for simulation were the differences in sample size, number of parameters in the model, regression coefficient, error variance, and distribution of independent variables. The results of the study showed that, for small to moderate sample sizes and the true model is somewhat difficult to identify, the performances of AIC and HQIC perform the best. However, they can identify the true model actually less accurate. As a result, the observed L_2 efficiency suggests that NIC is the best criterion for small to moderate sample sizes. For the large sample size and the true model is somewhat difficult to identify, the appropriate criteria are AIC and BIC. When the sample sizes are small to moderate and the true model can be specified more easily, the appropriate criterion is NIC. For the large sample size and the true model can be specified more easily, the appropriate criterion is BIC.

Keywords: Model selection criterion, performance, frequency of order being selected, observed L_2 efficiency.

1. Introduction

In regression analysis, the choice of an appropriate model from a class of candidate models to characterize the study data is a key issue. In real life, we may not know what the true model is, but we hope to find a model that is a reasonably accurate representation. A model selection criterion represents a useful tool to judge the propriety of a fitted model by assessing whether it offers an optimal balance between goodness of fit and parsimony. The first model selection criterion to gain widespread acceptance was Akaike information criterion, AIC (Akaike 1974). This serves as an asymptotically unbiased estimator of a variant of Kullback's directed divergence between the true and the candidate models. The directed divergence, also known as the I-divergence or the relative entropy, assesses the dissimilarity between two statistical models. Other well-known criterion was subsequently introduced and studied such as, Bayesian Information Criterion (BIC) (Schwarz 1978), Hannan and Quinn Information Criterion (HQIC) (Hannan and Quinn 1979), a corrected version of AIC (AICc) (Hurvich and Tsai 1989), a corrected version of AICc (AICu) (McQuarrie et al. 1997), a corrected version of HQIC (HQICc) (McQuarrie and Tsai 1998), Kullback Information Criterion (KIC) (Cavanaugh 1999), a corrected version of KIC by Cavanaugh ($KICc_C$) (Cavanaugh 2004), a corrected version of KIC by Seghouane and Bekara ($KICc_{SB}$) (Seghouane and Bekara 2004), and a corrected version of KIC by Hafidi and Mkhadri ($KICc_{HM}$) (Hafidi and Mkhadri 2006). Although AIC remains arguably the most widely used model selection criterion, BIC, HQIC, AICc, AICu, HQICc, KIC, $KICc_C$, $KICc_{SB}$, and $KICc_{HM}$ are the popular competitors. In fact, BIC is often preferred over AIC by practitioners who find appeal in either its Bayesian justification or its tendency to choose more parsimonious models than AIC (Neath and Cavanaugh 1997). Likewise, KIC is a symmetric measure which combines the information in two related, though distinct measures; it functions as a gauge of model disparity that is arguably more sensitive than AIC that corresponds to only individual components (Cavanaugh 1999; Cavanaugh 2004). When the sample size is small or the dimension of candidate model is large relative to the sample size, AIC suffers from a large negative bias. As a result, it has the problem of high probability of overfitting. In this setting, Hurvich and Tsai (1989) proposed a corrected AIC (AICc) for linear and non-linear regression. The AICc has been extended in a number of directions, including autoregressive moving average modeling (Hurvich et al. 1990), vector autoregressive modeling (Hurvich and Tsai 1993), and multivariate regression modeling (Bedrick and Tsai 1994). Although, Hurvich and Tsai (1989) showed that AICc is an unbiased estimator for the expected Kullback's directed divergence, McQuarrie et al. (1997) indicated that AICc tend to overfitted when the sample size increased. Therefore, they proposed AICu that is an approximately unbiased estimator of Kullback's directed divergence. It provided better model choices than AICc for moderate to large sample sizes except when the true model is of infinite order. Further, the KIC tends to underestimate the Kullback's symmetric divergence in small-sample applications, as indicated by Cavanaugh (2004), Seghouane and Bekara (2004), and Hafidi and Mkhadri (2006). Therefore, they proposed KICc in order to cope with this problem. However, AIC, BIC, HQIC, AICc, AICu, HQICc, KIC, $KICc_C$, $KICc_{SB}$, and $KICc_{HM}$ are advantages and disadvantages for each subject differing as shown in the literature reviews as follows. Keerativibool (2014) study the penalty functions of AIC, BIC, and KIC, which can unify their formulas as $\mathrm{APIC}\alpha = \log\left(\hat{\sigma}^2\right) + \alpha\left(p+1\right)/n,$ called adjusted penalty information criterion. The theoretical results show that, the probability of overfitting tends to zero and the signal-to-noise ratio tends to strong if the value of α tends to infinity. However, the simulation results show that, when the true model is weakly identifiable, the small value of α gives a high probability of correct order being selected. But, if the true model is very difficult to detect, the observed L_2 efficiency is a meaningful measurement than the probability of order selected. The observed L_2 efficiency suggests the large value of α causes the high efficiency of $\mathrm{APIC}\alpha$ which indicates that the correct model is also the closet model, except when the true model can be specified more easily and sample sizes are moderate to large, then the small value of α is preferable. For the strongly identifiable true model, the large value of α performs well, whereas if the regression coefficients are not large enough and the sample sizes are small to moderate, the value of α should be moderate. Keerativibool and Siripanich (2017) unify the justifications of AIC, AICc, KIC, $KICc_C$, $KICc_{SB}$, and $KICc_{HM}$. The results show that $KICc_C$ has the strongest penalty function under the condition $\left(1-p/n\right)\exp\left(p/n\right)<1,$ followed, respectively, by $KICc_{SB}$, $KICc_{HM}$, KIC and AIC. Also, KIC is greater than AICc under the condition $n-p>2,\ n>3$ and p belongs to the set of $\left[-1,\ n/3-2\right],$ but AICc always greater than AIC. The result of simulation shown that, the model selection with a larger penalty term may lead to underfitting and slow convergence while a smaller penalty term may lead to overfitting and inconsistency. When the sample size is small to moderate and the true model is somewhat difficult to identify, the performances of AIC and AICc are better than others. However, they can identify the true model actually less accurate. When the sample size is large, the performances of all model selection criteria are insignificant difference, but all criteria can identify the true model still less accurate. As a result, this paper used the observed L_2 efficiency to assess model selection criteria performances. On the average, this measure suggests that in a weakly identifiable true model, whether the sample size is small or large, $KICc_C$ is the best criterion. For the small sample size and the true model can be specified more easily with small error variance, every model selection criteria still have the ability to select the correct model. If the error variance increase, the performances of all model selection criteria are bad. When the sample sizes are moderate to large, KICc performs the best, it can identify a lot of true model for small error variance. But, if the error variance increases and the sample size is not large enough, all model selection criteria can identify a little true model.

The aims of this paper are to establish new criteria for regression model selection, called new information criterion (NIC). The performances of NIC are examined by the extensive simulation study against AIC, BIC, HQIC, AICc, AICu, HQICc, KIC, $KICc_C$, $KICc_{SB}$, and $KICc_{HM}$, under the difference various circumstances: sample sizes (n), regression coefficients (β), variances of error term (σ^2), and distribution of independent variables. All criteria performances are examined by a consistent measure which is a measure of counting the frequency of order being selected. Particularly for the case of small to moderate sample sizes, we use an efficient measure which is the observed L_2 efficiency. This is a useful measure when the criteria do not accomplish the correct model. The remainder of this paper is organized as follows. In Section 2, we propose the new criterion for selection in regression model, called NIC. Simulation study and results for 1,000 samples of multiple regression models to examine the performances of all criteria are shown in Section 3. Finally, Section 4 is the conclusions.

2. Methodology

The true and the candidate models to consider in this study are, respectively, given by

$$\mathbf{y} = \mathbf{X}_0\boldsymbol{\beta}_0 + \boldsymbol{\varepsilon}_0,\ \boldsymbol{\varepsilon}_0 \sim N_n\left(\mathbf{0},\ \sigma_0^2\mathbf{I}_n\right), \tag{1}$$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},\ \boldsymbol{\varepsilon} \sim N_n\left(\mathbf{0},\ \sigma^2\mathbf{I}_n\right), \tag{2}$$

where $\mathbf{y}$ is an $n\times 1$ dependent random vector of observations, $\mathbf{X}_0$ and $\mathbf{X}$ are $n\times p_0$ and $n\times p$ matrices of independent variables with full-column rank, respectively, $\boldsymbol{\beta}_0$ and $\boldsymbol{\beta}$ are $p_0\times 1$ and $p\times 1$ parameter vectors of regression coefficients, respectively, $\boldsymbol{\varepsilon}_0$ and $\boldsymbol{\varepsilon}$ are $n\times 1$ noise vectors, n represents the sample size, p_0 and p represent the number of parameters in the true and approximate models, respectively, including the constants. The true model is assumed to be correctly specified or overfitted by all the candidate models. This means that $\boldsymbol{\beta}_0$ has p_0 nonzero entries with $0 < p_0 \le p$ and the rest of the $(p - p_0)$ entries are equal to zero. The $(p+1)\times 1$ vector of parameters is $\boldsymbol{\theta}_0 = \left[\boldsymbol{\beta}_0'\ \ \sigma_0^2\right]'$ and the maximum likelihood estimator of $\boldsymbol{\theta}_0$ is $\hat{\boldsymbol{\theta}} = \left[\hat{\boldsymbol{\beta}}'\ \ \hat{\sigma}^2\right]'$ where

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{y} \ \text{ and } \ \hat{\sigma}^2 = \frac{1}{n}\left(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}\right)'\left(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}\right). \tag{3}$$

The unbiased estimator of σ_0^2 is

$$s^2 = \frac{1}{n-p}\left(\mathbf{y}-\mathbf{X}\hat{\beta}\right)'\left(\mathbf{y}-\mathbf{X}\hat{\beta}\right) = \frac{n\hat{\sigma}^2}{n-p}. \tag{4}$$

The observed distance L_2 or squared error distance, scaled by $1/n$, between the true and the candidate models is defined as (McQuarrie et al. 1997)

$$L_2(p) = \frac{1}{n}\left(\beta_0-\hat{\beta}\right)'\mathbf{X}'\mathbf{X}\left(\beta_0-\hat{\beta}\right). \tag{5}$$

Observed L_2 efficiency is defined by the ratio

$$\text{Observed } L_2 \text{ efficiency } = \frac{\min_{1\le p\le P} L_2(p)}{L_2(p_s)}, \tag{6}$$

where P is the class of all candidate models, p is the rank of fitted candidate model, and p_s is the model accomplished by specific model selection criterion. The closer the selected model is to the true model, the higher the efficiency. Therefore, the best model selection criterion will select a model which yields high efficiency even in small samples or if the true model is weakly identifiable.

