



การเปรียบเทียบประสิทธิภาพของเกณฑ์การคัดเลือกตัวแบบสำหรับตัวแบบการถดถอย

A Comparison of the Effectiveness of Model Selection Criteria for Regression Model

วรารังคณา เรียงสุทธิ^{*}

Warangkhan Riansut^{*}

สาขาวิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยทักษิณ

Department of Mathematics and Statistics, Faculty of Science, Thaksin University

Received : 11 March 2021

Revised : 29 April 2021

Accepted : 6 May 2021

บทคัดย่อ

วัตถุประสงค์ของการศึกษาค้นคว้าครั้งนี้คือ การเปรียบเทียบประสิทธิภาพของเกณฑ์การคัดเลือกตัวแบบสำหรับตัวแบบการถดถอย 10 เกณฑ์ ได้แก่ เกณฑ์ AIC, BIC, HQIC, AICc, AICu, HQICc, KIC, KICc_C, KICc_{SB} และ KICc_{HM} ภายใต้การจำลองข้อมูลที่มีความแตกต่างของขนาดตัวอย่าง จำนวนพารามิเตอร์ในตัวแบบ สัมประสิทธิ์การถดถอย ความแปรปรวนของความคลาดเคลื่อน และการแจกแจงของตัวแปรอิสระ ผลการศึกษาพบว่า กรณีตัวอย่างขนาดเล็ก ตัวแบบระบุยากถึงยากมาก เกณฑ์ที่เหมาะสม คือ AIC, HQIC กรณีตัวอย่างขนาดกลาง ตัวแบบระบุยากถึงยากมาก เกณฑ์ที่เหมาะสม คือ AIC และถ้าตัวแบบระบุง่ายถึงง่ายมากทั้งกรณีตัวอย่างขนาดเล็กและขนาดกลาง เกณฑ์ที่เหมาะสม คือ AICu และ KICc สำหรับกรณีตัวอย่างขนาดใหญ่ ตัวแบบระบุยากถึงยากมาก เกณฑ์ที่เหมาะสม คือ AIC, BIC และถ้าตัวแบบระบุง่ายถึงง่ายมาก เกณฑ์ที่เหมาะสม คือ BIC

คำสำคัญ : เกณฑ์การคัดเลือกตัวแบบ ; ตัวแบบการถดถอย ; ประสิทธิภาพ

Abstract

The aim of this study is to compare the effectiveness of the ten model selection criteria for regression model, 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 variable. The results of the study showed that a small sample case and the true model is somewhat difficult to identify, the appropriate criteria are AIC, HQIC. Medium sample cases and the true model is somewhat difficult to identify, the appropriate criterion is AIC. If the true model can be specified more easily for both small and medium samples, the appropriate criteria are AICu, KICc. For the large sample cases, and the true model is somewhat difficult to identify, the appropriate criteria are AIC, BIC. If the true model can be specified more easily, the appropriate criterion is BIC.

Keywords : model selection criterion, regression model, efficiency

*Corresponding author. E-mail : warang27@gmail.com



บทนำ

ปัญหาสำคัญของการคัดเลือกตัวแบบการถดถอยที่เหมาะสมจากตัวแบบที่เป็นไปได้ทั้งหมด เพื่อใช้ในการพยากรณ์ ข้อมูลในขนาดต้นฉบับเป็นปัญหาหนึ่งที่สำคัญ ในโลกความเป็นจริงเราไม่สามารถทราบได้เลยว่าตัวแบบที่แท้จริง (True Model) คือตัวแบบใด แต่เราคาดหวังว่าจะสามารถหาตัวแบบที่มาเป็นตัวแทนที่ถูกต้องและเหมาะสมได้ เกณฑ์การคัดเลือกตัวแบบ เป็นเครื่องมือหนึ่งที่มีประโยชน์ในการพิจารณาความเหมาะสมของตัวแบบ โดยการประเมินว่าตัวแบบใดที่ให้ความสมมูลที่ดี ที่สุด กล่าวคือ สามารถคัดเลือกได้ว่าตัวแบบใดที่มีจำนวนตัวแปรในสมการที่เพียงพอ ถูกต้อง และเหมาะสม ไม่มากและไม่ น้อยเกินไป (Goodness of Fit และ Parsimony) เกณฑ์การคัดเลือกตัวแบบเกณฑ์แรกที่เป็นที่รู้จัก คือ เกณฑ์สารสนเทศ ของอะกะอิเกะ (Akaike's Information Criterion: AIC) (Akaike, 1974) เกณฑ์นี้เป็นตัวประมาณไม่เอนเอียงเชิงเส้นกำกับ (Asymptotically Unbiased Estimator) ของความผันแปรของความแตกต่างแบบทางตรงของคูลแบค (Kullback's Directed Divergence) ซึ่งความแตกต่างแบบทางตรงเป็นที่รู้จักกันในนามของสารสนเทศของคูลแบคและลีเบลล (Kullback-Leibler Information) เพื่อประเมินความแตกต่างระหว่างตัวแบบทางสถิติ 2 ตัวแบบ คือ ตัวแบบที่แท้จริงและตัวแบบโดยประมาณ (Fitted Model) เกณฑ์การคัดเลือกตัวแบบเกณฑ์อื่น ๆ ที่เป็นที่รู้จักต่อมา ได้แก่ เกณฑ์สารสนเทศของเบส์ (Bayesian Information Criterion: BIC) (Schwarz, 1978) เกณฑ์สารสนเทศของ Hannan และ Quinn (Hannan and Quinn Information Criterion: HQIC) (Hannan and Quinn, 1979) เกณฑ์การปรับแก้ของ AIC (AICc) (Hurvich and Tsai, 1989) เกณฑ์การปรับแก้ของ AICc (AICu) (McQuarrie *et al.*, 1997) เกณฑ์การปรับแก้ของ HQIC (HQICc) (McQuarrie and Tsai, 1998) เกณฑ์สารสนเทศของคูลแบค (Kullback Information Criterion: KIC) (Cavanaugh, 1999) เกณฑ์การปรับแก้ ของ KIC โดย Cavanaugh (KICc_c) (Cavanaugh, 2004) เกณฑ์การปรับแก้ของ KIC โดย Seghouane และ Bekara (KICc_{SB}) (Seghouane and Bekara, 2004) และเกณฑ์การปรับแก้ของ KIC โดย Hafidi และ Mkhadri (KICc_{HM}) (Hafidi and Mkhadri, 2006) ถึงแม้ว่าเกณฑ์ AIC ยังคงเป็นที่นิยมใช้กันมากที่สุดในการคัดเลือกตัวแบบ แต่เกณฑ์ BIC, HQIC, AICc, AICu, HQICc, KIC, KICc_c, KICc_{SB} และ KICc_{HM} นับเป็นคู่แข่งที่ได้รับความนิยม เช่น เกณฑ์ BIC มีแนวโน้มที่จะคัดเลือก ตัวแบบที่มีจำนวนตัวแปรน้อยกว่าเกณฑ์ AIC (Neath and Cavanaugh, 1997) เกณฑ์ KIC เป็นเกณฑ์ในลักษณะสมมาตร ซึ่งรวมสารสนเทศระหว่างความแตกต่างแบบทางตรงทั้ง 2 รูปแบบหรือความแตกต่างแบบสมมาตรของคูลแบค (Kullback's Symmetric Divergence) ดังนั้นเกณฑ์ KIC จึงเป็นค่าวัดความแตกต่างของตัวแบบที่มีความละเอียดอ่อนกว่าเกณฑ์ AIC ซึ่งวัดเพียงองค์ประกอบเฉพาะแต่ละส่วนเท่านั้น (Cavanaugh, 1999; Cavanaugh, 2004) สำหรับในกรณีตัวอย่างขนาดเล็ก เกณฑ์ KICc มีแนวโน้มที่จะรับมือกับปัญหาของการประมาณค่าความแตกต่างแบบสมมาตรของคูลแบคที่ต่ำเกินไป (Cavanaugh, 2004; Seghouane and Bekara, 2004; Hafidi and Mkhadri, 2006) อย่างไรก็ตาม ทั้ง 10 เกณฑ์การคัดเลือก ตัวแบบที่กล่าวมายังมีข้อดีและข้อเสียในแต่ละประเด็นที่แตกต่างกันไป ดังจะเห็นได้จากวรรณกรรมที่ผ่านมา เช่น Keerativibool (2014a) ได้ศึกษาและเปรียบเทียบเกณฑ์การคัดเลือกตัวแบบบนพื้นฐานของความแตกต่างแบบสมมาตรของ คูลแบค ได้แก่ เกณฑ์ KIC, KICc_c, KICc_{SB} และ KICc_{HM} ผลการศึกษาพบว่า เกณฑ์ KICc_c ให้ค่าที่ใกล้เคียงกับค่าคาดหวัง ของตัวประมาณค่าวัดที่สมมาตร (Expected Estimated Symmetric Measure) มากที่สุด และมีค่าของฟังก์ชันการลงโทษ (Penalty Function) มากที่สุด เมื่อเงื่อนไข $(1-p/n)\exp(p/n) < 1$ เป็นจริง โดยเกณฑ์ที่มีค่าของฟังก์ชันการลงโทษรองลงมา คือ เกณฑ์ KICc_{SB}, KICc_{HM} และ KIC ตามลำดับ ซึ่งจากการมีค่าของฟังก์ชันการลงโทษมาก ส่งผลให้เกณฑ์ KICc_c

