

PERBANDINGAN HASIL MODEL REGRESI ROBUST ESTIMASI M (METHOD OF MOMENT), ESTIMASI M (MAXIMUM LIKELIHOOD TYPE), DAN ESTIMASI LTS (LEAST TRIMMED SQUARE) PADA PRODUKSI PADI DI KECAMATAN SEKARAN

Amanda Fatma Muamalah

Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Billfath, Lamongan, Indonesia

E-mail: amandafatma2233@gmail.com*

Pukky Tetralian Bintining Ngastiti ¹⁾, Ahmad Isro'il ²⁾

^{1), 2)} Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Billfath, Lamongan, Indonesia

E-mail: tetralian1010@gmail.com ¹⁾, ahmad.isroil@gmail.com ²⁾

Abstrak

Analisis regresi merupakan metode statistik yang digunakan untuk mengetahui pengaruh variabel dependen terhadap variabel independent. Tujuan dari analisis regresi adalah untuk mendapatkan model estimasi parameter model regresi dari suatu data. Salah satu metode yang digunakan untuk mengestimasi parameter regresi adalah metode MKT (Metode Kuadrat Terkecil). Metode tersebut kurang tepat digunakan pada data yang mengandung Outlier. Oleh karena itu diperlukan metode alternatif yang kekar terhadap kehadiran pencailan yaitu regresi robust, pada penelitian ini metode robust yang digunakan adalah regresi robust estimasi MM, estimasi M, dan estimasi LTS. Tujuan penelitian ini adalah membandingkan ketiga metode estimasi tersebut dan memilih model estimasi yang terbaik yang dilihat dari nilai koefisien determinasi dan mean square error. Studi kasus pada penelitian ini adalah data produksi padi di kecamatan Sekaran dengan variabel dependen produksi padi, variabel independent luas lahan, produktivitas, dan jumlah penduduk. Hasil penelitian menunjukkan bahwa regresi robust metode Least Trimmed Square (LTS) merupakan metode yang menghasilkan model terbaik, karena metode Least Trimmed Square (LTS) memiliki nilai determinasi lebih besar dan Mean Square Error (MSE) lebih kecil dibandingkan metode estimasi MM dan estimasi M.

Kata Kunci: MKT, Outlier, Regresi Robust.

Abstract

Regression analysis is a statistical method used to determine the effect of the dependent variable on the independent variable. The aim of regression analysis is to obtain an estimated model of regression model parameters from data. One of the methods used to estimate regression parameters is the MKT method (Least Squares Method). This method is not appropriate to use on data that contains outliers. Therefore, we need an alternative method that is robust to the presence of outliers, namely robust regression. In this study, the robust method used is robust regression, MM estimation, M estimation, and LTS estimation. The aim of this research is to compare the three estimation methods and select the best estimation model based on the coefficient of determination and mean square error. The case study in this research is rice production data in Sekaran sub-district with the dependent variable being rice production, the independent variables land area, productivity and population. The results of the research show that the Least Trimmed Square (LTS) robust regression method is the method that produces the best model, because the Least Trimmed Square (LTS) method has a greater determination value and a smaller Mean Square Error (MSE) compared to the MM estimation and M estimation methods.

Keyword: MKT, Outliers, Robust Regression.

PENDAHULUAN

Indonesia merupakan negara agraris yang mayoritas penduduknya bermata pencaharian sebagai petani. Kecamatan Sekaran merupakan salah satu kecamatan dari 27 kecamatan yang berada di kabupaten Lamongan. Padi merupakan tanaman

primadona di kecamatan Sekaran. Berdasarkan data yang dikumpulkan dari Badan Pusat Statistika (BPS) Lamongan, produksi padi pada tahun 2015 mencapai 41897 ton, sedangkan pada tahun 2016 jumlahnya turun mencapai 36897 ton. Pada tahun 2017 produksi padi mencapai 489,844 ton GKP dari luas panen 6,762 H_a . Naik di tahun 2021 dengan luas

panen 7,050 H_a mendapatkan hasil 489,448 ton GKP (Badan Pusat Statistik Lamongan, 2022). Laju pertumbuhan penduduk yang semakin meingkat dan berkembangnya bisnis non pertanian seperti bisnis property menyebabkan luas lahan pertanian berkurang sehingga terjadi penurunan produksi padi. Oleh karena itu, guna mencapai produksi padi yang optimal dapat digunakan analisis regresi.

