การวิจัยนี้มีจุดมุ่งหมายเพื่อศึกษาและเปรียบเทียบวิธีประมาณค่าสัมประสิทธิ์การ ถดถอยเชิงเส้นพหุ เมื่อเกิดปัญหาพหุสัมพันธ์ระหว่างตัวแปรอิสระและข้อมูลมีค่าผิดปกติ ด้วยวิธี ริดจ์ชนิดตัวประมาณจีเอ็ม (GMR) ซึ่งได้จากการบวกสมาชิกทุกตัวบนเส้นทแยงมุมของเมทริกซ์ $X'X$ ด้วยค่าคงที่ k ที่ได้จากการประมาณด้วยวิธี Hoerl – Kennard (HK) และวิธี Hoerl – Kennard – Baldwin (HKB) ใช้เกณฑ์ความแกร่ง 6 เกณฑ์ คือ เกณฑ์ความแกร่งของ Huber, Ramsay, Andrews, Hample's 17A, Barya และ Talwar กำหนดจำนวนตัวแปรอิสระที่ใช้ใน การศึกษาเท่ากับ 3 ตัวแปร ขนาดตัวอย่างเท่ากับ 20, 50 และ 100 ระดับความสัมพันธ์ของตัวแปร อิสระเป็น 3 ระดับ คือ ระดับต่ำ (ρ =0.1) ระดับปานกลาง (ρ =0.5) และระดับสูง (ρ = 0.9) กำหนดค่า ผิดปกติ 2 ระดับ คือ ระดับปานกลางและระดับรุนแรง เปอร์เซ็นต์การเกิดค่าผิดปกติ เท่ากับ 1% และ 5% ของขนาดตัวอย่าง โดยแบ่งกรณีศึกษาเป็น 5 กรณี ดังนี้ ข้อมูลมีค่าผิดปกติใน ตัวแปร Y อย่างเดียว ข้อมูลมีค่าผิดปกติในตัวแปร X_3 อย่างเดียว ข้อมูลมีค่าผิดปกติในตัวแปร $\left(X_1, X_2\right)$, $\left(Y, X_3\right)$ และ $\left(Y, X_1, X_2\right)$ ทำการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล และกระทำ ซ้ำ 1,000 รอบในแต่ละสถานการณ์ การเปรียบเทียบพิจารณาจากค่าเฉลี่ยความคลาดเคลื่อน กำลังสอง (MSE)

ผลการวิจัยพบว่า ภายใต้กรณีที่ศึกษาวิธีริดจ์ชนิดตัวประมาณจีเอ็มเมื่อประมาณค่าคง ที่ k ด้วยวิธี HKB ให้ค่า MSE ต่ำกว่าวิธี HK เป็นส่วนใหญ่ และเมื่อขนาดตัวอย่างเพิ่มขึ้น การประมาณค่าคงที่ k ทั้งสองวิธีให้ค่า MSE ใกล้เคียงกัน นอกจากนี้เมื่อพิจารณาทุกระดับความ รุนแรงของค่าผิดปกติ เปอร์เซ็นต์ค่าผิดปกติ ระดับขนาดของความสัมพันธ์ และทุกลักษณะของตัว แปรที่เกิดค่าผิดปกติ พบว่าส่วนใหญ่วิธีริดจ์ชนิดตัวประมาณจีเอ็มเมื่อใช้เกณฑ์ความแกร่งของ Ramsay จะให้ค่า MSE ต่ำที่สุด เมื่อขนาดตัวอย่างเท่ากับ 20 และวิธีริดจ์ชนิดตัวประมาณจีเอ็ม เมื่อใช้เกณฑ์ความแกร่งของ Huber และ Ramsay ให้ค่า MSE ใกล้เคียงกันและมีค่าต่ำที่สุดเป็น ส่วนใหญ่ เมื่อขนาดตัวอย่างเท่ากับ 50 และ 100 นอกจากนี้ยังพบว่า MSE มีแนวโน้มสูงขึ้น เมื่อ ระดับความรุนแรงของค่าผิดปกติและเปอร์เซ็นต์การเกิดค่าผิดปกติเพิ่มขึ้น แต่ MSE มีแนวโน้ม ลดลง เมื่อขนาดตัวอย่างเพิ่มขึ้น

The objective of this research is to study and compare the estimation methods of regression coefficients in the multiple linear regression with the occurrences of multicollinearity and outliers using Generalized M – estimator Ridge Methods (GMR), with the addition of all elements in diagonal matrix of $X'X$ with constant k estimated by Hoerl – Kennard Method (HK) and Hoerl – Kennard – Baldwin Method (HKB) and using 6 robust criteria: Huber, Ramsay, Andrews, Hample's 17A, Barya and Talwar. There are 3 independent variables with the sample sizes of 20, 50 and 100, the correlation coefficients of 0.1, 0.5 and 0.9. Two levels of outliers, mild and extreme, are considered and occurred in 1% and 5% of the sample sizes. The study is classified according to outlying – specified variables as following: variable Y only, variable X_3 only, (X_1, X_2), (Y, X_3) and (Y, X_1, X_2). The simulation is performed using Monte Carlo technique and repeated 1,000 times in each situation, The comparison is based on the mean square error (MSE).

The results indicate that under the situation of this study, the estimation of constant k in GMR by HKB mostly provides smaller MSE than that by HK and the MSEs of both methods tend to get closer as the sample sizes increase. At all serious levels and percentages of outliers, correlation coefficients and outlying – specified variables, the GMR using Ramsay's robustness, in general, provides smallest MSE when the sample size is equal to 20, while the GMR using Huber and Ramsay robustness mostly provide similar and smallest MSE when the sample sizes are equal to 50 and 100. In addition, it is also revealed that the MSE tends to be highest as the serious levels and percentages of outliers increase, but MSE tends to be lower as the sample sizes increase.