As mentioned in Keerativibool and Siripanich (2017),

$$KICc_C = \log\left(\hat{\sigma}^2\right) + \log\left(\frac{n}{n-p}\right) + \frac{\left[(n-p)(2p+3)-2\right]}{(n-p-2)(n-p)}, \tag{7}$$

is the best criterion, it can identify a lot of true models for small error variance. Including, the observed L_2 efficiency suggests that in a weakly identifiable true model, whether the sample size is small or large, $KICc_C$ is still the best criterion. Also, McQuarrie et al. (1997) replaced the maximum likelihood estimator of σ_0^2 in (3) by the unbiased estimator in (4) in

$$AICc = \log\left(\hat{\sigma}^2\right) + \frac{2(p+1)}{n-p-2}, \tag{8}$$

therefore

$$AICu = \log\left(s^2\right) + \frac{2(p+1)}{n-p-2} \tag{9}$$

had a greater penalty for overfitting, especially as the sample size increase. From the two-above knowledges, we replace $\hat{\sigma}^2$ in (3) by s^2 in (4) to the $KICc_C$ formula in (7), yield the new criterion for selection in regression model, called NIC as follows:

$$NIC = \log\left(s^2\right) + \log\left(\frac{n}{n-p}\right) + \frac{\left[(n-p)(2p+3)-2\right]}{(n-p-2)(n-p)}. \tag{10}$$

In the next section, we compare the performance of NIC in (10) against the well-known model selection criteria: AIC in (11), BIC in (12), HQIC in (13), AICc in (8), AICu in (9), HQICc in (14), KIC in (15), $KICc_C$ in (7), $KICc_{SB}$ in (16), and $KICc_{HM}$ in (17), whose formulas are as follows:

$$AIC = \log\left(\hat{\sigma}^2\right) + \frac{2(p+1)}{n} \tag{11}$$

$$BIC = \log\left(\hat{\sigma}^2\right) + \frac{(p+1)\log(n)}{n} \tag{12}$$

$$HQIC = \log\left(\hat{\sigma}^2\right) + \frac{2(p+1)\log\log(n)}{n} \tag{13}$$

$$\text{HQICc} = \log\left(\hat{\sigma}^2\right) + \frac{2(p+1)\log\log(n)}{n-p-2} \tag{14}$$

$$\text{KIC} = \log\left(\hat{\sigma}^2\right) + \frac{3(p+1)}{n} \tag{15}$$

$$\text{KICc}_{\text{SB}} = \log\left(\hat{\sigma}^2\right) + \frac{(p+1)(3n-p-2)}{n(n-p-2)} + \frac{p}{n(n-p)} \tag{16}$$

$$\text{KICc}_{\text{HM}} = \log\left(\hat{\sigma}^2\right) + \frac{(p+1)(3n-p-2)}{n(n-p-2)}. \tag{17}$$

3. Simulation Study

In order to examine the performance of NIC against AIC, BIC, HQIC, AICc, AICu, HQICc, KIC, KICc_C, KICc_SB, and KICc_HM, we generated the true multiple regression models in (1) as the (18) until (21). We assume that all regressors are fixed, not a random variable.

Model I represents a very weakly identifiable true model with the true order $p_0 = 5$:

$$y = 1 + 0.5X_1 + 0.4X_2 + 0.3X_3 + 0.2X_4 + \varepsilon_0. \tag{18}$$

Model II represents a weakly identifiable true model with the true order $p_0 = 3$:

$$y = 1 + 0.5X_1 + 0.4X_2 + \varepsilon_0. \tag{19}$$

Model III represents a very strongly identifiable true model with the true order $p_0 = 3$:

$$y = 1 + 2X_1 + 2X_2 + \varepsilon_0. \tag{20}$$

Model IV represents a strongly identifiable true model with the true order $p_0 = 5$:

$$y = 1 + 2X_1 + 2X_2 + 2X_3 + 2X_4 + \varepsilon_0. \tag{21}$$

Model I and Model II represent the weakly identifiable true models which mean they are not easily identified compared to the strongly identifiable true models such as Model III and Model IV. The error terms ε_0 in (18) until (21) are assumed to be normally distributed with zero mean and variances σ_0^2 equal to one of the following three levels: 0.25, 1, and 9. For all true models in (18) until (21), we consider 1,000 samples for three levels of the sample sizes which are n = 15 (small), n = 30 (moderate), and n = 100 (large). The steps for simulation and all results are as follows.

3.1 Generate the error terms ε_0 in (18) until (21) about 100,000 observations from normal population with zero mean and variances equal to 0.25, 1, and 9.

3.2 Split the series of error terms in step 1 into 1,000 samples, each of which consists of three levels of sample sizes, n = 15, 30, 100 observations.

3.3 Generate the independent variables X_1 to X_6 about 100,000 observations from standard normal population $(\mathbf{X} \sim N(0, \mathbf{I}))$ and uniform population as follows:

$$X_1 \sim U(5, 10),\ X_2 \sim U(10, 20),\ X_3 \sim U(7, 9),\ X_4 \sim U(6, 11),\ X_5 \sim U(9, 19),\ X_6 \sim U(4, 8)$$

where the relevant independent variables of Model I and Model IV are X_1 to X_4 and irrelevant independent variables are X_5, X_6, whereas the relevant independent variables of Model II and Model III are X_1, X_2 and irrelevant independent variables are X_3 to X_6.

3.4 Split the series of independent variables in step 3 into 1,000 samples, each of which consists of 15, 30, 100 observations.

3.5 Use the corresponding relevant independent variables obtained in Step 4 and the error terms obtained in Step 2 to construct the dependent variables described in (18) until (21).

3.6 Use the concept of sequentially nested fashion as a potential candidate model in (2); i.e., we consider 6 subsets, 1) constant and X_1 (p = 2), 2) constant, X_1, X_2 (p = 3), 3) constant, X_1, X_2, X_3 (p = 4), 4) constant, X_1, X_2, X_3, X_4 (p = 5), 5) constant, X_1, X_2, X_3, X_4, X_5 (p = 6), 6) constant, X_1, X_2, X_3, X_4, X_5, X_6 (p = 7). For each subset, calculate NIC in (10) against AIC in (11), BIC in (12), HQIC in (13), AICc in (8), AICu in (9), HQICc in (14), KIC in (15), $KICc_C$ in (7), $KICc_{SB}$ in (16), and $KICc_{HM}$ in (17).

The subset with the minimum value of model selection criterion can be classified to be the best model. Due to the large number of subsets, it is impractical to summarize the individual models chosen. Hence, Tables 1 to 6 summarize p = rank(X) of the selected subset to be three groups: the selected order less than p is called underfitted order, the selected order equals to p is called correct order, and the selected order greater than p is called overfitted order. Tables 7 to 10 display the candidate models that are closest to the true model in the L_2 sense Equations (5) and (6) for small and moderate sample size. The ave. and S.D. L_2 rows denote, respectively, the average and standard deviation of observed L_2 efficiency in (6) over 1,000 samples.