มีประสิทธิภาพสูงที่สุดถึงแม้ว่าเกณฑ์นี้อาจจะคัดเลือกได้ตัวแบบที่มีจำนวนตัวแปรน้อยเกินไป Keerativibool (2014b) ได้ศึกษาฟังก์ชันการลงโทษของเกณฑ์การคัดเลือกตัวแบบที่เป็นที่รู้จัก 3 เกณฑ์ ได้แก่ เกณฑ์ AIC, BIC และ KIC ซึ่งทั้ง 3 เกณฑ์นี้สามารถเขียนให้อยู่ในรูปเกณฑ์ใหม่ คือ $APIC\alpha = \log(\hat{\sigma}^2) + \alpha(p+1)/n$ เรียกชื่อว่า Adjusted Penalty Information Criterion ผลการศึกษาเชิงทฤษฎีพบว่า ถ้าค่าของ α มีค่าเข้าสู่อนันต์จะทำให้ความน่าจะเป็นที่จะมีจำนวนตัวแปรมากเกินไป ลู่เข้าสู่ศูนย์และอัตราส่วนสัญญาณต่อการรบกวน (Signal-to-Noise Ratio) มีแนวโน้มที่จะมีค่าสูง ผลการศึกษาจากการจำลองข้อมูลพบว่า เมื่อตัวแบบที่แท้จริงระบุได้ยาก ค่าของ α ที่ต่ำจะทำให้ได้ความน่าจะเป็นสูงที่จะคัดเลือกตัวแบบได้ถูกต้อง อย่างไรก็ตาม เมื่อตัวแบบที่แท้จริงระบุได้ยาก เกณฑ์ประสิทธิภาพ L_2 เป็นเกณฑ์ที่ดีกว่าเกณฑ์ความน่าจะเป็นที่จะคัดเลือกตัวแบบได้ถูกต้อง โดยเกณฑ์ประสิทธิภาพ L_2 แนะนำว่าค่าของ α ที่สูง ส่งผลให้เกณฑ์ APIC α มีประสิทธิภาพสูง ยกเว้นเมื่อสามารถระบุตัวแบบที่แท้จริงได้ไม่ยากมากนักและตัวอย่างมีขนาดปานกลางถึงขนาดใหญ่ ($n = 30 - 100$) ควรใช้ค่าของ α ที่ต่ำ สำหรับกรณีตัวแบบที่แท้จริงที่สามารถระบุได้อย่างง่าย ๆ ควรใช้ค่าของ α ที่สูง แต่ถ้าสัมประสิทธิ์การถดถอยมีค่าไม่มากเพียงพอและตัวอย่างมีขนาดเล็กถึงขนาดปานกลาง ($n = 15 - 30$) ค่าของ α ควรจะอยู่ในระดับปานกลาง Keerativibool and Siripanich (2017) ได้ศึกษาและเปรียบเทียบเกณฑ์การคัดเลือกตัวแบบบนพื้นฐานของสารสนเทศของ คูลแบคและลีเบลอ์ ได้แก่ เกณฑ์ AIC, AICc, KIC, KICc, KICc_{SB} และ KICc_{HM} ซึ่งเป็นการศึกษาที่ได้ผลต่อเนื่องจากการศึกษาของ Keerativibool (2014a) คือ เกณฑ์ KICc มีค่าของฟังก์ชันการลงโทษมากกว่าเกณฑ์ KICc_{SB}, KICc_{HM}, KIC และ AIC ตามลำดับ ขณะที่เกณฑ์ KIC จะมีค่าของฟังก์ชันการลงโทษมากกว่าเกณฑ์ AICc เมื่อมีเงื่อนไข $n - p > 2$, $n > 3$ และ p อยู่ในเซต $[-1, n/3 - 2]$ เป็นจริง แต่เกณฑ์ AICc จะมีค่าของฟังก์ชันการลงโทษมากกว่าเกณฑ์ AIC เสมอ ผลการศึกษาจากการจำลองข้อมูลพบว่า เกณฑ์การคัดเลือกตัวแบบที่มีค่าของฟังก์ชันการลงโทษมากจะทำให้ได้ตัวแบบที่มีจำนวนตัวแปรน้อยเกินไปและมีค่าของเกณฑ์การคัดเลือกตัวแบบลู่เข้าสู่ค่าจริงช้า ขณะที่เกณฑ์การคัดเลือกตัวแบบที่มีค่าของฟังก์ชันการลงโทษน้อยจะทำให้ได้ตัวแบบที่มีจำนวนตัวแปรมากเกินไปและไม่คงเส้นคงวา เกณฑ์ประสิทธิภาพ L_2 แนะนำว่า กรณีตัวแบบที่แท้จริงระบุได้ยาก ไม่ว่าจะตัวอย่างจะมีขนาดเล็กหรือขนาดใหญ่ เกณฑ์ KICc เป็นเกณฑ์ที่ดีที่สุด สำหรับกรณีตัวแบบที่แท้จริงระบุได้ง่าย ตัวอย่างขนาดเล็ก ($n = 15$) และความแปรปรวนของความคลาดเคลื่อนมีค่าต่ำ ทุกเกณฑ์การคัดเลือกตัวแบบมีความสามารถในการคัดเลือกตัวแบบได้ถูกต้อง แต่ถ้าความแปรปรวนของความคลาดเคลื่อนมีค่าเพิ่มขึ้น ประสิทธิภาพของทุกเกณฑ์การคัดเลือกตัวแบบจะลดลง เมื่อตัวอย่างมีขนาดกลางถึงขนาดใหญ่ ($n = 30 - 100$) และความแปรปรวนของความคลาดเคลื่อนมีค่าต่ำ เกณฑ์ KICc เป็นเกณฑ์ที่ดีที่สุด เนื่องจากสามารถระบุตัวแบบได้ถูกต้องจำนวนมาก แต่ถ้าความแปรปรวนของความคลาดเคลื่อนมีค่าเพิ่มขึ้น และขนาดตัวอย่างไม่มากเพียงพอ ทุกเกณฑ์การคัดเลือกตัวแบบสามารถคัดเลือกตัวแบบได้ถูกต้องน้อยลง และ Sangthong (2019) ได้ศึกษาประสิทธิภาพของเกณฑ์การคัดเลือกตัวแบบการวิเคราะห์พหุระดับ 2 วิธี คือ เกณฑ์ AIC และ BIC ผลการศึกษาพบว่า เมื่อจำนวนกลุ่มมีขนาดเล็กประสิทธิภาพของเกณฑ์การคัดเลือกตัวแบบค่อนข้างต่ำ และเมื่อจำนวนกลุ่มมากขึ้น ประสิทธิภาพการคัดเลือกตัวแบบจะสูงขึ้น ในกรณีจำนวนกลุ่มเท่ากับ 100 พบว่า โดยส่วนใหญ่ประสิทธิภาพการคัดเลือกตัวแบบด้วยเกณฑ์ BIC ดีกว่าเกณฑ์ AIC เมื่อประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุดแบบจำกัด (Restricted Maximum Likelihood: RML)

