

ROBUST JACKKNIFE RIDGE REGRESSION DENGAN ESTIMATOR LEAST ABSOLUTE DEVIATIONS (LAD) UNTUK MENGATASI MASALAH MULTIKOLINEARITAS DAN PENCILAN PADA REGRESI LINEAR BERGANDA

Oleh

Gabriella Tumiur Agustina Lumbantobing

18/430365/PA/18878

Dalam ilmu statistika, salah satu metode analisis yang paling sering digunakan adalah analisis regresi. Analisis regresi merupakan metode yang digunakan untuk mengetahui pengaruh antara variabel independen terhadap variabel dependen. Untuk mengestimasi parameter model regresi digunakan metode kuadrat terkecil. Metode ini memerlukan beberapa asumsi klasik yang harus terpenuhi, salah satunya adalah tidak terjadi multikolinearitas antar variabel independen. Jika terjadi multikolinearitas pada model regresi, penggunaan metode kuadrat terkecil akan menghasilkan estimator yang bersifat bias dan memberikan kesimpulan yang kurang tepat.

Metode *ridge regression* dapat digunakan untuk mengatasi masalah multikolinearitas. Metode ini ditempuh dengan cara menambah tetapan bias k pada diagonal utama matriks $Z'Z$. Meskipun *ridge regression* memiliki sifat yang optimal dalam menangani masalah multikolinearitas, estimatornya bersifat bias. Metode *jackknife ridge regression* dapat digunakan untuk mengurasi bias dari *ridge regression*. Akan tetapi *jackknife ridge regression* kurang tepat digunakan untuk data yang memiliki pencilan. Untuk mengatasinya dapat menggunakan *robust regression* dengan estimator *Least Absolute Deviations* (LAD). Jadi untuk mengatasi masalah multikolinearitas dan pencilan secara bersamaan digunakan metode *robust jackknife ridge regression* dengan estimator LAD, di mana *robust jackknife ridge regression* merupakan hasil pengembangan dari *robust ridge regression*.

Penelitian ini menggunakan data persentase penduduk miskin berdasarkan kabupaten/kota di Provinsi Sumatera Utara pada tahun 2019 dan faktor-faktor yang mempengaruhinya sebagai studi kasus. Berdasarkan kriteria pemilihan model terbaik, yaitu nilai *Mean Square Error* (MSE), *Akaike Information Criterion* (AIC), dan *Bayesian Information Criterion* (BIC), metode *robust jackknife ridge regression* dengan estimator LAD lebih baik dibandingkan metode *robust ridge regression* dengan estimator LAD.

Kata kunci : multikolinearitas, pencilan, *ridge regression*, *jackknife*, *robust*, estimator *Least Absolute Deviations* (LAD)

ROBUST JACKKNIFE RIDGE REGRESSION BASED ON LEAST ABSOLUTE DEVIATIONS (LAD) ESTIMATOR TO HANDLE THE PROBLEM OF MULTICOLLINEARITY AND OUTLIERS IN MULTIPLE LINEAR REGRESSIONS

by

Gabriella Tumiur Agustina Lumbantobing

18/430365/PA/18878

In statistics, one of the most frequently used analytical method is regression analysis. Regression analysis is the method that used to determine the effect of the independent variables on the dependent variable. To estimate the parameters of the regression model, the ordinary least squares method is used. This method requires several classical assumptions that must be fulfilled, one of which is that there is no multicollinearity between independent variables. If there is multicollinearity in the regression model, the use of ordinary least squares method will result in a biased estimator and give inaccurate conclusions.

The ridge regression method is used to solve the multicollinearity problem. The concept is the addition of bias constant k on the main diagonal of the $Z'Z$ matrix. Although ridge regression has optimal properties in dealing with multicollinearity problems, the estimator is biased. Jackknife ridge regression method can be used to reduce the bias of ridge regression. However, jackknife ridge regression is not appropriate for data that has outliers. To overcome this, robust regression can be used with the Least Absolute Deviations (LAD) estimator. So to overcome the problem of multicollinearity and outliers simultaneously, the robust jackknife ridge regression method is used with the LAD estimator, where robust jackknife ridge regression is the development of the robust ridge regression.

This study uses data on the percentage of poor people based on districts/cities in North Sumatra Province in 2019 and the factors that influence it as a case study. Based on the criteria for selecting the best model, such as the value of Mean Square Error (MSE), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC), the robust jackknife ridge regression method with the LAD estimator is better than the robust ridge regression method with the LAD estimator.

Keywords : multicollinearity, outliers, ridge regression, jackknife, robust, Least Absolute Deviations (LAD) estimator