Table 1 Percentage of model selection for small sample size (n = 15) and the independent variables have a standard normal distribution

σ_0^2	Model	Order	AIC	BIC	HQIC	AICc	AICu	HQICc	KIC	$KICc_C$	$KICc_{SB}$	$KICc_{HM}$	NIC
0.25	I	order < p	34.2	46.7	34.2	82.4	89.8	82.4	51.3	90.2	88.5	87.6	95.0
		p = 5	**30.2**	29.9	30.0	16.1	9.7	16.1	28.9	9.3	11.0	11.9	4.7
		order > p	35.6	23.4	35.8	1.5	0.5	1.5	19.8	0.5	0.5	0.5	0.3
0.25	II	order < p	5.1	9.4	5.0	18.2	27.7	17.9	10.6	27.8	26.8	26.1	37.0
		p = 5	47.3	58.4	47.3	73.1	67.8	**73.2**	62.1	68.0	68.3	68.5	60.7
		order > p	47.6	32.2	47.7	8.7	4.5	8.9	27.3	4.2	4.9	5.4	2.3
0.25	III	order < p	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		p = 5	52.0	68.2	51.9	90.9	94.2	90.9	72.6	94.4	93.8	93.4	**97.0**
		order > p	48.0	31.8	48.1	9.1	5.8	9.1	27.4	5.6	6.2	6.6	3.0
0.25	IV	order < p	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		p = 5	60.9	70.0	60.8	95.1	97.6	95.1	73.4	97.8	96.7	96.5	**98.5**
		order > p	39.1	30.0	39.2	4.9	2.4	4.9	26.6	2.2	3.3	3.5	1.5
1	I	order < p	56.5	72.0	56.3	97.1	98.8	97.1	77.1	98.9	98.7	98.6	99.5
		p = 5	14.5	10.9	**14.6**	2.2	1.1	2.2	9.8	1.0	1.1	1.1	0.5
		order > p	29.0	17.1	29.1	0.7	0.1	0.7	13.1	0.1	0.2	0.3	0.0
1	II	order < p	29.5	43.1	29.4	63.6	73.5	63.5	47.6	74.0	72.9	72.2	81.9
		p = 5	29.4	31.0	29.4	31.3	24.1	**31.4**	31.2	23.8	24.7	25.2	17.4
		order > p	41.1	25.9	41.2	5.1	2.4	5.1	21.2	2.2	2.4	2.6	0.7
1	III	order < p	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.1	0.1	0.3
		p = 5	51.1	67.1	50.9	91.3	95.2	91.1	72.4	95.8	94.8	94.4	**96.8**
		order > p	48.9	32.9	49.1	8.7	4.7	8.9	27.6	4.1	5.1	5.5	2.9
1	IV	order < p	0.1	0.1	0.1	0.8	2.0	0.8	0.3	2.1	1.9	1.8	3.5
		p = 5	58.1	69.1	57.9	95.0	**96.4**	94.9	72.4	**96.4**	95.9	95.8	95.9
		order > p	41.8	30.8	42.0	4.2	1.6	4.3	27.3	1.5	2.2	2.4	0.6
9	I	order < p	68.2	83.4	68.0	98.3	99.7	98.2	87.0	99.7	99.6	99.3	100.0
		p = 5	**7.6**	4.3	**7.6**	1.1	0.2	1.2	3.3	0.2	0.3	0.5	0.0
		order > p	24.2	12.3	24.4	0.6	0.1	0.6	9.7	0.1	0.1	0.2	0.0
9	II	order < p	50.4	65.9	50.4	85.5	91.6	85.5	71.7	91.8	91.1	90.5	96.0
		p = 5	**13.0**	12.1	**13.0**	10.8	6.8	10.8	11.3	6.8	7.1	7.5	3.3
		order > p	36.6	22.0	36.6	3.7	1.6	3.7	17.0	1.4	1.8	2.0	0.7
9	III	order < p	11.2	17.5	11.1	30.7	42.1	30.7	20.4	42.7	41.3	40.8	52.6
		p = 5	44.7	54.0	44.6	**63.3**	55.3	63.2	55.1	54.8	55.7	56.0	46.1
		order > p	44.1	28.5	44.3	6.0	2.6	6.1	24.5	2.5	3.0	3.2	1.3
9	IV	order < p	16.4	24.5	16.4	61.3	75.1	61.0	29.0	76.1	73.0	71.8	87.2
		p = 5	45.6	**48.1**	45.4	36.7	24.0	36.9	47.8	23.0	25.9	27.1	12.4
		order > p	38.0	27.4	38.2	2.0	0.9	2.1	23.2	0.9	1.1	1.1	0.4

Note: Boldface type indicates the maximum percentage of correct order being selected.

Table 2 Percentage of model selection for small sample size (n = 15) and the independent variables have a uniform distribution

σ_0^2	Model	Order	AIC	BIC	HQIC	AICc	AICu	HQICc	KIC	$KICc_C$	$KICc_{SB}$	$KICc_{HM}$	NIC
0.25	I	order < p	66.3	81.1	65.9	98.4	99.4	98.4	85.2	99.5	99.2	99.2	99.9
		p = 5	7.7	6.0	**7.8**	1.1	0.5	1.1	5.0	0.4	0.6	0.6	0.1
		order > p	26.0	12.9	26.3	0.5	0.1	0.5	9.8	0.1	0.2	0.2	0.0
0.25	II	order < p	41.1	56.9	40.6	79.2	85.7	78.9	61.9	86.1	85.3	85.2	92.0
		p = 5	18.4	**18.6**	18.4	17.3	12.7	17.4	18.5	12.3	13.1	13.2	7.2
		order > p	40.5	24.5	41.0	3.5	1.6	3.7	19.6	1.6	1.6	1.6	0.8
0.25	III	order < p	0.4	1.1	0.4	3.2	4.8	3.2	1.3	4.9	4.4	4.2	8.5
		p = 5	49.8	65.1	49.6	86.5	89.4	86.3	68.2	**89.6**	89.4	89.3	88.0
		order > p	49.8	33.8	50.0	10.3	5.8	10.5	30.5	5.5	6.2	6.5	3.5
0.25	IV	order < p	0.7	2.0	0.7	13.0	20.9	13.0	3.0	22.3	18.4	17.6	37.2
		p = 5	59.0	67.0	58.7	**83.9**	77.0	**83.9**	68.8	75.7	79.4	79.9	61.6
		order > p	40.3	31.0	40.6	3.1	2.1	3.1	28.2	2.0	2.2	2.5	1.2
1	I	order < p	69.9	83.1	69.8	99.1	99.9	99.0	87.3	99.9	99.8	99.7	99.9
		p = 5	**9.1**	5.9	**9.1**	0.7	0.1	0.8	4.7	0.1	0.2	0.3	0.1
		order > p	21.0	11.0	21.1	0.2	0.0	0.2	8.0	0.0	0.0	0.0	0.0
1	II	order < p	48.7	64.5	48.4	83.6	90.1	83.3	70.7	90.4	90.1	89.5	93.8
		p = 5	13.3	**13.9**	13.3	12.0	7.5	12.1	12.9	7.2	7.5	8.0	5.3
		order > p	38.0	21.6	38.3	4.4	2.4	4.6	16.4	2.4	2.4	2.5	0.9
1	III	order < p	15.5	23.5	15.5	37.6	50.5	37.5	26.5	51.4	49.2	48.3	61.4
		p = 5	37.4	44.7	37.3	**53.5**	45.8	**53.5**	47.3	45.0	46.5	47.0	36.8
		order > p	47.1	31.8	47.2	8.9	3.7	9.0	26.2	3.6	4.3	4.7	1.8
1	IV	order < p	21.3	32.8	21.2	72.6	84.3	72.4	38.1	85.0	82.7	82.0	92.4
		p = 5	39.2	**41.5**	39.1	24.9	15.3	25.1	40.1	14.6	16.6	17.3	7.5
		order > p	39.5	25.7	39.7	2.5	0.4	2.5	21.8	0.4	0.7	0.7	0.1
9	I	order < p	71.0	83.6	71.0	98.4	99.5	98.4	87.4	99.6	99.4	99.4	99.9
		p = 5	**7.9**	5.7	**7.9**	1.2	0.5	1.2	4.5	0.4	0.5	0.5	0.1
		order > p	21.1	10.7	21.1	0.4	0.0	0.4	8.1	0.0	0.1	0.1	0.0
9	II	order < p	50.7	69.6	50.6	87.3	93.1	87.2	75.7	93.3	92.6	92.3	96.0
		p = 5	**13.3**	12.0	**13.3**	8.7	5.4	8.7	10.3	5.3	5.7	5.8	3.6
		order > p	36.0	18.4	36.1	4.0	1.5	4.1	14.0	1.4	1.7	1.9	0.4
9	III	order < p	46.3	61.9	46.2	82.1	88.4	82.1	67.0	88.6	88.2	87.4	92.6
		p = 5	**14.9**	14.7	**14.9**	14.1	9.4	14.1	13.6	9.2	9.4	10.1	6.4
		order > p	38.8	23.4	38.9	3.8	2.2	3.8	19.4	2.2	2.4	2.5	1.0
9	IV	order < p	63.5	78.4	63.1	97.6	99.1	97.6	82.9	99.1	98.9	98.8	99.6
		p = 5	11.5	7.7	**11.8**	1.8	0.7	1.8	6.2	0.7	0.8	0.8	0.4
		order > p	25.0	13.9	25.1	0.6	0.2	0.6	10.9	0.2	0.3	0.4	0.0

Note: Boldface type indicates the maximum percentage of correct order being selected.