จากการศึกษาที่ผ่านมา พบว่า ยังไม่เคยมีการศึกษาเปรียบเทียบประสิทธิภาพของเกณฑ์การคัดเลือกตัวแบบสำหรับตัวแบบการถดถอยทั้ง 10 เกณฑ์ ได้แก่ เกณฑ์ AIC, BIC, HQIC, AICc, AICu, HQICc, KIC, KICc, KICc_{SB} และ KICc_{HM}



ดังนั้นการศึกษาคำนี้จึงต้องการเปรียบเทียบประสิทธิภาพของเกณฑ์ทั้งหมดดังกล่าว เพื่อสรุปว่าควรใช้เกณฑ์ใดภายใต้สถานการณ์ใด สำหรับการคัดเลือกตัวแบบการถดถอยที่ถูกต้อง ซึ่งจะส่งผลดีต่อค่าพยากรณ์ในอนาคตต่อไป โดยการศึกษาทั้งหมดจะกระทำภายใต้การจำลองข้อมูลในสถานการณ์ต่าง ๆ ได้แก่ ความแตกต่างในขนาดตัวอย่าง จำนวนตัวแปรอิสระในสมการ สัมประสิทธิ์การถดถอย ความแปรปรวนของความคลาดเคลื่อน และการแจกแจงของตัวแปรอิสระ เกณฑ์ที่จะถูกจัดให้เป็นเกณฑ์ที่ดีที่สุดเมื่อมีความน่าจะเป็นต่ำที่จะมีจำนวนตัวแปรในสมการมากหรือน้อยเกินไปและมีความน่าจะเป็นสูงที่สุดในการเลือกคัดเลือกรูปแบบได้ถูกต้อง

วิธีดำเนินการวิจัย

การศึกษาคำนี้จะศึกษาเฉพาะตัวแบบการถดถอย โดยมีตัวแบบที่แท้จริงหรือตัวแบบที่เกิดจากการจำลองข้อมูลในรูปแบบสมการที่ (1) ดังนี้ (Montgomery *et al.*, 2006)

$$y = X_0\beta_0 + \epsilon_0 \quad (1)$$

และตัวแบบโดยประมาณหรือตัวแบบที่เกิดจากการประมาณค่าพารามิเตอร์ด้วยวิธีการวิเคราะห์การถดถอยในรูปแบบสมการที่ (2) ดังนี้

$$y = X\beta + \epsilon \quad (2)$$

โดยที่ y แทนเวกเตอร์เชิงสุมขนาด $n \times 1$ X_0 และ X แทนเมตริกซ์ของตัวแปรอิสระขนาด $n \times p_0$ และ $n \times p$ ตามลำดับ β_0 และ β แทนเวกเตอร์ของสัมประสิทธิ์การถดถอยขนาด $p_0 \times 1$ และ $p \times 1$ ตามลำดับ ϵ_0 และ ϵ แทนเวกเตอร์ของความคลาดเคลื่อนที่มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ $\sigma_0^2 I_n$ และ $\sigma^2 I_n$ ตามลำดับ n แทนขนาดตัวอย่าง p_0 และ p แทนจำนวนพารามิเตอร์ในตัวแบบที่แท้จริงและตัวแบบโดยประมาณ ตามลำดับ โดยรวมค่าคงที่ด้วย ตัวประมาณภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimator) ของ β และ σ^2 แสดงดังสมการที่ (3) และ (4) ตามลำดับ

$$\hat{\beta} = (X'X)^{-1} X'y \quad \text{และ} \quad \hat{\sigma}^2 = \frac{1}{n} (y - X\hat{\beta})' (y - X\hat{\beta}) = \frac{SSE}{n} \quad (3)$$

$$s^2 = \frac{1}{n-p} (y - X\hat{\beta})' (y - X\hat{\beta}) = \frac{SSE}{n-p} = \frac{n\hat{\sigma}^2}{n-p} \quad (4)$$

