

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบความถูกต้องของการประมาณค่าสัมประสิทธิ์การถดถอยของตัวแบบถดถอยพหุคูณเมื่อเกิดความคลาดเคลื่อนในตัวแปรอิสระ โดยจะเปรียบเทียบวิธีการประมาณค่าสัมประสิทธิ์การถดถอย 3 วิธี ได้แก่ วิธีการกำลังสองน้อยสุด (Ordinary Least Squares method (OLS)) วิธีการรวมกำลังสองน้อยสุด (Total Least Squares method (TLS)) และวิธีการถดถอยของตัวแปรแฝงเชิงเบย์ (Bayesian Latent Variable Regression method (BLVR)) เกณฑ์ที่ใช้ในการตัดสินใจคือเกณฑ์ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (Average Mean Squares Error (AMSE)) และเกณฑ์ที่เป็นส่วนประกอบของการตัดสินใจคือค่าอัตราส่วนผลต่างของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (Ratio of Different Average Mean Squares Error (RDAMSE)) สถานการณ์ที่ศึกษา คือจะกำหนด $\beta = (1, 1, \dots, 1)'$ สำหรับวิธี OLS และ TLS แต่สำหรับกรณีวิธี BLVR จะกำหนด β ให้มีการแจกแจงแบบปกติซึ่งมีค่าเฉลี่ยเท่ากับ 1 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.5 จำนวนตัวแปรอิสระ (p) ที่ศึกษาคือ 3 5 และ 7 และขนาดตัวอย่าง (n) ที่ศึกษาเท่ากับ 30 50 70 100 และ 150 ตัวแปรอิสระทุกตัวมีความคลาดเคลื่อนซึ่งความคลาดเคลื่อนสุ่มในตัวแปรอิสระแต่ละตัว (e_{-j}) มีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 และมีส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.1 0.3 0.5 0.7 และ 1.0 ตามลำดับ ความคลาดเคลื่อนในตัวแปรตาม (e_{-y}) มีการแจกแจงแบบปกติซึ่งมีค่าเฉลี่ยเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.1 0.3 0.5 0.7 และ 1.0 ตามลำดับ ข้อมูลที่ใช้ในการวิจัยได้จากการจำลองแบบโดยวิธีมอนติคาร์โลซึ่งกระทำซ้ำ 500 รอบ ในแต่ละสถานการณ์ ผลของการวิจัยสามารถสรุปได้ดังนี้

การเปรียบเทียบค่า AMSE ของทั้ง 3 วิธีพบว่า เมื่อส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนในตัวแปรอิสระเท่ากับ 0.1 และส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนในตัวแปรตามเท่ากับ 0.1 0.3 และ 0.5 วิธี OLS จะให้ค่า AMSE ต่ำที่สุด แต่ในกรณีที่ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนในตัวแปรตามเท่ากับ 0.7 และ 1.0 วิธี OLS จะดีที่สุดเมื่อขนาดตัวอย่างใหญ่ และวิธี BLVR จะดีที่สุดในกรณีที่ขนาดตัวอย่างเล็ก ส่วนกรณีที่ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนในตัวแปรอิสระเท่ากับ 0.3 วิธี OLS จะให้ค่า AMSE ต่ำที่สุดเมื่อจำนวนตัวแปรอิสระเท่ากับ 3 และขนาดตัวอย่างใหญ่ โดยขนาดตัวอย่างจะแปรผันตามส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนในตัวแปรตาม ส่วนกรณีที่จำนวนตัวแปรอิสระเท่ากับ 5 และ 7 ค่า AMSE ของวิธี BLVR จะมีค่าต่ำสุด ส่วนกรณีที่ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนในตัวแปรอิสระเท่ากับ 0.5 0.7 และ 1.0 จะให้ผลสรุปเหมือนกันคือวิธี BLVR จะให้ค่า AMSE ต่ำที่สุด รองลงมาคือวิธี TLS และวิธี OLS ตามลำดับ

ปัจจัยที่มีผลต่อค่า AMSE ของทุกวิธี คือ ขนาดตัวอย่าง จำนวนตัวแปรอิสระ ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนในตัวแปรอิสระ และส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนในตัวแปรตาม

The objective of this research is the comparison on the accuracy of regression-coefficient estimation of multiple regression with errors in independent variables. This research compares three multiple-regression-coefficient estimation methods, Ordinary Least Squares method (OLS), Total Least Squares method (TLS), and Bayesian Latent Variable Regression method (BLVR). The measurement for making decision is Average Mean Squares Error (AMSE) and Ratio of Different Average Mean Squares Error (RDAMSE). As for the case study, we specify $\beta = (1, 1, \dots, 1)'$ for OLS and TLS methods. However, in the case of BLVR method, the distribution of β is assumed to be normal distribution with mean equal to 1 and standard deviation equal to 0.5, the numbers of independent variables are equivalent to 3, 5 and 7, and the sample sizes are 30, 50, 70, 100 and 150. All independent variables have errors which random error in each independent variable (e_{x_j}) has the average of normal distribution equivalent to 0 and standard deviations are equal to 0.1, 0.3, 0.5, 0.7 and 1.0, respectively. The distribution of error in the dependent variable (ε_y) is normal distribution with mean equal to 0 and standard deviations equal to 0.1, 0.3, 0.5, 0.7 and 1.0, respectively. The data for this research is simulated by using the Monte Carlo simulation technique with 500 repetitions for each case. The results of this research are as follows:

According to the comparison of AMSE from three referred methods, we found that when the standard deviation of the errors in independent variables is equal to 1 and the standard deviations of the error in dependent variable are equal to 0.1, 0.3 and 0.5, OLS method would give the lowest AMSE. Conversely, when standard deviations of error in dependent variable are equal to 0.7 and 1.0, OLS method is the best method for large sample size while BLVR method is the most appropriate for small sample size. Additionally, in case of the standard deviation of errors in independent variables is equal to 0.3, OLS method would give the lowest AMSE when the number of independent variables is equal to 3 and with large sample size. As well, when the numbers of independent variables are equal to 5 and 7, AMSE value from BLVR method is the lowest. Furthermore, when the standard deviations of errors in independent variables are equal to 0.5, 0.7, and 1.0, from the result, the lowest AMSE is from BLVR method, while TLS and OLS are the second and the third lowest, respectively.

From all methods, the factors that effect to AMSE are sample size, the number of independent variables, the standard deviation of errors in independent variables and the standard deviation of error in dependent variable.