Table 3 Percentage of model selection for moderate sample size (n = 30) and the independent variables have a standard normal distribution

σ_0^2	Model	Order	AIC	BIC	HQIC	AICc	AICu	HQICc	KIC	$KICc_C$	$KICc_{SB}$	$KICc_{HM}$	NIC
0.25	I	order < p	20.2	37.3	25.7	34.4	48.0	42.1	32.3	49.0	46.9	46.6	62.0
		p = 5	53.2	52.1	53.8	**55.4**	47.3	51.6	54.4	46.6	47.9	48.2	35.4
		order > p	26.6	10.6	20.5	10.2	4.7	6.3	13.3	4.4	5.2	5.2	2.6
0.25	II	order < p	0.4	1.4	0.5	1.0	1.9	1.2	1.0	1.9	1.9	1.8	3.5
		p = 5	65.5	84.7	73.8	80.0	89.1	87.3	79.9	89.4	88.8	88.4	**91.9**
		order > p	34.1	13.9	25.7	19.0	9.0	11.5	19.1	8.7	9.3	9.8	4.6
0.25	III	order < p	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		p = 5	66.1	87.1	74.6	83.6	91.2	88.9	83.5	91.4	91.0	90.9	**95.9**
		order > p	33.9	12.9	25.4	16.4	8.8	11.1	16.5	8.6	9.0	9.1	4.1
0.25	IV	order < p	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		p = 5	70.2	87.3	76.9	88.3	94.1	93.0	84.7	94.3	93.5	93.5	**96.7**
		order > p	29.8	12.7	23.1	11.7	5.9	7.0	15.3	5.7	6.5	6.5	3.3
1	I	order < p	56.3	80.2	66.9	76.3	88.8	84.8	75.5	89.4	88.1	87.6	94.5
		p = 5	**22.4**	13.7	19.2	17.1	9.5	11.9	16.1	8.9	9.9	10.3	5.2
		order > p	21.3	6.1	13.9	6.6	1.7	3.3	8.4	1.7	2.0	2.1	0.3
1	II	order < p	19.5	36.2	23.7	27.8	40.9	35.8	30.8	41.4	39.7	39.0	51.9
		p = 5	48.4	51.0	51.2	**54.6**	52.2	53.9	51.6	52.0	52.3	52.9	44.7
		order > p	32.1	12.8	25.1	17.6	6.9	10.3	17.6	6.6	8.0	8.1	3.4
1	III	order < p	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		p = 5	63.2	84.1	73.2	81.1	89.8	86.9	80.6	90.0	89.3	88.9	**95.1**
		order > p	36.8	15.9	26.8	18.9	10.2	13.1	19.4	10.0	10.7	11.1	4.9
1	IV	order < p	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		p = 5	69.8	86.4	76.9	87.6	93.4	91.6	82.9	93.4	92.7	92.4	**96.0**
		order > p	30.2	13.6	23.1	12.4	6.6	8.4	17.1	6.6	7.3	7.6	4.0
9	I	order < p	80.9	95.2	87.4	93.9	98.1	96.6	93.2	98.2	97.8	97.4	99.3
		p = 5	**7.0**	2.8	5.5	3.6	1.3	2.2	3.4	1.3	1.5	1.8	0.4
		order > p	12.1	2.0	7.1	2.5	0.6	1.2	3.4	0.5	0.7	0.8	0.3
9	II	order < p	58.5	80.2	68.1	73.2	84.0	79.9	76.1	84.1	83.1	82.8	90.7
		p = 5	**17.1**	13.5	16.4	16.6	12.1	14.5	14.5	12.0	12.4	12.5	7.9
		order > p	24.4	6.3	15.5	10.2	3.9	5.6	9.4	3.9	4.5	4.7	1.4
9	III	order < p	2.0	4.2	2.5	2.8	5.5	3.9	3.3	5.8	5.3	5.2	8.9
		p = 5	66.8	83.9	74.5	81.8	87.5	86.3	80.9	87.3	87.4	87.3	**88.3**
		order > p	31.2	11.9	23.0	15.4	7.0	9.8	15.8	6.9	7.3	7.5	2.8
9	IV	order < p	3.2	9.0	4.9	7.3	14.6	11.5	6.9	14.8	13.6	13.1	22.9
		p = 5	68.2	77.4	72.7	**80.2**	78.1	79.8	76.8	77.9	78.8	79.3	73.7
		order > p	28.6	13.6	22.4	12.5	7.3	8.7	16.3	7.3	7.6	7.6	3.4

Note: Boldface type indicates the maximum percentage of correct order being selected.

Table 4 Percentage of model selection for moderate sample size (n = 30) and the independent variables have a uniform distribution

σ_0^2	Model	Order	AIC	BIC	HQIC	AICc	AICu	HQICc	KIC	$KICc_{C}$	$KICc_{SB}$	$KICc_{HM}$	NIC
0.25	I	order < p	72.5	92.0	81.5	89.0	96.4	94.6	88.3	96.5	96.3	96.2	98.8
		p = 5	**11.5**	5.1	8.6	7.3	2.9	4.2	6.5	2.8	3.0	3.1	1.1
		order > p	16.0	2.9	9.9	3.7	0.7	1.2	5.2	0.7	0.7	0.7	0.1
0.25	II	order < p	44.6	65.6	52.2	57.6	69.8	64.7	60.6	70.2	69.4	68.7	79.6
		p = 5	29.4	26.8	29.9	**31.9**	25.7	29.2	29.4	25.4	25.9	26.0	18.9
		order > p	26.0	7.6	17.9	10.5	4.5	6.1	10.0	4.4	4.7	5.3	1.5
0.25	III	order < p	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		p = 5	68.5	89.3	77.8	86.2	93.0	90.7	86.0	93.2	92.6	92.3	**96.5**
		order > p	31.5	10.7	22.2	13.8	7.0	9.3	14.0	6.8	7.4	7.7	3.5
0.25	IV	order < p	0.0	0.1	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.3
		p = 5	69.7	85.4	76.5	86.1	93.2	91.2	82.1	93.2	92.4	92.1	**95.7**
		order > p	30.3	14.5	23.5	13.8	6.7	8.7	17.8	6.7	7.5	7.8	4.0
1	I	order < p	81.8	95.8	88.5	94.0	98.4	97.0	93.7	98.6	98.3	98.3	99.8
		p = 5	**7.0**	2.5	5.2	3.5	0.9	2.0	3.6	0.9	1.0	1.0	0.2
		order > p	11.2	1.7	6.3	2.5	0.7	1.0	2.7	0.5	0.7	0.7	0.0
1	II	order < p	56.9	82.5	66.6	74.3	86.4	83.0	76.1	86.8	86.1	85.7	92.3
		p = 5	15.8	11.0	**15.9**	15.6	9.4	11.5	13.6	9.3	9.6	9.9	6.1
		order > p	27.3	6.5	17.5	10.1	4.2	5.5	10.3	3.9	4.3	4.4	1.6
1	III	order < p	4.1	9.8	5.5	7.0	12.3	9.6	8.3	12.7	11.7	11.4	20.3
		p = 5	63.4	78.5	71.1	77.5	80.2	**80.5**	76.0	79.9	80.4	80.4	76.2
		order > p	32.5	11.7	23.4	15.5	7.5	9.9	15.7	7.4	7.9	8.2	3.5
1	IV	order < p	5.5	14.1	7.5	11.4	20.0	16.3	10.6	20.4	18.9	18.5	31.7
		p = 5	64.8	72.9	70.1	**75.8**	72.6	74.8	72.6	72.4	73.1	73.3	64.4
		order > p	29.7	13.0	22.4	12.8	7.4	8.9	16.8	7.2	8.0	8.2	3.9
9	I	order < p	84.4	97.2	90.7	95.8	99.2	98.2	95.3	99.5	99.0	98.8	99.9
		p = 5	**6.0**	1.7	3.8	2.4	0.3	1.1	2.0	0.2	0.4	0.6	0.0
		order > p	9.6	1.1	5.5	1.8	0.5	0.7	2.7	0.3	0.6	0.6	0.1
9	II	order < p	63.4	86.1	73.1	79.6	89.1	85.6	81.9	89.9	88.8	88.4	94.3
		p = 5	**12.4**	8.9	11.5	12.1	8.4	10.3	10.5	7.8	8.3	8.5	5.0
		order > p	24.2	5.0	15.4	8.3	2.5	4.1	7.6	2.3	2.9	3.1	0.7
9	III	order < p	45.0	72.4	56.5	63.8	77.7	72.1	66.4	78.3	77.0	76.6	86.9
		p = 5	**26.4**	20.0	24.6	25.1	18.4	21.6	22.8	18.2	18.7	19.0	12.3
		order > p	28.6	7.6	18.9	11.1	3.9	6.3	10.8	3.5	4.3	4.4	0.8
9	IV	order < p	63.9	87.5	74.1	84.2	94.9	90.8	83.3	94.9	93.9	93.5	98.1
		p = 5	**20.4**	9.5	15.9	12.3	4.5	8.0	11.6	4.5	5.3	5.7	1.6
		order > p	15.7	3.0	10.0	3.5	0.6	1.2	5.1	0.6	0.8	0.8	0.3

Note: Boldface type indicates the maximum percentage of correct order being selected.