โดยที่ SSE แทนผลบวกกำลังสองของความคลาดเคลื่อน สำหรับขั้นตอนการดำเนินการวิจัย แสดงดังนี้

1. จำลองข้อมูลภายใต้สถานการณ์ต่าง ๆ จำนวน 1,000 ชุด ได้แก่ ข้อมูลที่มีความแตกต่างของขนาดตัวอย่าง จำนวนพารามิเตอร์ในตัวแบบ สัมประสิทธิ์การถดถอย ความแปรปรวนของความคลาดเคลื่อน และการแจกแจงของตัวแปรอิสระ ดังนี้



• ขนาดตัวอย่าง 3 ระดับ ได้แก่ ตัวอย่างขนาดเล็ก ($n = 15$) ตัวอย่างขนาดกลาง ($n = 30$) และตัวอย่างขนาดใหญ่ ($n = 100$)

• จำนวนพารามิเตอร์ในตัวแบบ 2 ระดับ ได้แก่ $p_0 = 3$ และ $p_0 = 5$

• สัมประสิทธิ์การถดถอย 4 ตัวแบบ ได้แก่ ตัวแบบที่ 1: $\beta_0 = 1, \beta_1 = 0.5, \beta_2 = 0.4, \beta_3 = 0.3, \beta_4 = 0.2$ ตัวแบบที่ 2: $\beta_0 = 1, \beta_1 = 0.5, \beta_2 = 0.4$ ตัวแบบที่ 3: $\beta_0 = 1, \beta_1 = \beta_2 = 2$ และตัวแบบที่ 4: $\beta_0 = 1, \beta_1 = \beta_2 = \beta_3 = \beta_4 = 2$

• ความคลาดเคลื่อนมีการแจกแจงปกติ มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวน 3 ระดับ ได้แก่ ความแปรปรวนต่ำ ($\sigma_0^2 = 0.25$) ความแปรปรวนปานกลาง ($\sigma_0^2 = 1$) และความแปรปรวนสูง ($\sigma_0^2 = 9$)

• การแจกแจงของตัวแปรอิสระทั้งหมด 6 ตัว จำนวน 2 การแจกแจง ได้แก่ การแจกแจงปกติมาตรฐาน (Standard Normal Distribution) นั่นคือ $\mathbf{X}_0 \sim N(0, \mathbf{I})$ และการแจกแจงเอกรูป (Uniform Distribution) ในช่วง (a, b) ดังนี้

$$X_{01} \sim U(5, 10), X_{02} \sim U(10, 20), X_{03} \sim U(7, 9), X_{04} \sim U(6, 11), X_{05} \sim U(9, 19), X_{06} \sim U(4, 8)$$

ดังนั้นตัวแบบการถดถอยหรือตัวแบบที่แท้จริงตามสมการที่ (1) ที่พิจารณาในการศึกษาค้างนี้ คือ

ตัวแบบที่ 1 หรือตัวแบบระบุนยากมาก โดยมีตัวแปรอิสระจำนวน 4 ตัว และมีค่าสัมประสิทธิ์การถดถอยต่ำ

$$y = 1 + 0.5X_1 + 0.4X_2 + 0.3X_3 + 0.2X_4 + \varepsilon_0$$

ตัวแบบที่ 2 หรือตัวแบบระบุนยาก โดยมีตัวแปรอิสระจำนวน 2 ตัว และมีค่าสัมประสิทธิ์การถดถอยต่ำ

$$y = 1 + 0.5X_1 + 0.4X_2 + \varepsilon_0$$

ตัวแบบที่ 3 หรือตัวแบบระบุนง่ายมาก โดยมีตัวแปรอิสระจำนวน 2 ตัว และมีค่าสัมประสิทธิ์การถดถอยสูง

$$y = 1 + 2X_1 + 2X_2 + \varepsilon_0$$

ตัวแบบที่ 4 หรือตัวแบบระบุนง่าย โดยมีตัวแปรอิสระจำนวน 4 ตัว และมีค่าสัมประสิทธิ์การถดถอยสูง

$$y = 1 + 2X_1 + 2X_2 + 2X_3 + 2X_4 + \varepsilon_0$$

2. จากการสร้างตัวแปรอิสระทั้งหมด 6 ตัว จะพิจารณาตัวแปรอิสระเข้าตัวแบบครั้งละ 1 ตัว ตามลำดับ ทั้งหมด 6 ครั้ง (Sequentially Nested) โดยไม่ใช้การนำเข้าของตัวแปรอิสระทุกตัวที่เป็นไปได้ (All Possible Subset) เนื่องจากการทบทวนวรรณกรรมพบว่า วิธีการนำตัวแปรอิสระเข้าตัวแบบครั้งละ 1 ตัว ตามลำดับ หรือการนำเข้าทุกตัวที่เป็นไปได้มีผลการทดลองที่ไม่แตกต่างกัน (Keerativibool and Siripanich, 2017) โดยการพิจารณาตัวแปรอิสระเข้าตัวแบบครั้งละ 1 ตัว ตามลำดับ ดำเนินการดังนี้ ครั้งที่ 1 มีค่าคงที่ และ X_1 ($p = 2$) ครั้งที่ 2 มีค่าคงที่, X_1 และ X_2 ($p = 3$) ครั้งที่ 3 มีค่าคงที่, X_1 , X_2 และ X_3 ($p = 4$) ครั้งที่ 4 มีค่าคงที่, X_1 , X_2 , X_3 และ X_4 ($p = 5$) ครั้งที่ 5 มีค่าคงที่, X_1 , X_2 , X_3 , X_4 และ X_5 ($p = 6$) และครั้งที่ 6 มีค่าคงที่, X_1 , X_2 , X_3 , X_4 , X_5 และ X_6 ($p = 7$) สำหรับตัวแบบที่ 1 และ 4 จำนวนตัวแปรน้อยเกินไป หมายถึง ตัวแบบที่มีตัวแปรอิสระไม่ครบทุกตัว คือ มีค่าคงที่ และ X_1 หรือ มีค่าคงที่, X_1 และ X_2 หรือ มีค่าคงที่, X_1 , X_2 และ X_3 โดยไม่มีตัวแปร X_5 และ X_6 รวมอยู่ในตัวแบบ ส่วนจำนวนตัวแปรมากเกินไป หมายถึง ตัวแบบที่มีตัวแปรอิสระครบทุกตัว และมีตัวแปรที่ไม่เกี่ยวข้องรวมอยู่ในตัวแบบ คือ X_5 และ X_6 นั่นคือ มีค่าคงที่, X_1 , X_2 , X_3 , X_4 และ X_5 หรือ มีค่าคงที่, X_1 , X_2 , X_3 , X_4 , X_5 และ X_6 สำหรับตัวแบบที่ 2 และ 3 จำนวนตัวแปรน้อยเกินไป หมายถึง ตัวแบบที่มีค่าคงที่ และ X_1 ส่วนจำนวนตัวแปรมากเกินไป หมายถึง



