

ABSTRACT

COMPARISON OF KERNEL ESTIMATOR ON OPTIMUM *BANDWIDTH* FOR MOTORIZED VEHICLE TAX (MVT) REVENUE DATA IN LAMPUNG PROVINCE

By

DONY CHERISTIAN VENESIA

Regression analysis is used to build a functional model of several variables to explain or predict a natural phenomenon based on other phenomena. The pattern of motorized vehicle tax distribution is assumed not to follow the patterns of certain distribution. Therefore, the nonparametric regression is used to estimate the distribution. In nonparametric regression, smoothing techniques are used to estimate the distribution of data, one of which is the Kernel estimator. There are four Kernel functions, namely Biweight, Triangular, Epanechnikov, and Rectangular. The most important in the Kernel estimator is to determine the optimum bandwidth based on the minimum GCV value. The results of the data analysis show that estimating motorized vehicle tax revenue data using the four Kernel estimators will produce a different optimum bandwidth. The optimum bandwidth for the Kernel Biweight estimator is at 5,1 with a MSE value of $1,640395 \times 10^{19}$, the optimum bandwidth for the Kernel Triangular estimator is at 4,8 with a MSE value of $1,536755 \times 10^{19}$, the optimum bandwidth for the Kernel Epanechnikov estimator is at a value of 4,4 with a MSE value of $1,705738 \times 10^{19}$, while the optimum bandwidth for the Rectangular Kernel estimator is at value 2 with a MSE value of $1,606819 \times 10^{19}$. From the results of the comparison above, it can be concluded that the best estimator in the motorized vehicle tax revenue data is the Triangular Kernel estimator, because it has the smallest MSE value.

Keywords: Nonparametric Regression, Kernel Estimator, Kernel Function, Bandwidth, Motorized Vehicle Tax

ABSTRAK

PERBANDINGAN ESTIMATOR KERNEL PADA *BANDWIDTH* OPTIMUM UNTUK DATA PENERIMAAN PAJAK KENDARAAN BERMOTOR (PKB) PROVINSI LAMPUNG

Oleh

DONY CHERISTIAN VENESIA

Analisis regresi merupakan ilmu statistik yang mempelajari bagaimana membangun sebuah model fungsional dari beberapa variabel sehingga dapat menjelaskan atau meramalkan suatu fenomena alami berdasarkan fenomena-fenomena yang lain. Pola sebaran data penerimaan pajak kendaraan bermotor diasumsikan tidak mengikuti pola sebaran data tertentu. Oleh karena itu, untuk mengestimasi sebaran data digunakan regresi nonparametrik. Dalam regresi nonparametrik, untuk mengestimasi sebaran data digunakan teknik *smoothing*, salah satunya yaitu estimator Kernel. Terdapat empat fungsi Kernel, yaitu *Biweight*, *Triangular*, *Epanechnikov*, dan *Rectangular*. Hal terpenting dalam estimator Kernel adalah penentuan *bandwidth* optimum didasarkan pada nilai GCV minimum. Hasil analisis data menunjukkan bahwa mengestimasi data penerimaan pajak kendaraan bermotor menggunakan keempat estimator Kernel akan menghasilkan *bandwidth* optimum yang berbeda. *Bandwidth* optimum untuk estimator *Kernel Biweight* berada pada nilai 5,1 dengan nilai MSE sebesar $1,640395 \times 10^{19}$, *bandwidth* optimum untuk estimator *Kernel Triangular* berada pada nilai 4,8 dengan nilai MSE sebesar $1,536755 \times 10^{19}$, *bandwidth* optimum untuk estimator *Kernel Epanechnikov* berada pada nilai 4,4 dengan nilai MSE sebesar $1,705738 \times 10^{19}$, sedangkan *bandwidth* optimum untuk estimator *Kernel Rectangular* berada pada nilai 2 dengan nilai MSE sebesar $1,606819 \times 10^{19}$. Dari hasil perbandingan nilai MSE untuk keempat estimator, disimpulkan bahwa estimator terbaik pada data penerimaan pajak kendaraan bermotor adalah estimator *Kernel Triangular*, karena memiliki nilai MSE terkecil.

Kata Kunci: Regresi Nonparametrik, Estimator Kernel, Fungsi Kernel, *Bandwidth*, Pajak Kendaraan Bermotor