Analisis regresi merupakan suatu proses statistik dalam memperkirakan hubungan antar variabel dependen dan independent (Amstrong, 2012). Model regresi linier seringkali memiliki data outlier yaitu data yang mana letaknya jauh dari pola kumpulan data. Data yang mengandung outlier melanggar asumsi kenormalan dalam regresi linier. Metode Kuadrat Terkecil (MKT) dalam analisis regresi kurang tepat dilakukan untuk data yang mengandung outlier karena akan mempengaruhi nilai koefisien regresi. Untuk menyelesaikan masalah ini diperlukan metode regresi robust. Menurut (Chen, 2002) dalam regresi robust terdapat beberapa metode estimasi diantaranya: estimasi-M, estimasi Least Median Square (LMS), estimasi Least Trimmed Square (LTS), estimasi-S, estimasi Method Of Moment (MM).

Penelitian yang dilakukan oleh Nurcahyadi (2010) memberikan kesimpulan bahwa estimasi LTS lebih baik dari pada estimasi MM. Estimasi LTS menghasilkan sisaan yang lebih kecil dari estimasi MM. Selain itu penelitian yang dilakukan oleh Pangesti (2021) mengenai perbandingan regresi ridge robust S, M, mm-estimator menyimpulkan regresi Ridge Robust S-estimator merupakan metode efektif dalam kasus faktor-faktor yang mempengaruhi kemiskinan di Jawa Tengah tahun 2020. Sedangkan penelitian mengenai perbandingan metode estimasi S dan M yang dilakukan oleh (Atamia et al., 2021) menyatakan metode estimasi M lebih baik dari estimasi S karena memiliki nilai Adj R-Square terbesar dan nilai MSE terkecil. Sehingga dari penelitian tersebut penulis termotivasi untuk memecahkan masalah mengenai model estimasi MM, estimasi M, dan estimasi LTS pada kasus produksi padi.

Berdasarkan latar belakang yang diuraikan di atas, rumusan masalah yang dapat diambil dari penelitian ini adalah bagaimana hasil perbandingan regresi robust estimasi-MM, estimasi-M, dan estimasi-LTS untuk mengestimasi produksi padi di

kecamatan Sekaran. Sehingga tujuan dari penelitian ini adalah mendapatkan hasil perbandingan regresi robust estimasi-MM, estimasi-M, dan estimasi-LTS untuk produksi padi di Kecamatan Sekaran.

KAJIAN TEORI

Produksi Padi

Padi merupakan salah satu subsektor pertanian yang paling produktif. Pada tahun 2021 provinsi Jawa Timur bertahan sebagai penghasil padi terbesar di Indonesia, dengan jumlah produksi padi sebesar 9,91 juta ton. Dari angka tersebut, kabupaten Lamongan merupakan salah satu penyumbang terbesar produksi padi di Jawa Timur. Menurut kepala dinas ketahanan pangan dan pertanian Kabupaten Lamongan, produksi padi pada tahun 2021 mengalami kenaikan sebanyak 23,345 ton atau 1,9 persen dibandingkan 2020 yang sebesar 1.172.965 juta ton GKG (Pemerintah Kabupaten Lamongan, 2022).

Analisis Regresi Linier Berganda

Regresi linier merupakan metode yang digunakan untuk menyatakan pola hubungan antar variabel respon dan variabel prediktor. Bentuk persamaan regresi linier berganda adalah sebagai berikut (Walpole & Myers, 1995):

$$Y_i = \beta_0 + \beta_1 X_{1k} + \beta_2 X_{2k} + \dots + \beta_k X_{ik} + \varepsilon_i \quad (1)$$

Dengan:

Y_i = Nilai observasi variabel respon ke- i

X_{ik} = Nilai observasi variabel prediktor ke- k

Pada pengamatan ke- i

β_0 = Nilai intersep model regresi

β_k = Koefisien regresi variabel prediktor ke- k

ε_i = Error pada pengamatan ke- i

Metode Kuadrat Terkecil (MKT)

Metode Kuadrat Terkecil merupakan salah satu metode untuk mengestimasi parameter pada regresi linier yaitu $\beta_0, \beta_1, \dots, \beta_k$. Tujuan MKT adalah meminimumkan jumlah kuadrat dari kesalahan (Supranto, 2001).