ตัวแบบที่มีตัวแปรอิสระครบทุกตัว และมีตัวแปรที่ไม่เกี่ยวข้องรวมอยู่ในตัวแบบ คือ X_3, X_4, X_5 และ X_6 นั่นคือ มีค่าคงที่, X_1, X_2 และ X_3 หรือ มีค่าคงที่, X_1, X_2, X_3 และ X_4 หรือ มีค่าคงที่, X_1, X_2, X_3, X_4 และ X_5 หรือ มีค่าคงที่, X_1, X_2, X_3, X_4, X_5 และ X_6

3. จากแต่ละครั้งที่น่าตัวแปรอิสระเข้าตัวแบบ จะคำนวณค่าของเกณฑ์การคัดเลือกตัวแบบการถดถอยทั้ง 10 เกณฑ์ ได้แก่ AIC, BIC, HQIC, AICc, AICu, HQICc, KIC, KICc_C, KICc_{SB} และ KICc_{HM} แสดงดังสมการที่ (5) – (14) ตามลำดับ

$$AIC = \log(\hat{\sigma}^2) + \frac{2(p+1)}{n} \quad (5)$$

$$BIC = \log(\hat{\sigma}^2) + \frac{(p+1)\log(n)}{n} \quad (6)$$

$$HQIC = \log(\hat{\sigma}^2) + \frac{2(p+1)\log\log(n)}{n} \quad (7)$$

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

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

$$HQICc = \log(\hat{\sigma}^2) + \frac{2(p+1)\log\log(n)}{n-p-2} \quad (10)$$

$$KIC = \log(\hat{\sigma}^2) + \frac{3(p+1)}{n} \quad (11)$$

$$KICc_C = \log(\hat{\sigma}^2) + \log\left(\frac{n}{n-p}\right) + \frac{[(n-p)(2p+3)-2]}{(n-p-2)(n-p)} \quad (12)$$

$$KICc_{SB} = \log(\hat{\sigma}^2) + \frac{(p+1)(3n-p-2)}{n(n-p-2)} + \frac{p}{n(n-p)} \quad (13)$$

$$KICc_{HM} = \log(\hat{\sigma}^2) + \frac{(p+1)(3n-p-2)}{n(n-p-2)} \quad (14)$$



4. จากแต่ละเกณฑ์จะสรุปผลว่าทั้งหมด 6 ครั้งที่น่าตัวแปรอิสระเข้าตัวแบบ ครั้งใดมีค่าของเกณฑ์การคัดเลือกตัวแบบการถดถอยต่ำที่สุด จะได้ครั้งนั้นเป็นตัวแบบที่คัดเลือกไว้หรือตัวแบบโดยประมาณ เช่น เมื่อพิจารณาเกณฑ์ AIC ทั้งหมด 6 ครั้ง พบว่า ครั้งที่ 5 ให้ค่า AIC ต่ำที่สุด จะได้ตัวแบบที่คัดเลือกประกอบด้วยค่าคงที่, X_1 , X_2 , X_3 , X_4 และ X_5 ($p = 6$)
5. ดำเนินการข้อ 2 – 4 ซ้ำในแต่ละชุดข้อมูลที่จำลองขึ้น จำนวน 1,000 ชุด เพื่อสรุปผลว่าเกณฑ์ใดที่มีโอกาสการคัดเลือกตัวแบบได้ถูกต้องมากที่สุด นั่นคือ มีโอกาสได้ตัวแบบที่คัดเลือกไว้ในข้อ 4 ตรงกับตัวแบบที่สร้างขึ้นในข้อ 1 มากที่สุด

ผลการวิจัย

จากผลการจำลองข้อมูล จำนวน 1,000 ชุด ภายใต้สถานการณ์ต่าง ๆ ได้แก่ ตัวอย่างขนาดเล็ก ($n = 15$) ตัวอย่างขนาดกลาง ($n = 30$) และตัวอย่างขนาดใหญ่ ($n = 100$) จำนวนพารามิเตอร์ในตัวแบบ 2 ระดับ คือ $p_0 = 3$ และ $p_1 = 5$ สัมประสิทธิ์การถดถอยแตกต่างกัน 4 ตัวแบบ คือ ตัวแบบระบุนยากมาก มีตัวแปรอิสระ 4 ตัว มีค่าสัมประสิทธิ์การถดถอยต่ำ ตัวแบบระบุนยาก มีตัวแปรอิสระ 2 ตัว มีค่าสัมประสิทธิ์การถดถอยต่ำ ตัวแบบระบุนง่ายมาก มีตัวแปรอิสระ 2 ตัว มีค่าสัมประสิทธิ์การถดถอยสูง และตัวแบบระบุนง่าย มีตัวแปรอิสระ 4 ตัว มีค่าสัมประสิทธิ์การถดถอยสูง กำหนดความคลาดเคลื่อนมีการแจกแจงปกติ มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวน 3 ระดับ คือ ความแปรปรวนต่ำ ($\sigma_0^2 = 0.25$) ความแปรปรวนปานกลาง ($\sigma_0^2 = 1$) และความแปรปรวนสูง ($\sigma_0^2 = 9$) และการแจกแจงของตัวแปรอิสระ 2 การแจกแจง คือ การแจกแจงปกติ และการแจกแจงเอกรูป ได้ผลการวิจัยแสดงดังตารางที่ 1 – 6 ซึ่งสรุปเป็นเกณฑ์การคัดเลือกตัวแบบที่เหมาะสมภายใต้สถานการณ์ต่าง ๆ ดังตารางที่ 7 โดยพบว่า กรณีตัวอย่างขนาดเล็ก ตัวแบบระบุนยากถึงยากมาก เกณฑ์ที่เหมาะสม คือ AIC, HQIC และถ้าตัวแบบระบุนง่ายถึงง่ายมาก เกณฑ์ที่เหมาะสม คือ AICu, KICc กรณีตัวอย่างขนาดกลาง ตัวแบบระบุนยากถึงยากมาก เกณฑ์ที่เหมาะสม คือ AIC และถ้าตัวแบบระบุนง่ายถึงง่ายมาก เกณฑ์ที่เหมาะสม คือ AICu และ KICc สำหรับกรณีตัวอย่างขนาดใหญ่ ตัวแบบระบุนยากถึงยากมาก เกณฑ์ที่เหมาะสม คือ AIC, BIC และถ้าตัวแบบระบุนง่ายถึงง่ายมาก เกณฑ์ที่เหมาะสม คือ BIC