Table 5 Percentage of model selection for large sample size (n = 100) and the independent variables have a standard normal distribution

σ_0^2	Model	Order	AIC	BIC	HQIC	AICc	AICu	HQICc	KIC	$KICc_C$	$KICc_{SB}$	$KICc_{HM}$	NIC
0.25	I	order < p	0.8	4.0	2.2	1.0	2.6	2.6	2.0	2.6	2.6	2.6	3.8
		p = 5	78.6	**91.9**	88.4	82.0	90.4	90.8	88.6	90.4	90.1	90.1	**91.9**
		order > p	20.6	4.1	9.4	17.0	7.0	6.6	9.4	7.0	7.3	7.3	4.3
0.25	II	order < p	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		p = 5	73.1	**95.7**	87.3	78.0	89.1	89.9	86.6	89.1	88.8	88.7	94.4
		order > p	26.9	4.3	12.7	22.0	10.9	10.1	13.4	10.9	11.2	11.3	5.6
0.25	III	order < p	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		p = 5	74.0	**97.3**	89.1	78.8	91.2	91.7	88.2	91.2	90.8	90.7	96.3
		order > p	26.0	2.7	10.9	21.2	8.8	8.3	11.8	8.8	9.2	9.3	3.7
0.25	IV	order < p	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		p = 5	74.2	**93.8**	86.7	79.2	89.3	90.0	86.2	89.3	89.1	89.1	93.6
		order > p	25.8	6.2	13.3	20.8	10.7	10.0	13.8	10.7	10.9	10.9	6.4
1	I	order < p	25.1	60.0	40.5	29.9	43.8	44.9	39.8	43.8	43.3	43.1	57.5
		p = 5	**53.6**	37.1	48.3	52.6	46.9	46.7	48.6	46.9	47.0	47.1	39.1
		order > p	21.3	2.9	11.2	17.5	9.3	8.4	11.6	9.3	9.7	9.8	3.4
1	II	order < p	0.9	4.7	2.0	1.0	2.3	2.4	2.0	2.3	2.2	2.2	4.0
		p = 5	70.5	**91.2**	85.7	75.4	87.2	87.8	84.8	87.3	87.1	86.9	90.7
		order > p	28.6	4.1	12.3	23.6	10.5	9.8	13.2	10.4	10.7	10.9	5.3
1	III	order < p	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		p = 5	71.6	**95.6**	87.1	75.6	89.1	89.8	86.4	89.1	89.1	88.9	94.7
		order > p	28.4	4.4	12.9	24.4	10.9	10.2	13.6	10.9	10.9	11.1	5.3
1	IV	order < p	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		p = 5	77.1	**96.3**	89.2	82.1	91.0	91.9	88.6	91.0	90.8	90.8	95.4
		order > p	22.9	3.7	10.8	17.9	9.0	8.1	11.4	9.0	9.2	9.2	4.6
9	I	order < p	78.1	98.0	92.3	82.6	94.0	94.7	92.1	94.1	93.6	93.6	97.6
		p = 5	**11.5**	1.9	5.7	10.0	4.9	4.5	5.9	4.8	5.0	5.0	2.2
		order > p	10.4	0.1	2.0	7.4	1.1	0.8	2.0	1.1	1.4	1.4	0.2
9	II	order < p	45.7	78.8	61.4	48.4	63.4	65.0	59.8	63.6	63.2	63.1	76.4
		p = 5	35.2	19.5	31.5	**36.0**	30.3	29.4	32.4	30.2	30.4	30.3	21.5
		order > p	19.1	1.7	7.1	15.6	6.3	5.6	7.8	6.2	6.4	6.6	2.1
9	III	order < p	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		p = 5	73.1	**96.1**	87.5	77.6	89.2	90.2	87.0	89.3	89.2	89.0	95.1
		order > p	26.9	3.9	12.5	22.4	10.8	9.8	13.0	10.7	10.8	11.0	4.9
9	IV	order < p	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		p = 5	77.5	**95.9**	89.2	83.0	91.5	92.5	88.7	91.8	91.0	90.9	95.7
		order > p	22.5	4.1	10.8	17.0	8.5	7.5	11.3	8.2	9.0	9.1	4.3

Note: Boldface type indicates the maximum percentage of correct order being selected.

Table 6 Percentage of model selection for large sample size (n = 100) and the independent variables have a uniform distribution

σ_0^2	Model	Order	AIC	BIC	HQIC	AICc	AICu	HQICc	KIC	$KICc_C$	$KICc_{SB}$	$KICc_{HM}$	NIC
0.25	I	order < p	57.0	90.1	76.8	63.8	78.7	80.7	76.1	78.9	78.7	78.5	88.5
		p = 5	**27.5**	8.3	17.5	24.9	16.7	15.3	17.6	16.5	16.7	16.9	9.6
		order > p	15.5	1.6	5.7	11.3	4.6	4.0	6.3	4.6	4.6	4.6	1.9
0.25	II	order < p	14.9	43.4	28.4	16.4	29.6	30.4	27.8	29.6	29.4	29.2	39.9
		p = 5	58.2	52.9	61.5	61.4	61.1	61.0	**62.1**	61.1	61.2	61.3	55.6
		order > p	26.9	3.7	10.1	22.2	9.3	8.6	10.1	9.3	9.4	9.5	4.5
0.25	III	order < p	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		p = 5	70.9	**95.3**	86.1	75.1	88.3	89.1	85.5	88.4	88.1	94.3	94.3
		order > p	29.1	4.7	13.9	24.9	11.7	10.9	14.5	11.6	11.9	5.7	5.7
0.25	IV	order < p	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		p = 5	75.5	**94.7**	88.4	80.8	90.7	91.2	87.8	90.7	90.7	90.6	94.6
		order > p	24.5	5.3	11.6	19.2	9.3	8.8	12.2	9.3	9.3	9.4	5.4
1	I	order < p	81.7	98.8	95.1	86.5	95.8	96.4	94.5	95.9	95.7	95.7	98.3
		p = 5	**10.4**	1.0	3.4	8.8	3.1	2.7	3.8	3.0	3.2	3.2	1.3
		order > p	7.9	0.2	1.5	4.7	1.1	0.9	1.7	1.1	1.1	1.1	0.4
1	II	order < p	48.7	83.9	68.6	52.6	70.6	72.1	67.5	70.8	70.4	70.1	82.3
		p = 5	**30.9**	15.2	24.6	30.8	23.7	22.7	25.0	23.5	23.9	24.2	16.2
		order > p	20.4	0.9	6.8	16.6	5.7	5.2	7.5	5.7	5.7	5.7	1.5
1	III	order < p	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		p = 5	72.3	**96.7**	88.8	77.4	89.9	91.2	88.3	90.1	89.9	89.8	95.6
		order > p	27.7	3.3	11.2	22.6	10.1	8.8	11.7	9.9	10.1	10.2	4.4
1	IV	order < p	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		p = 5	75.7	**94.7**	87.8	81.1	89.3	90.4	87.6	89.6	89.3	89.3	93.9
		order > p	24.3	5.3	12.2	18.9	10.7	9.6	12.4	10.4	10.7	10.7	6.1
9	I	order < p	87.2	99.5	97.3	91.6	97.7	98.2	97.2	97.8	97.7	97.7	99.3
		p = 5	**5.2**	0.4	1.8	3.9	1.7	1.4	1.8	1.7	1.7	1.7	0.6
		order > p	7.6	0.1	0.9	4.5	0.6	0.4	1.0	0.5	0.6	0.6	0.1
9	II	order < p	67.9	95.5	86.2	72.6	87.6	88.3	85.4	87.7	87.6	87.5	94.2
		p = 5	**14.4**	3.9	8.6	13.7	8.4	8.1	9.0	8.4	8.4	8.5	4.9
		order > p	17.7	0.6	5.2	13.7	4.0	3.6	5.6	3.9	4.0	4.0	0.9
9	III	order < p	25.8	56.9	40.3	28.7	42.3	43.4	39.4	42.3	42.1	41.9	53.5
		p = 5	49.4	40.2	50.5	50.5	50.0	49.4	**50.7**	50.1	50.0	50.1	43.1
		order > p	24.8	2.9	9.2	20.8	7.7	7.2	9.9	7.6	7.9	8.0	3.4
9	IV	order < p	35.6	78.6	57.4	41.4	62.0	63.7	56.5	62.2	61.6	61.5	74.8
		p = 5	**46.4**	19.9	35.8	44.8	32.5	31.6	36.1	32.3	32.3	32.4	23.3
		order > p	18.0	1.5	6.8	13.8	5.5	4.7	7.4	5.5	6.1	6.1	1.9

Note: Boldface type indicates the maximum percentage of correct order being selected.

Table 7 Observed L_2 efficiency for small sample size (n = 15) and the independent variables have a standard normal distribution