Model estimasi dari persamaan (1) sebagai berikut :

$$JKG = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2} - \dots - \beta_k x_{ik})^2 \quad (2)$$

Jika persamaan (2) diturunkan maka diperoleh persamaan sebagai berikut (Soryowati, 2016):

$$(X'X)^{-1}X'X\hat{\beta} = (X'X)^{-1}X'Y$$

$$\hat{\beta} = (X'X)^{-1}X'Y$$

Uji Asumsi Klasik

Uji asumsi klasik dilakukan untuk menunjukkan pengujian yang dilakukan telah lolos, sehingga pengujian dapat dilakukan ke analisis regresi linier.

1. Uji Normalitas

Uji normalitas bertujuan untuk mengetahui residual pada model regresi mempunyai distribusi normal atau tidak. Uji normalitas dilakukan dengan uji Kolmogorov-Smirnov. Dengan hipotesis, jika

H_0 : Residual berdistribusi normal

H_1 : residual bukan berdistribusi normal

Uji Statistik :

$$D = \max|F_0(X_i) - F_n(X_i)| \quad (3)$$

Dimana :

$F_0(X_i)$ = Fungsi distribusi frekuensi kumulatif dari distribusi normal

$F_n(X_i)$ = Distribusi frekuensi kumulatif

Pengamatan sebanyak n sampel

Apabila nilai $D > D_{tabel}$ maka asumsi kenormalan tidak terpenuhi (Semae et al, 2020).

2. Uji Heterokedastisitas

Uji heterokedastisitas bertujuan untuk menguji apakah varian dan residual dalam model regresi tidak sama antara satu pengamatan dengan pengamatan lainnya. Uji heterokedastisitas dilakukan dengan uji Breusch-Pangan. Hipotesis yang digunakan dalam uji ini yaitu:

H_0 : regresi bersifat heterokedastisitas

H_1 : regresi tidak bersifat heterokedastisitas

Statistik uji yang digunakan adalah:

$$BP = n \times R_{baru}^2 \quad (4)$$

Dimana n merupakan banyaknya pengamatan dan R_{baru}^2 adalah nilai koefisien determinasi yang diperoleh dari meregresikan variabel bebas dengan variabel terikat. Dengan daerah kritis dalam uji heterokedastisitas adalah H_0 ditolak jika nilai $BP > x_{tabel}^2$ dan nilai $p - value < \alpha$ berarti terjadi heterokedastisitas (Situmorang & Susanti, 2020).

3. Uji Multikolinearitas

Uji multikolinearitas bertujuan untuk melihat hubungan antara masing-masing variabel. Uji multikolinearitas dilihat dari nilai TOL (Tolerance) dan VIF (Variance Inflation Factor). Suatu model dikatakan

bebas dari multikolinearitas, jika mempunyai nilai TOL lebih dari 0,1 dan nilai VIF tidak lebih dari 10. Adapun nilai TOL dan VIF dapat dihitung dengan rumus sebagai berikut (Gujarati, 2004):

$$VIF = \frac{1}{1-R_j^2}; j = 1,2, \dots, k \quad (5)$$

Dan

$$TOL = \frac{1}{VIF} \quad (6)$$

Dimana R_j^2 adalah nilai koefisien determinasi yang diperoleh dari meregresikan variabel bebas x_j dengan variabel lainnya.