ตารางที่ 1 ร้อยละของการคัดเลือกตัวแบบ เมื่อตัวอย่างมีขนาดเล็ก ($n = 15$) และตัวแปรอิสระมีการแจกแจงปกติมาตรฐาน

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

หมายเหตุ ตัวทึบและขีดเส้นใต้ หมายถึง เกณฑ์ที่มีร้อยละของการคัดเลือกตัวแบบสูง

ตารางที่ 2 ร้อยละของการคัดเลือกตัวแบบ เมื่อตัวอย่างมีขนาดเล็ก ($n = 15$) และตัวแปรอิสระมีการแจกแจงเอกรูป

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

หมายเหตุ ตัวทึบและขีดเส้นใต้ หมายถึง เกณฑ์ที่มีร้อยละของการคัดเลือกตัวแบบสูง

ตารางที่ 3 ร้อยละของการคัดเลือกตัวแบบ เมื่อตัวอย่างมีขนาดกลาง ($n = 30$) และตัวแปรอิสระมีการแจกแจงปกติมาตรฐาน

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

หมายเหตุ ตัวทึบและขีดเส้นใต้ หมายถึง เกณฑ์ที่มีร้อยละของการคัดเลือกตัวแบบสูง

ตารางที่ 4 ร้อยละของการคัดเลือกตัวแบบ เมื่อตัวอย่างมีขนาดกลาง ($n = 30$) และตัวแปรอิสระมีการแจกแจงเอกรูป

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

หมายเหตุ ตัวทึบและขีดเส้นใต้ หมายถึง เกณฑ์ที่มีร้อยละของการคัดเลือกตัวแบบสูง



ตารางที่ 5 ร้อยละของการคัดเลือกตัวแบบ เมื่อตัวอย่างมีขนาดใหญ่ ($n = 100$) และตัวแปรอิสระมีการแจกแจงปกติมาตรฐาน

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

หมายเหตุ ตัวทึบและขีดเส้นใต้ หมายถึง เกณฑ์ที่มีร้อยละของการคัดเลือกตัวแบบสูง

ตารางที่ 6 ร้อยละของการคัดเลือกตัวแบบ เมื่อตัวอย่างมีขนาดใหญ่ ($n = 100$) และตัวแปรอิสระมีการแจกแจงเอกรูป

σ_0^2	ตัวแบบระบุ	จำนวนตัวแปร	AIC	BIC	HQIC	AICc	AICu	HQICc	KIC	KICc _c	KICc _{SB}	KICc _{HM}
ต่ำ	ยากมาก	น้อยเกินไป	57.0	90.1	76.8	63.8	78.7	80.7	76.1	78.9	78.7	78.5
		ถูกต้อง	<u>27.5</u>	8.3	17.5	24.9	16.7	15.3	17.6	16.5	16.7	16.9
		มากเกินไป	15.5	1.6	5.7	11.3	4.6	4.0	6.3	4.6	4.6	4.6
ต่ำ	ยาก	น้อยเกินไป	14.9	43.4	28.4	16.4	29.6	30.4	27.8	29.6	29.4	29.2
		ถูกต้อง	58.2	52.9	61.5	61.4	61.1	61.0	<u>62.1</u>	61.1	61.2	61.3
		มากเกินไป	26.9	3.7	10.1	22.2	9.3	8.6	10.1	9.3	9.4	9.5
ต่ำ	ง่ายมาก	น้อยเกินไป	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		ถูกต้อง	70.9	<u>95.3</u>	86.1	75.1	88.3	89.1	85.5	88.4	88.1	87.9
		มากเกินไป	29.1	4.7	13.9	24.9	11.7	10.9	14.5	11.6	11.9	12.1
ต่ำ	ง่าย	น้อยเกินไป	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		ถูกต้อง	75.5	<u>94.7</u>	88.4	80.8	90.7	91.2	87.8	90.7	90.7	90.6
		มากเกินไป	24.5	5.3	11.6	19.2	9.3	8.8	12.2	9.3	9.3	9.4
กลาง	ยากมาก	น้อยเกินไป	81.7	98.8	95.1	86.5	95.8	96.4	94.5	95.9	95.7	95.7
		ถูกต้อง	<u>10.4</u>	1.0	3.4	8.8	3.1	2.7	3.8	3.0	3.2	3.2
		มากเกินไป	7.9	0.2	1.5	4.7	1.1	0.9	1.7	1.1	1.1	1.1
กลาง	ยาก	น้อยเกินไป	48.7	83.9	68.6	52.6	70.6	72.1	67.5	70.8	70.4	70.1
		ถูกต้อง	<u>30.9</u>	15.2	24.6	30.8	23.7	22.7	25.0	23.5	23.9	24.2
		มากเกินไป	20.4	0.9	6.8	16.6	5.7	5.2	7.5	5.7	5.7	5.7
กลาง	ง่ายมาก	น้อยเกินไป	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		ถูกต้อง	72.3	<u>96.7</u>	88.8	77.4	89.9	91.2	88.3	90.1	89.9	89.8
		มากเกินไป	27.7	3.3	11.2	22.6	10.1	8.8	11.7	9.9	10.1	10.2
กลาง	ง่าย	น้อยเกินไป	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		ถูกต้อง	75.7	<u>94.7</u>	87.8	81.1	89.3	90.4	87.6	89.6	89.3	89.3
		มากเกินไป	24.3	5.3	12.2	18.9	10.7	9.6	12.4	10.4	10.7	10.7
สูง	ยากมาก	น้อยเกินไป	87.2	99.5	97.3	91.6	97.7	98.2	97.2	97.8	97.7	97.7
		ถูกต้อง	<u>5.2</u>	0.4	1.8	3.9	1.7	1.4	1.8	1.7	1.7	1.7
		มากเกินไป	7.6	0.1	0.9	4.5	0.6	0.4	1.0	0.5	0.6	0.6
สูง	ยาก	น้อยเกินไป	67.9	95.5	86.2	72.6	87.6	88.3	85.4	87.7	87.6	87.5
		ถูกต้อง	<u>14.4</u>	3.9	8.6	13.7	8.4	8.1	9.0	8.4	8.4	8.5
		มากเกินไป	17.7	0.6	5.2	13.7	4.0	3.6	5.6	3.9	4.0	4.0
สูง	ง่ายมาก	น้อยเกินไป	25.8	56.9	40.3	28.7	42.3	43.4	39.4	42.3	42.1	41.9
		ถูกต้อง	49.4	40.2	50.5	50.5	50.0	49.4	<u>50.7</u>	50.1	50.0	50.1
		มากเกินไป	24.8	2.9	9.2	20.8	7.7	7.2	9.9	7.6	7.9	8.0
สูง	ง่าย	น้อยเกินไป	35.6	78.6	57.4	41.4	62.0	63.7	56.5	62.2	61.6	61.5
		ถูกต้อง	<u>46.4</u>	19.9	35.8	44.8	32.5	31.6	36.1	32.3	32.3	32.4
		มากเกินไป	18.0	1.5	6.8	13.8	5.5	4.7	7.4	5.5	6.1	6.1

