

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ (สัมประสิทธิ์การถดถอย) ในสมการการถดถอยเชิงเส้นพหุคูณเมื่อตัวอย่างมีขนาดเล็ก ในที่นี้ทำการศึกษาวิธีการประมาณ 4 วิธีคือวิธีกำลังสองน้อยสุด (OLS) วิธีบุคคลแปร (BS) วิธีตัวประมาณ M เมื่อใช้เกณฑ์ความแกร่งของแรมเซย์ (M/R) และวิธีตัวประมาณ M เมื่อใช้เกณฑ์ความแกร่งของตุ๊กกี (M/T) ซึ่งเกณฑ์ที่ใช้ในการเปรียบเทียบคือค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ภายใต้สถานการณ์ต่างๆที่ศึกษา ดังนี้

ก) จำนวนตัวแปรอิสระเท่ากับ 3 และ 6

ข) ขนาดตัวอย่างเท่ากับ 10, 15, 20, 25 และ 30 ตามลำดับ

ค) การแจกแจงของความคลาดเคลื่อนที่ศึกษามี 3 การแจกแจงคือการแจกแจงปกติ การแจกแจงปกติปลอมปน และการแจกแจงแกมมา การแจกแจงปกติจะใช้ค่าเฉลี่ยเท่ากับ 0 และ 1 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1, 2 และ 3 ส่วนการแจกแจงปกติปลอมปนจะใช้ค่าสเกลแฟกเตอร์เท่ากับ 10 เปอร์เซนต์การปลอมปนเท่ากับ 25%, 50% และ 75% และการแจกแจงแกมมาจะใช้ค่าพารามิเตอร์เบต้าเท่ากับ 1 และค่าพารามิเตอร์แอลฟาเท่ากับ 0.5, 1, 2 และ 4 ตามลำดับ

ข้อมูลที่ใช้ในงานวิจัยนี้ได้จากการจำลองด้วยเทคนิคมอนติ-คาร์โล ซึ่งกระทำซ้ำ 1,000 ครั้งในแต่ละสถานการณ์ โดยใช้โปรแกรม S-PLUS2000 ผลการวิจัยสรุปได้ดังนี้

กรณีความคลาดเคลื่อนมีการแจกแจงปกติ และการแจกแจงปกติปลอมปน วิธี BS มีค่า AMSE ต่ำกว่าวิธีอื่นเมื่อขนาดตัวอย่างน้อย (น้อยกว่า 20) แต่เมื่อขนาดตัวอย่างมากขึ้นและจำนวนตัวแปรอิสระมากขึ้น วิธี M/T จะมีค่า AMSE ต่ำสุด แต่เมื่อความคลาดเคลื่อนมีการแจกแจงแกมมาวิธี M/R มีค่า AMSE ต่ำสุดทุกระดับขนาดตัวอย่างและจำนวนตัวแปรอิสระ

โดยทั่วไปค่า AMSE จะแปรผันตามปัจจัยดังต่อไปนี้ เรียงลำดับจากมากไปน้อยดังนี้ ค่าความแปรปรวนของค่าความคลาดเคลื่อน จำนวนตัวแปรอิสระ และเปอร์เซนต์การปลอมปนของค่าผิดปกติในค่าความคลาดเคลื่อน แต่จะแปรผกผันกับขนาดตัวอย่าง

The objective of this research is to compare the estimation of parameters (regression coefficients) in the multiple linear regression with small sample size. In this study, the estimation methods are the Ordinary Least Square method (OLS), Bootstrap method (BS), M-estimator method using Ramsay's robust criteria (M/R) and M-estimator method using Tukey's robust criteria (M/T). They were compared by average of mean square error (AMSE). The comparison was done under several situations which are as follows :

- a) The number of independent variables are 3 and 6
- b) The sample sizes are 10, 15, 20, 25 and 30
- c) This study used three residual distributions which are normal distribution, scale-contaminated normal distribution and gamma distribution. For normal distribution, the means of 0 and 1, the standard deviations of 1, 2, and 3 are used. For scale-contaminated normal distribution, the scale factor of 10, the percent of contaminations of 25%, 50% and 75% are used. For gamma distribution, the beta parameter of 1, the alpha parameter of 0.5, 1, 2 and 4 are used.

The data for this research is simulated by using the Monte-Carlo simulation technique with 1,000 repetitions for each situation by S-PLUS 2000 program. The results of this research are as follows :

In case that residuals have normal distribution and scale-contaminated normal distribution, the AMSE of BS method is lower than the others when the sample sizes are small (lower than 20) but when the sample size and the number of independent variables are increased, the AMSE of M/T method is lower than the others. But in case that residuals have gamma distribution, the AMSE of M/R method is lower than the others every levels of the sample size and the number of independent variables.

In general, the AMSE varies with, most to least respectively, the variance of the residuals, the number of independent variables and the percent of contaminations of outliers in the residuals but coverse to sample size.