4. Uji Autokorelasi

Autokorelasi bertujuan mengetahui adanya korelasi antara kesalahan pengganggu pada periode sebelumnya. Pengujian yang digunakan adalah uji Durbin-Watson. Kriteria yang digunakan dalam pengujian adalah sebagai berikut (L. Febrianto et al, 2018):

Tabel 1. Kriteria Pengujian Autokorelasi

DW	Kesimpulan
$< dL$	Ada autokorelasi
$dL - dU$	Ragu-ragu
$dU - (4 - dU)$	Tidak ada autokorelasi
$(4 - dU) - (4 - dL)$	Ragu-ragu
$> (4 - dL)$	Ada autokorelasi negatif

Dimana hipotesis uji yang digunakan adalah H_0 : Tidak terjadi autokorelasi

H_1 : terjadi autokorelasi

Adapun uji statistik uji Durbin-Watson adalah:

$$DW = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2} \quad (7)$$

Data Pencilan

Data pencilan adalah pengamatan yang jauh dari pusat data yang mungkin berpengaruh besar terhadap koefisien regresi. Pengujian pencilan (outlier) dapat diketahui melalui uji DfFITS. Uji DfFITS digunakan untuk mendeteksi apakah nilai pengamatan tertentu berpengaruh terhadap nilai duga \hat{y}_i . Besarnya nilai DfFITS adalah:

$$DfFITS_i = t_i \sqrt{\frac{h_{ii}}{1-h_{ii}}} \quad (8)$$

Dengan:

$$t_i = e_i \sqrt{\frac{n-k-1}{JKG(1-h_{ii})-e_i^2}} \quad (9)$$

e_i adalah residual ke- i dan JKG adalah jumlah kuadrat galat.

Data dikatakan outlier ketika nilai $|DfFITS| > 2\sqrt{p/n}$, dimana $p = k + 1$, dan k merupakan banyaknya variabel bebas dan n banyaknya pengamatan (Soemartini, 2007).

Regresi Robust

Menurut Chen (2002), regresi robust adalah sebuah alat yang paling penting untuk menganalisis data yang mengandung pencilan. Regresi robust memiliki beberapa metode, antara lain:

1. Estimasi Method Of Moment (MM)

Estimasi MM adalah perpaduan estimasi-M dan estimasi-S. Menurut Semar dkk (2020) berikut langkah-langkah menduga parameter dengan metode estimasi-MM:

Menaksir β awal yaitu $\hat{\beta}_0$

- a. Menghitung nilai residual $e_i = y_i - \hat{y}_i$
- b. Menghitung standar deviasi $\hat{\sigma}_s$

$$\hat{\sigma}_s = \sqrt{\frac{n \sum_{i=1}^n (e_i^2) - (\sum_{i=1}^n e_i)^2}{n(n-1)}} \quad (10)$$

- c. Menghitung nilai $u_i = \frac{e_i}{\hat{\sigma}_s}$
- d. Menghitung nilai pembobot Tukey Bisquare (w_i)

$$w_i(u_i) = \frac{\psi(u_i)}{u_i} = \begin{cases} \left[1 - \left(\frac{u_i}{c}\right)^2\right]^2, & |u_i| \leq 4,685 \\ 0, & |u_i| > 4,685 \end{cases} \quad (11)$$

- e. Menghitung MKT terbobot untuk mendapatkan penaksiran kuadrat terkecil terbobot

$$\hat{\beta}_{mm} = (X'WX)^{-1}X'WY \quad (12)$$

- f. Mengulangi langkah 2-5 sampai menghasilkan nilai yang konvergen

Dimana:

w_i = Nilai pembobot data ke- i

$\hat{\beta}_{mm}$ = Koefisien parameter regresi yang Baru

c = Nilai fungsi pembobot Huber Konstanta

2. Estimasi Maximum Likelihood Type (M)

Metode penaksiran M merupakan metode penaksiran dalam regresi robust untuk mengestimasi parameter yang disebabkan adanya outlier. Tahapan estimasi parameter

regresi menggunakan estimasi M adalah sebagai berikut Azizah et al (2022):

- a. Menghitung estimasi parameter $\hat{\beta}$ menggunakan MKT sehingga didapatkan \hat{y}_i
- b. Menghitung nilai residualnya dengan $e_i = Y_i - \hat{Y}_i$
- c. Menghitung nilai $S = \frac{\text{median}|e_i - \text{median}(e_i)|}{0,6745}$ (13)