หมายเหตุ ตัวทึบและขีดเส้นใต้ หมายถึง เกณฑ์ที่มีร้อยละของการคัดเลือกตัวแบบสูง

ตารางที่ 7 เกณฑ์การคัดเลือกตัวแบบที่เหมาะสมภายใต้สถานการณ์ต่าง ๆ

n	σ_0^2	การระบุตัวแบบ	เกณฑ์ที่แนะนำ	การระบุตัวแบบ	เกณฑ์ที่แนะนำ	
เล็ก	ต่ำ	ยากมาก	AIC, HQIC	ง่ายมาก	AICu, KICc	
		ยาก	AIC, BIC, HQIC, AICc, HQICc, KIC	ง่าย	AICc, AICu, HQICc, KICc	
	กลาง	ยากมาก	AIC, HQIC	ง่ายมาก	AICc, AICu, HQICc, KICc	
		ยาก	AIC, BIC, HQIC, AICc, HQICc	ง่าย	BIC, AICu, KICc	
	สูง	ยากมาก	AIC, HQIC	ง่ายมาก	AIC, HQIC, AICc, HQICc	
		ยาก	AIC, HQIC	ง่าย	AIC, BIC, HQIC	
กลาง	ต่ำ	ยากมาก	AIC, AICc	ง่ายมาก	AICu, KICc	
		ยาก	AICc, AICu, KICc	ง่าย	AICu, KICc	
	กลาง	ยากมาก	AIC	ง่ายมาก	AICu, HQICc, KICc	
		ยาก	AIC, HQIC, AICc	ง่าย	AICc, AICu, KICc	
	สูง	ยากมาก	AIC	ง่ายมาก	AIC, AICu, KICc	
		ยาก	AIC	ง่าย	AIC, AICc	
	ใหญ่	ต่ำ	ยากมาก	AIC, BIC	ง่ายมาก	BIC
			ยาก	BIC, KIC	ง่าย	BIC
		กลาง	ยากมาก	AIC	ง่ายมาก	BIC
			ยาก	AIC, BIC	ง่าย	BIC
สูง		ยากมาก	AIC	ง่ายมาก	BIC, KIC	
		ยาก	AIC, AICc	ง่าย	AIC, BIC	

วิจารณ์ผลการวิจัย

จากภาพรวมของตารางที่ 1 – 6 พบว่า เมื่อความแปรปรวนของความคลาดเคลื่อนเพิ่มขึ้น ร้อยละของการคัดเลือกตัวแบบถูกต้องจะลดลง สอดคล้องกับการศึกษาของ Keerativibool and Siripanich (2017) ที่พบว่า ประสิทธิภาพของทุกเกณฑ์การคัดเลือกตัวแบบจะลดลงเมื่อความแปรปรวนของความคลาดเคลื่อนมีค่าเพิ่มขึ้น ซึ่งการศึกษารังนี้ยังพบว่า ถ้าตัวแบบระบุยากถึงยากมาก ความแปรปรวนของความคลาดเคลื่อนที่เพิ่มขึ้น จะทำให้ร้อยละของการคัดเลือกตัวแบบถูกต้องลดลงอย่างมาก และมีโอกาสมีจำนวนตัวแปรน้อยเกินไปมากกว่าการมีจำนวนตัวแปรมากเกินไป เมื่อการแจกแจงของตัวแปรอิสระเปลี่ยนจากการแจกแจงปกติมาตรฐานเป็นการแจกแจงเอกรูป ร้อยละของการคัดเลือกตัวแบบถูกต้องจะลดลง และถ้าความแปรปรวนของความคลาดเคลื่อนต่ำ ตัวแบบระบุยากถึงยากมาก เมื่อการแจกแจงของตัวแปรอิสระเปลี่ยน จะมีโอกาสมีจำนวนตัวแปรน้อยเกินไปมากกว่าการมีจำนวนตัวแปรมากเกินไป แต่ถ้าตัวแบบระบุง่ายถึงง่ายมาก จะมีโอกาสมีจำนวนตัวแปรมากเกินไปมากกว่าการมีจำนวนตัวแปรน้อยเกินไป ขณะที่ความแปรปรวนของความคลาดเคลื่อนปานกลางถึงสูงในทุกลักษณะของตัวแบบ จะมีโอกาสมีจำนวนตัวแปรน้อยเกินไปมากกว่าการมีจำนวนตัวแปรมากเกินไป กรณีตัวอย่างมีขนาดเล็ก ร้อยละของการคัดเลือกตัวแบบถูกต้องจะต่ำ และเมื่อขนาดตัวอย่างเพิ่มขึ้น ร้อยละของการคัดเลือกตัวแบบถูกต้องจะเพิ่มขึ้น สอดคล้องกับการศึกษาของ Sangthong (2019) ที่พบว่า ประสิทธิภาพของเกณฑ์การคัดเลือกตัวแบบค่อนข้างต่ำเมื่อจำนวน



กลุ่มมีขนาดเล็ก และประสิทธิภาพการคัดเลือกตัวแบบจะสูงขึ้นเมื่อจำนวนกลุ่มมากขึ้น อย่างไรก็ตาม การศึกษาครั้งนี้ให้ผล การศึกษาขัดแย้งกับการศึกษาของ Keerativibool (2014a) ที่พบว่า เกณฑ์ $KICc_0$ มีประสิทธิภาพสูงที่สุด อาจเนื่องมาจากการศึกษาครั้งนี้ได้เปรียบเทียบประสิทธิภาพของเกณฑ์การคัดเลือกตัวแบบที่หลากหลายเกณฑ์มากกว่า