σ_0^2	Model	Stat.	AIC	BIC	HQIC	AICc	AICu	HQICc	KIC	$KICc_C$	$KICc_{SB}$	$KICc_{HM}$	NIC
0.25	I	Ave. L_2 eff.	**0.7327**	0.7261	0.7319	0.6428	0.5750	0.6445	0.7209	0.5683	0.5845	0.5904	0.5020
		S.D. L_2 eff.	**0.2451**	0.2510	0.2454	0.2670	0.2771	0.2663	0.2552	0.2767	0.2770	0.2762	0.2758
0.25	II	Ave. L_2 eff.	0.6714	0.7373	0.6716	0.8377	0.7938	**0.8391**	0.7598	0.7948	0.7985	0.8006	0.7359
		S.D. L_2 eff.	0.3503	0.3405	0.3501	0.2900	0.3226	**0.2885**	0.3343	0.3225	0.3189	0.3171	0.3537
0.25	III	Ave. L_2 eff.	0.6968	0.7988	0.6966	0.9391	0.9601	0.9391	0.8252	0.9616	0.9573	0.9547	**0.9793**
		S.D. L_2 eff.	0.3445	0.3155	0.3444	0.1989	0.1653	0.1989	0.3031	0.1622	0.1707	0.1751	**0.1201**
0.25	IV	Ave. L_2 eff.	0.8326	0.8658	0.8323	0.9744	0.9867	0.9744	0.8782	0.9877	0.9823	0.9817	**0.9923**
		S.D. L_2 eff.	0.2423	0.2297	0.2422	0.1195	0.0889	0.1195	0.2241	0.0858	0.1006	0.1017	**0.0647**
1	I	Ave. L_2 eff.	0.6386	0.6487	0.6383	0.6738	0.6574	**0.6739**	0.6555	0.6567	0.6571	0.6578	0.6423
		S.D. L_2 eff.	0.2549	0.2534	0.2550	0.2405	0.2429	**0.2404**	0.2507	0.2430	0.2432	0.2427	0.2423
1	II	Ave. L_2 eff.	0.5975	0.6367	0.5972	0.6818	0.6613	**0.6827**	0.6464	0.6604	0.6625	0.6636	0.6382
		S.D. L_2 eff.	0.3396	0.3280	0.3397	0.3005	0.2901	0.3002	0.3237	0.2896	0.2907	0.2914	**0.2842**
1	III	Ave. L_2 eff.	0.6902	0.7890	0.6895	0.9451	0.9687	0.9438	0.8254	0.9720	0.9663	0.9642	**0.9776**
		S.D. L_2 eff.	0.3498	0.3253	0.3497	0.1895	0.1469	0.1914	0.3043	0.1406	0.1517	0.1557	**0.1296**
1	IV	Ave. L_2 eff.	0.8220	0.8579	0.8216	0.9709	**0.9758**	0.9702	0.8691	0.9752	0.9731	0.9728	0.9683
		S.D. L_2 eff.	0.2489	0.2429	0.2488	0.1340	**0.1308**	0.1357	0.2382	0.1334	0.1364	0.1358	0.1560
9	I	Ave. L_2 eff.	0.6205	0.7344	0.6196	0.8561	0.8947	0.8550	0.7681	0.8956	0.8941	0.8903	**0.9202**
		S.D. L_2 eff.	0.3484	0.3291	0.3486	0.2439	0.2047	0.2449	0.3131	0.2041	0.2052	0.2110	**0.1723**
9	II	Ave. L_2 eff.	0.6244	0.7256	0.6244	0.8579	0.8934	0.8579	0.7636	0.8947	0.8901	0.8860	**0.9209**
		S.D. L_2 eff.	0.3565	0.3349	0.3565	0.2475	0.2126	0.2475	0.3171	0.2108	0.2166	0.2219	**0.1744**
9	III	Ave. L_2 eff.	0.6766	0.7377	0.6761	**0.7889**	0.7331	0.7884	0.7408	0.7287	0.7370	0.7394	0.6688
		S.D. L_2 eff.	0.3338	0.3205	0.3336	**0.3056**	0.3295	**0.3056**	0.3219	0.3319	0.3270	0.3259	0.3440
9	IV	Ave. L_2 eff.	**0.7837**	0.7748	0.7832	0.6529	0.5445	0.6539	0.7664	0.5360	0.5602	0.5683	0.4416
		S.D. L_2 eff.	0.2451	0.2581	**0.2449**	0.3197	0.3215	0.3198	0.2678	0.3201	0.3237	0.3256	0.2919

Note: Boldface type indicates the maximum average and minimum standard deviation of observed L_2 efficiency.

Table 8 Observed L_2 efficiency for small sample size (n = 15) and the independent variables have a uniform distribution

σ_0^2	Model	Stat.	AIC	BIC	HQIC	AICc	AICu	HQICc	KIC	$KICc_C$	$KICc_{SB}$	$KICc_{HM}$	NIC
0.25	I	Ave. L_2 eff.	0.6061	0.6797	0.6043	0.7733	0.7945	0.7726	0.7012	0.7954	0.7913	0.7919	**0.8105**
		S.D. L_2 eff.	0.2988	0.2885	0.2994	0.2344	0.2207	0.2347	0.2793	0.2200	0.2234	0.2234	**0.2118**
0.25	II	Ave. L_2 eff.	0.5803	0.6608	0.5777	0.7658	0.7776	0.7645	0.6825	0.7788	0.7764	0.7768	**0.7965**
		S.D. L_2 eff.	0.3247	0.3059	0.3254	0.2494	0.2429	0.2502	0.2969	0.2424	0.2430	0.2431	**0.2372**
0.25	III	Ave. L_2 eff.	0.6819	0.7700	0.6807	0.9108	0.9273	0.9095	0.7897	**0.9284**	0.9271	0.9267	0.9121
		S.D. L_2 eff.	0.3523	0.3361	0.3523	0.2369	0.2198	0.2384	0.3282	**0.2187**	0.2204	0.2209	0.2447
0.25	IV	Ave. L_2 eff.	0.8261	0.8522	0.8252	**0.8997**	0.8405	**0.8997**	0.8571	0.8289	0.8597	0.8649	0.7056
		S.D. L_2 eff.	0.2421	0.2384	0.2421	0.2401	0.3039	0.2401	**0.2381**	0.3140	0.2869	0.2812	0.3864
1	I	Ave. L_2 eff.	0.6259	0.7256	0.6250	0.8746	0.9192	0.8728	0.7619	0.9230	0.9168	0.9137	**0.9350**
		S.D. L_2 eff.	0.3535	0.3377	0.3535	0.2340	0.1804	0.2365	0.3221	0.1753	0.1838	0.1892	**0.1521**
1	II	Ave. L_2 eff.	0.6157	0.7257	0.6133	0.8578	0.8982	0.8563	0.7677	0.8999	0.8982	0.8947	**0.9213**
		S.D. L_2 eff.	0.3670	0.3431	0.3674	0.2595	0.2193	0.2607	0.3254	0.2179	0.2193	0.2228	**0.1892**
1	III	Ave. L_2 eff.	0.6342	0.6798	0.6336	**0.7449**	0.6938	0.7447	0.6987	0.6886	0.6985	0.7016	0.6327
		S.D. L_2 eff.	0.3380	0.3370	0.3378	**0.3173**	0.3295	0.3176	0.3321	0.3301	0.3287	0.3286	0.3361
1	IV	Ave. L_2 eff.	**0.7448**	0.7324	0.7447	0.5999	0.5151	0.6015	0.7181	0.5073	0.5255	0.5321	0.4475
		S.D. L_2 eff.	0.2544	0.2707	**0.2542**	0.2953	0.2842	0.2957	0.2774	0.2830	0.2873	0.2890	0.2623
9	I	Ave. L_2 eff.	0.6363	0.7577	0.6358	0.9132	0.9538	0.9108	0.7928	0.9549	0.9517	0.9480	**0.9739**
		S.D. L_2 eff.	0.3936	0.3633	0.3934	0.2399	0.1808	0.2431	0.3479	0.1779	0.1834	0.1895	**0.1347**
9	II	Ave. L_2 eff.	0.6367	0.7729	0.6362	0.9025	0.9473	0.9016	0.8136	0.9490	0.9435	0.9406	**0.9678**
		S.D. L_2 eff.	0.3911	0.3553	0.3909	0.2575	0.1918	0.2589	0.3358	0.1883	0.1983	0.2044	**0.1504**
9	III	Ave. L_2 eff.	0.5957	0.6755	0.5953	0.7891	0.8111	0.7891	0.6996	0.8121	0.8098	0.8075	**0.8350**
		S.D. L_2 eff.	0.3344	0.3180	0.3342	0.2563	0.2457	0.2563	0.3112	0.2449	0.2467	0.2481	**0.2287**
9	IV	Ave. L_2 eff.	0.6344	0.6908	0.6319	0.7950	0.8190	0.7945	0.7120	0.8198	0.8162	0.8144	**0.8366**
		S.D. L_2 eff.	0.2852	0.2794	0.2847	0.2292	0.2130	0.2293	0.2722	0.2132	0.2153	0.2162	**0.2020**

Note: Boldface type indicates the maximum average and minimum standard deviation of observed L_2 efficiency.

Table 9 Observed L_2 efficiency for moderate sample size (n = 30) and the independent variables have a standard normal distribution

