Abstract

The objective of this research was to study methods of estimating multiple regression coefficients with multicollinearity and outliers problems by comparing Ordinary Least Square method (OLS) , Ridge Least Absolute Value method (RLAV) and Generalized M – estimator Ridge method (GMR) under the criteria of Huber 's robustness (1964). The comparison of parameter efficiency was made by means of least MSE. In this study the residual distribution was assumed normally distributed with $\mu = 1$ and coefficients of variation (C.V.) of 5 % and 10 % with the number of 3 and 5 independent variables. Three levels of the strength of correlation coefficient were used : weak , moderate and strong which were identified by $\rho = 0.10$ $\rho = 0.50$ and $\rho = 0.90$, respectively. The percent of outliers were 10 % 20 % and 30 % of the sample sizes and the sample sizes were 20 and 50. Significant levels (α) used in the hypothesis testing were 0.01 and 0.05. The data were obtained by using Monte Carlo simulation technique repeating 1,000 times for each case. The results of this research are as follows :

GMR method gave least MSE for all cases except when we use the sample sizes of 20 C.V. of 10 % , the number of independent variables of 5 , high level of correlation and the percent of the outliers of 30 % of the sample sizes RLAV method also gave the least MSE. However , the least MSE varied directly with the number of independent variables , the level of correlation , the percent of the outliers and coefficients of variation , but conversed to the sample size.

Based on Bradley's test , probability of type I could be controled by OLS and RLAV methods. But GMR method generally could not control probability of type I error except when we increase the significant level , its capability to control probability of type I error for all estimation methods was improved. OLS method gave the highest power of the test for all cases. The power of the test varied directly with the number of independent variables , the level of correlation , the sample size and significant level , but conversed to the percent of the outliers and coefficients of variation.