สรุปผลการวิจัย

การศึกษานี้ได้ศึกษาเปรียบเทียบประสิทธิภาพของเกณฑ์การคัดเลือกตัวแบบสำหรับตัวแบบการถดถอย 10 เกณฑ์ ได้แก่ เกณฑ์ AIC, BIC, HQIC, AICc, AICu, HQICc, KIC, $KICc_0$, $KICc_{SB}$ และ $KICc_{HM}$ ด้วยการจำลองข้อมูล จำนวน 1,000 ชุด ภายใต้สถานการณ์ต่าง ๆ ได้แก่ ตัวอย่างขนาดเล็ก ($n = 15$) ตัวอย่างขนาดกลาง ($n = 30$) และตัวอย่างขนาดใหญ่ ($n = 100$) จำนวนพารามิเตอร์ในตัวแบบ 2 ระดับ คือ $p_0 = 3$ และ $p_0 = 5$ สัมประสิทธิ์การถดถอยแตกต่างกัน 4 ตัวแบบ คือ ตัวแบบระบุนยากมาก มีตัวแปรอิสระ 4 ตัว มีค่าสัมประสิทธิ์การถดถอยต่ำ ตัวแบบระบุนยาก มีตัวแปรอิสระ 2 ตัว มีค่าสัมประสิทธิ์การถดถอยต่ำ ตัวแบบระบุนง่ายมาก มีตัวแปรอิสระ 2 ตัว มีค่าสัมประสิทธิ์การถดถอยสูง และตัวแบบระบุนง่าย มีตัวแปรอิสระ 4 ตัว มีค่าสัมประสิทธิ์การถดถอยสูง กำหนดความคลาดเคลื่อนมีการแจกแจงปกติ มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวน 3 ระดับ คือ ความแปรปรวนต่ำ ($\sigma_0^2 = 0.25$) ความแปรปรวนปานกลาง ($\sigma_0^2 = 1$) และความแปรปรวนสูง ($\sigma_0^2 = 9$) ผลการศึกษาพบว่า กรณีตัวอย่างขนาดเล็ก เกณฑ์ AIC, HQIC เหมาะสมกับกรณีตัวแบบระบุนยากถึงยากมาก แต่เกณฑ์ AICu, $KICc$ เหมาะสมกับกรณีตัวแบบระบุนง่ายถึงง่ายมาก กรณีตัวอย่างขนาดกลาง เกณฑ์ AIC เหมาะสมกับกรณีตัวแบบระบุนยากถึงยากมาก แต่เกณฑ์ AICu และ $KICc$ เหมาะสมกับกรณีตัวแบบระบุนง่ายถึงง่ายมาก สำหรับกรณีตัวอย่างขนาดใหญ่ เกณฑ์ AIC, BIC เหมาะสมกับกรณีตัวแบบระบุนยากถึงยากมาก แต่เกณฑ์ BIC เหมาะสมกับกรณีตัวแบบระบุนง่ายถึงง่ายมาก ดังนั้นหากพิจารณาในภาพรวมพบว่า เกณฑ์ที่เหมาะสมในกรณีตัวอย่างขนาดเล็กถึงขนาดกลาง คือ เกณฑ์ AIC แต่กรณีตัวอย่างขนาดใหญ่ควรใช้เกณฑ์ BIC การศึกษานี้ยังพบว่า เมื่อความคลาดเคลื่อนมีความแปรปรวนเพิ่มขึ้น ทำให้ร้อยละของการคัดเลือกตัวแบบถูกต้องลดลง โดยเฉพาะถ้าตัวแบบระบุนยากถึงยากมาก ความแปรปรวนของความคลาดเคลื่อนที่เพิ่มขึ้น ทำให้ร้อยละของการคัดเลือกตัวแบบถูกต้องลดลงอย่างมาก และมีโอกาสที่จะมีจำนวนตัวแปรน้อยเกินไปมากกว่าการมีจำนวนตัวแปรมากเกินไป กรณีการแจกแจงของตัวแปรอิสระเป็นการแจกแจงเอกรูป จะมีร้อยละของการคัดเลือกตัวแบบถูกต้องต่ำกว่าการแจกแจงปกติมาตรฐาน เกณฑ์การคัดเลือกตัวแบบสำหรับตัวแบบการถดถอยทุกเกณฑ์ กรณีตัวอย่างมีขนาดเล็ก จะมีโอกาสคัดเลือกตัวแบบได้ถูกต้องน้อย และคัดเลือกได้ถูกต้องเพิ่มขึ้นเมื่อตัวอย่างมีขนาดใหญ่ขึ้น

กิตติกรรมประกาศ

การศึกษานี้ได้รับงบประมาณสนับสนุนจากสำนักงานสภานโยบายการอุดมศึกษา วิทยาศาสตร์ วิจัยและนวัตกรรมแห่งชาติ โดยมหาวิทยาลัยทักษิณ ประจำปีงบประมาณ พ.ศ. 2564



เอกสารอ้างอิง

- Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, 19(6), 716-723.
- Cavanaugh, J.E. (1999). A Large-Sample Model Selection Criterion Based on Kullback's Symmetric Divergence. *Statistics and Probability Letters*, 42(4), 333-343.
- Cavanaugh, J.E. (2004). Criteria for Linear Model Selection Based on Kullback's Symmetric Divergence. *Australian and New Zealand Journal of Statistics*, 46(2), 257-274.
- Hafidi, B. & Mkhadri, A. (2006). A Corrected Akaike Criterion Based on Kullback's Symmetric Divergence: Applications in Time Series, Multiple and Multivariate Regression. *Computational Statistics and Data Analysis*, 50(6), 1524-1550.
- Hannan, E.J. & Quinn, B.G. (1979). The Determination of the Order of an Autoregression. *Journal of the Royal Statistical Society Series B*, 41(2), 190-195.
- Hurvich, C.M. & Tsai, C.L. (1989). Regression and Time Series Model Selection in Small Samples. *Biometrika*, 76(2), 297-307.
- Keerativibool, W. (2014a). Unifying the Derivations of Kullback Information Criterion and Corrected Versions. *Thailand Statistician Journal of Thai Statistical Association*, 12(1), 37-53.
- Keerativibool, W. (2014b). Study on the Penalty Functions of Model Selection Criteria. *Thailand Statistician Journal of Thai Statistical Association*, 12(2), 161-178.
- Keerativibool, W. & Siripanich, P. (2017). Comparison of the Model Selection Criteria for Multiple Regression Based on Kullback-Leibler's Information. *Chiang Mai Journal of Science*, 44(2), 699-714.
- McQuarrie, A.D.R. & Tsai, C.L. (1998). Regression and Time Series Model Selection. World Scientific, Singapore.



- McQuarrie, A.D.R., Shumway, R.H. & Tsai, C.L. (1997). The Model Selection Criterion AICu. *Statistics and Probability Letters*, 34(3), 285-292.
- Montgomery, D.C., Peck, E.A. & Vining, G.G. (2006). *Introduction to Linear Regression Analysis*. (4th ed.). New York: John Wiley & Sons.
- Neath, A. & Cavanaugh, J.E. (1997). Regression and Time Series Model Selection Using Variants of the Schwarz Information Criterion. *Communication in Statistic-Theory and Method*, 26(3), 559-580.
- Sangthong, M. (2019). A Study of the Effectiveness of Model Selection Criteria for Multilevel Analysis. *Burapha Science Journal*, 24(1), 156-169. (in Thai)
- Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics* 6(2), 461-464.
- Seghouane, A.K. & Bekara, M. (2004). A Small Sample Model Selection Criterion Based on Kullback's Symmetric Divergence. *IEEE Transactions on Signal Processing*, 52(12), 3314-3323.