σ_0^2	Model	Stat.	AIC	BIC	HQIC	AICc	AICu	HQICc	KIC	$KICc_C$	$KICc_{SB}$	$KICc_{HM}$	NIC
0.25	I	Ave. L_2 eff.	0.8055	0.7880	0.8052	**0.8092**	0.7524	0.7817	0.8026	0.7485	0.7562	0.7567	0.6700
		S.D. L_2 eff.	**0.2495**	0.2668	0.2533	0.2566	0.2832	0.2716	0.2603	0.2842	0.2821	0.2820	0.3032
0.25	II	Ave. L_2 eff.	0.7728	0.8948	0.8252	0.8686	0.9266	0.9135	0.8663	0.9288	0.9245	0.9222	**0.9430**
		S.D. L_2 eff.	0.3329	0.2574	0.3083	0.2768	0.2193	0.2369	0.2801	0.2161	0.2222	0.2248	**0.1985**
0.25	III	Ave. L_2 eff.	0.7835	0.9141	0.8347	0.8928	0.9397	0.9258	0.8909	0.9407	0.9382	0.9373	**0.9708**
		S.D. L_2 eff.	0.3255	0.2342	0.3017	0.2548	0.2024	0.2198	0.2581	0.2013	0.2049	0.2068	**0.1465**
0.25	IV	Ave. L_2 eff.	0.8670	0.9368	0.8940	0.9421	0.9689	0.9644	0.9257	0.9701	0.9662	0.9662	**0.9815**
		S.D. L_2 eff.	0.2266	0.1759	0.2119	0.1690	0.1302	0.1370	0.1873	0.1276	0.1348	0.1348	**0.1037**
1	I	Ave. L_2 eff.	**0.6983**	0.6450	0.6824	0.6807	0.6234	0.6469	0.6628	0.6215	0.6268	0.6284	0.5741
		S.D. L_2 eff.	**0.2397**	0.2470	0.2430	0.2418	0.2456	0.2431	0.2475	0.2454	0.2457	0.2446	0.2456
1	II	Ave. L_2 eff.	0.6918	0.7176	0.7156	**0.7427**	0.7265	0.7391	0.7195	0.7248	0.7269	0.7307	0.6731
		S.D. L_2 eff.	0.3438	0.3315	0.3359	**0.3228**	0.3283	0.3232	0.3326	0.3294	0.3285	0.3277	0.3419
1	III	Ave. L_2 eff.	0.7556	0.8915	0.8216	0.8747	0.9312	0.9115	0.8698	0.9324	0.9272	0.9247	**0.9676**
		S.D. L_2 eff.	0.3419	0.2613	0.3117	0.2737	0.2131	0.2387	0.2792	0.2112	0.2194	0.2224	**0.1477**
1	IV	Ave. L_2 eff.	0.8580	0.9295	0.8894	0.9355	0.9643	0.9547	0.9147	0.9643	0.9603	0.9586	**0.9779**
		S.D. L_2 eff.	0.2366	0.1882	0.2189	0.1811	0.1407	0.1570	0.2018	0.1407	0.1481	0.1517	**0.1138**
9	I	Ave. L_2 eff.	0.6877	0.7895	0.7279	0.7663	0.8176	0.7957	0.7680	0.8183	0.8155	0.8103	**0.8375**
		S.D. L_2 eff.	0.2911	0.2445	0.2762	0.2524	0.2219	0.2362	0.2565	0.2215	0.2238	0.2269	**0.2082**
9	II	Ave. L_2 eff.	0.6693	0.7775	0.7210	0.7506	0.7995	0.7795	0.7555	0.7997	0.7937	0.7915	**0.8281**
		S.D. L_2 eff.	0.3217	0.2650	0.2999	0.2827	0.2478	0.2605	0.2813	0.2479	0.2526	0.2539	**0.2289**
9	III	Ave. L_2 eff.	0.7822	0.8942	0.8327	0.8837	0.9193	0.9106	0.8759	0.9180	0.9182	0.9174	**0.9201**
		S.D. L_2 eff.	0.3299	0.2518	0.3021	0.2592	**0.2224**	0.2340	0.2677	0.2238	0.2244	0.2255	0.2254
9	IV	Ave. L_2 eff.	0.8591	0.8820	0.8718	**0.8993**	0.8739	0.8891	0.8856	0.8726	0.8793	0.8832	0.8321
		S.D. L_2 eff.	0.2299	0.2348	0.2287	**0.2173**	0.2527	0.2345	0.2259	0.2537	0.2469	0.2428	0.2964

Note: Boldface type indicates the maximum average and minimum standard deviation of observed L_2 efficiency.

Table 10 Observed L_2 efficiency for moderate sample size (n = 30) and the independent variables have a uniform distribution

σ_0^2	Model	Stat.	AIC	BIC	HQIC	AICc	AICu	HQICc	KIC	$KICc_C$	$KICc_{SB}$	$KICc_{HM}$	NIC
0.25	I	Ave. L_2 eff.	0.6450	0.6676	0.6532	0.6718	0.6810	0.6798	0.6633	0.6811	0.6812	**0.6816**	0.6810
		S.D. L_2 eff.	0.2574	0.2369	0.2492	0.2403	0.2303	0.2331	0.2427	0.2298	0.2298	0.2291	**0.2240**
0.25	II	Ave. L_2 eff.	0.6518	0.6822	0.6722	**0.6989**	0.6900	0.6961	0.6867	0.6889	0.6899	0.6884	0.6778
		S.D. L_2 eff.	0.3192	0.2893	0.3080	0.2953	0.2795	0.2873	0.2953	0.2788	0.2803	0.2813	**0.2671**
0.25	III	Ave. L_2 eff.	0.7954	0.9277	0.8551	0.9098	0.9521	0.9378	0.9068	0.9532	0.9488	0.9471	**0.9748**
		S.D. L_2 eff.	0.3215	0.2184	0.2871	0.2373	0.1816	0.2032	0.2424	0.1800	0.1882	0.1906	**0.1373**
0.25	IV	Ave. L_2 eff.	0.8579	0.9247	0.8857	0.9290	0.9630	0.9530	0.9113	0.9630	0.9591	0.9570	**0.9750**
		S.D. L_2 eff.	0.2407	0.1970	0.2270	0.1919	0.1452	0.1619	0.2079	0.1452	0.1520	0.1565	**0.1226**
1	I	Ave. L_2 eff.	0.6769	0.8005	0.7328	0.7761	0.8208	0.8044	0.7812	0.8234	0.8202	0.8191	**0.8591**
		S.D. L_2 eff.	0.3122	0.2594	0.2965	0.2727	0.2420	0.2523	0.2715	0.2396	0.2425	0.2432	**0.2101**
1	II	Ave. L_2 eff.	0.6610	0.8053	0.7194	0.7686	0.8278	0.8118	0.7731	0.8309	0.8264	0.8241	**0.8647**
		S.D. L_2 eff.	0.3356	0.2685	0.3162	0.2890	0.2508	0.2607	0.2903	0.2481	0.2514	0.2526	**0.2182**
1	III	Ave. L_2 eff.	0.7681	0.8609	0.8164	0.8580	0.8726	**0.8750**	0.8453	0.8712	0.8739	0.8742	0.8404
		S.D. L_2 eff.	0.3294	0.2800	0.3078	0.2796	0.2695	**0.2673**	0.2914	0.2699	0.2682	0.2676	0.2984
1	IV	Ave. L_2 eff.	0.8410	0.8584	0.8574	**0.8792**	0.8436	0.8656	0.8649	0.8423	0.8480	0.8511	0.7711
		S.D. L_2 eff.	0.2419	0.2530	0.2409	**0.2336**	0.2739	0.2502	0.2418	0.2746	0.2701	0.2665	0.3289
9	I	Ave. L_2 eff.	0.7356	0.9004	0.8024	0.8509	0.9333	0.9060	0.8653	0.9361	0.9284	0.9273	**0.9561**
		S.D. L_2 eff.	0.3722	0.2540	0.3406	0.3029	0.2077	0.2455	0.2909	0.2021	0.2160	0.2183	**0.1685**
9	II	Ave. L_2 eff.	0.7187	0.8941	0.7921	0.8451	0.9170	0.8923	0.8623	0.9229	0.9143	0.9118	**0.9539**
		S.D. L_2 eff.	0.3796	0.2615	0.3474	0.3077	0.2328	0.2617	0.2939	0.2248	0.2370	0.2398	**0.1750**
9	III	Ave. L_2 eff.	0.6416	0.7048	0.6745	0.7040	0.7190	0.7150	0.6991	0.7204	0.7174	0.7173	**0.7319**
		S.D. L_2 eff.	0.3210	0.2706	0.3038	0.2824	0.2572	0.2676	0.2820	0.2550	0.2593	0.2599	**0.2429**
9	IV	Ave. L_2 eff.	0.6588	0.6727	0.6620	0.6718	0.6810	0.6800	0.6683	0.6810	0.6804	0.6803	**0.6874**
		S.D. L_2 eff.	0.2449	0.2304	0.2378	0.2311	0.2281	**0.2271**	0.2336	0.2282	0.2290	0.2290	0.2307

Note: Boldface type indicates the maximum average and minimum standard deviation of observed L_2 efficiency.

From Tables 1 to 10, the results of comparing the model selection criteria performances can be concluded as Table 11.

Table 11 The appropriate criteria under various circumstances

n	σ_0^2	Model	Independent Variable							
			Normal				Uniform			
			Criteria	%Max Correct	Max. Ave. L_2 eff.	Min. S.D. L_2 eff.	Criteria	%Max Correct	Max. Ave. L_2 eff.	Min. S.D. L_2 eff.
15	0.25	I	AIC	30.2	AIC	AIC	HQIC	7.8	NIC	NIC
		II	HQICc	73.2	HQICc	HQICc	BIC	18.6	NIC	NIC
		III	NIC	97.0	NIC	NIC	$KICc_C$	89.6	$KICc_C$	$KICc_C$
		IV	NIC	98.5	NIC	NIC	AICc, HQICc	83.9	AICc, HQICc	KIC
	1	I	HQIC	14.6	HQICc	HQICc	AIC, HQIC	9.1	NIC	NIC
		II	HQICc	31.4	HQICc	NIC	BIC	13.9	NIC	NIC
		III	NIC	96.8	NIC	NIC	AICc, HQICc	53.5	AICc	AICc
		IV	AICu, $KICc_C$	96.4	AICu	AICu	BIC	41.5	AIC	HQIC
	9	I	AIC, HQIC	7.6	NIC	NIC	AIC, HQIC	7.9	NIC	NIC
		II	AIC, HQIC	13.0	NIC	NIC	AIC, HQIC	13.3	NIC	NIC
		III	AICc	63.3	AICc	AICc, HQIC	AIC, HQIC	14.9	NIC	NIC
		IV	BIC	48.1	AIC	HQIC	HQIC	11.8	NIC	NIC
30	0.25	I	AICc	55.4	AICc	AIC	AIC	11.5	$KICc_{HM}$	NIC
		II	NIC	91.9	NIC	NIC	AICc	31.9	AICc	NIC
		III	NIC	95.9	NIC	NIC	NIC	96.5	NIC	NIC
		IV	NIC	96.7	NIC	NIC	NIC	95.7	NIC	NIC
	1	I	AIC	22.4	AIC	AIC	AIC	7.0	NIC	NIC
		II	AICc	54.6	AICc	AICc	HQIC	15.9	NIC	NIC
		III	NIC	95.1	NIC	NIC	HQICc	80.5	HQICc	HQICc
		IV	NIC	96.0	NIC	NIC	AICc	75.8	AICc	AICc
	9	I	AIC	7.0	NIC	NIC	AIC	6.0	NIC	NIC
		II	AIC	17.1	NIC	NIC	AIC	12.4	NIC	NIC
		III	NIC	88.3	NIC	AICu	AIC	26.4	NIC	NIC
		IV	AICc	80.2	AICc	AICc	AIC	20.4	NIC	HQICc

Note: - means not considering the observed L_2 efficiency for the large sample size.

Table 11 (Continued)

n	σ_0^2	Model	Independent Variable							
			Normal				Uniform			
			Criteria	%Max Correct	Max. Ave. L_2 eff.	Min. S.D. L_2 eff.	Criteria	%Max Correct	Max. Ave. L_2 eff.	Min. S.D. L_2 eff.
100	0.25	I	BIC, NIC	91.9	-	-	AIC	27.5	-	-
		II	BIC	95.7	-	-	KIC	62.1	-	-
		III	BIC	97.3	-	-	BIC	95.3	-	-
		IV	BIC	93.8	-	-	BIC	94.7	-	-
	1	I	AIC	53.6	-	-	AIC	10.4	-	-
		II	BIC	91.2	-	-	AIC	30.9	-	-
		III	BIC	95.6	-	-	BIC	96.7	-	-
		IV	BIC	96.3	-	-	BIC	94.7	-	-
	9	I	AIC	11.5	-	-	AIC	5.2	-	-
		II	AICc	36.0	-	-	AIC	14.4	-	-
		III	BIC	96.1	-	-	KIC	50.7	-	-
		IV	BIC	95.9	-	-	AIC	46.4	-	-

Note: - means not considering the observed L_2 efficiency for the large sample size.

From Table 11 we can conclude that,

(1) The true model is very weakly identifiable as Model I for n = 15, 30, in most cases, AIC and HQIC select the most accurate models. These criteria can identify the true model about 8–55% of the time for σ_0^2 = 0.25, about 7–22% of the time for σ_0^2 = 1, and about 6–8% of the time for σ_0^2 = 9. For n = 100, in most cases, AIC select the most accurate models. This criterion can identify the true model about 28–92% of the time for σ_0^2 = 0.25, about 10–54% of the time for σ_0^2 = 1, and about 5–12% of the time for σ_0^2 = 9.

(2) The true model is weakly identifiable as Model II for n = 15, 30, in most cases, AIC and HQIC select the most accurate models. These criteria can identify the true model about 19–92% of the time for σ_0^2 = 0.25, about 14–55% of the time for σ_0^2 = 1, and about 12–17% of the time for σ_0^2 = 9. For n = 100, in most cases, AIC and BIC select the most accurate models. These criteria can identify the true model about 62–96% of the time for σ_0^2 = 0.25, about 31–91% of the time for σ_0^2 = 1, and about 14–36% of the time for σ_0^2 = 9.

(3) The true model is very strongly identifiable as Model III for n = 15, 30, in most cases, NIC select the most accurate models. This criterion can identify the true model about 90–97% of the time for σ_0^2 = 0.25, about 54–97% of the time for σ_0^2 = 1, and about 15–88% of the time for σ_0^2 = 9. For n = 100, in most cases, BIC select the most accurate models. This criterion can identify the true model about 95–97% of the time for σ_0^2 = 0.25, about 96–97% of the time for σ_0^2 = 1, and about 51–96% of the time for σ_0^2 = 9.

(4) The true model is strongly identifiable as Model IV for n = 15, 30, in most cases, NIC select the most accurate models. This criterion can identify the true model about 84–99% of the time for σ_0^2 = 0.25, about 42–96% of the time for σ_0^2 = 1, and about 12–80% of the time for σ_0^2 = 9. For n =

100, in most cases, BIC select the most accurate models. This criterion can identify the true model about 94–95% of the time for σ_0^2 = 0.25, about 95–96% of the time for σ_0^2 = 1, and about 6 – 96% of the time for σ_0^2 = 9.

(5) In most cases, the observed L_2 efficiency suggests that NIC is the best criterion for small to moderate sample sizes.

(6) When the sample size increases or the model is strongly identifiable, it is more likely to select the correct order and the observed L_2 efficiency is also increased. In addition, the error variance affects the correction rate of order being selected and the efficiency of the observed L_2.

(7) When the independent variables have a uniform distribution, they present a lower percentage of correct order being selected than the normal distribution.

4. Conclusions

In this paper, we propose the new criteria for regression model selection, called New Information Criterion (NIC). The performances of NIC are examined by the extensive simulation study against AIC, BIC, HQIC, AICc, AICu, HQICc, KIC, $KICc_C$, $KICc_{SB}$, and $KICc_{HM}$, under the difference various circumstances: sample sizes (n), regression coefficients (β), variances of error term (σ^2), and distribution of independent variables. For 1,000 samples of simulation, the results of comparing the model selection criteria performances can be concluded as follows. When the sample sizes are small to moderate and the true model is somewhat difficult to identify, the performances of AIC and HQIC perform the best. However, they can identify the true model actually less accurate about 6–55%. As a result, we used the observed L_2 efficiency to assess model selection criteria performances. In most cases, the observed L_2 efficiency suggests that NIC is the best criterion for small to moderate sample sizes. For the large sample size and the true model is somewhat difficult to identify, AIC and BIC select the most accurate models. When the sample sizes are small to moderate and the true model can be specified more easily, the performances of NIC perform the best. For the large sample size and the true model can be specified more easily, BIC select the most accurate models. When the sample size increases or the model is strongly identifiable, it is more likely to select the correct order and the observed L_2 efficiency is also increased. In addition, the error variance affects the correction rate of order being selected and the efficiency of the observed L_2. When the independent variables have a uniform distribution, they present a lower percentage of correct order being selected than the normal distribution.

Acknowledgements

This work was supported by Thailand Science Research and Inno Office of National Higher Education Science Research and Innovation Policy Council, Thaksin University (research project grant) Fiscal Year 2021.

References

Akaike H. A new look at the statistical model identification. IEEE Trans Automat Contr. 1974; 19(6): 716-723.

Bedrick EJ, Tsai CL. Model selection for multivariate regression in small samples. Biometrics. 1994; 50(1): 226-231.

Cavanaugh JE. A large-sample model selection criterion based on Kullback's symmetric divergence. Stat Prob Lett. 1999; 42(4): 333-343.

Cavanaugh JE. Criteria for linear model selection based on Kullback's symmetric divergence. Aust N Z J Stat. 2004; 46(2): 257-274.

Hafidi B, Mkhadri A. A corrected Akaike criterion based on Kullback's symmetric divergence: applications in time series, multiple and multivariate regression. Comput Stat Data Anal. 2006; 50(6): 1524-1550.

Hannan EJ, Quinn BG. The determination of the order of an autoregression. J Roy Stat Soc B Met. 1979; 41(2): 190-195.

Hurvich CM, Tsai CL. Regression and time series model selection in small samples. Biometrika. 1989; 76(2): 297-307.

Hurvich CM, Tsai CL. A Corrected Akaike information criterion for vector autoregressive model selection. J Time Anal. 1993; 14(3): 271-279.

Hurvich CM, Shumway RH, Tsai CL. Improved estimators of Kullback-Leibler information for autoregressive model selection in small samples. Biometrika. 1990; 77(4): 709-719.

Keerativibool W. Study on the penalty functions of model selection criteria. Thail Stat. 2014; 12(2): 161-178.

Keerativibool W, Siripanich P. Comparison of the model selection criteria for multiple regression based on Kullback-Leibler's information. Chiang Mai J Sci. 2017; 44(2): 699-714.

McQuarrie ADR, Tsai CL. Regression and time series model selection. Singapore: World Scientific Publishing Company; 1998.

McQuarrie ADR, Shumway RH, Tsai CL. The model selection criterion AICu. Stat Prob Lett. 1997; 34(3): 285-292.

Neath AA, Cavanaugh JE. Regression and time series model selection using variants of the Schwarz information criterion. Commun Stat Theory Methods. 1997; 26(3): 559-580.

Schwarz G. Estimating the dimension of a model. Ann Stat. 1978; 6(2): 461-464.

Seghouane AK, Bekara M. A small sample model selection criterion based on Kullback's symmetric divergence. IEEE Trans Signal Process. 2004; 52(12): 3314-3323.