- d. Menghitung nilai $u_i = \frac{e_i}{s_i}$
- e. Menghitung nilai fungsi pembobot $w_i = w(u_i)$ dengan fungsi pembobot Huber Konstanta yang digunakan adalah $c = 1,345$

$$w(u_i) = \begin{cases} 1 & |u_i| \leq c \\ \frac{c}{|u_i|} & |u_i| > c \end{cases} \quad (14)$$

- f. Mengestimasi nilai $\hat{\beta}_m$ menggunakan MKT terboboti menggunakan pembobot w_i sehingga diperoleh estimasi M tahap satu. Pada setiap iterasi ke- t dihitung residual $e_i^{(t-1)}$ dan menggunakan pembobot $w_i^{(t-1)} = w(u_i^{(t-1)})$ dari iterasi sebelumnya sehingga didapatkan estimasi parameter $\hat{\beta}_m$ yang baru.
- g. Ulangi langkah 5 dan 6 sehingga didapatkan estimasi parameter $\hat{\beta}_m$ yang konvergen

3. Estimasi Least Trimmed Square (LTS)

Estimasi LTS merupakan regresi robust yang menggunakan konsep pengepasan metode kuadrat terkecil untuk meminimumkan jumlah kuadrat sisaan. Langkah-langkah estimasi LTS adalah sebagai berikut (Chen, 2002):

- a. Langkah-langkah estimasi parameter $\hat{\beta}$ menggunakan MKT sehingga didapatkan \hat{y}_i
- b. Menghitung nilai residualnya dengan $e_i = Y_i - \hat{Y}_i$, kemudian menghitung jumlah $h_0 = \frac{n+p+1}{2}$ observasi dengan nilai e_i^2 terkecil.
- c. Menghitung $\sum_i^h e_i^2$ dan mencari nilai estimasi parameter $\hat{\beta}_{new}$ dari $\hat{\beta}_{sebelum}$
- d. Menentukan n dari jumlah kuadrat residual $e_i = Y_i - \hat{Y}_i$ yang berkesesuaian

dengan $(\hat{\beta}_{new})$ kemudian menghitung sejumlah $\hat{\beta}_{new}$ observasi dengan e_i^2 terkecil.

- e. Dihitung $\sum_i^{h_{new}} e_i^2$
- f. Melakukan langkah 4 sampai 5 hingga konvergen

Dimana:

- n = Banyaknya pengamatan
- e_i^2 = Kuadrat residual
- β_0 = Parameter koefisien
- Y_i = variabel dependen ke- i

Pengujian Parameter Regresi

Pengujian signifikansi parameter model regresi bertujuan untuk mengetahui ada atau tidaknya pengaruh variabel bebas terhadap variabel tidak bebas, baik secara simultan maupun parsial.

1. Uji Simultan

Menurut Wardani dkk (2021) uji F digunakan untuk menunjukkan apakah semua variabel bebas memiliki pengaruh secara bersama-sama terhadap variabel terikat. Hipotesis yang digunakan untuk uji F adalah:

$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$ (tidak terdapat hubungan antara variabel terikat y dengan variabel bebas x_i secara bersama)

H_1 : minimal ada satu $\beta_i \neq 0, i = 1, 2, \dots, k$ (terdapat hubungan antara variabel terikat y dengan variabel bebas x_i secara bersama)

Statistik Uji

$$F = \frac{R^2/k}{(1-R^2)/(n-k-1)} \tag{15}$$

Kriteria Uji

H_0 ditolak jika $F_{hitung} > F_{(\alpha, k, n-k-1)}$

2. Uji Parsial

Menurut Atamia dkk (2021) uji t bertujuan menguji pengaruh masing-masing variabel bebas terhadap variabel terikat. Hipotesis yang digunakan adalah:

$H_0: \beta_1 = 0, i = 1, 2, 3, \dots, k$ (koefisien parameter variabel x_i tidak signifikan terhadap y)

$H_1: \beta_i \neq 0, i = 1, 2, 3, \dots, k$ (koefisien parameter variabel x_i signifikan terhadap y)

Statistik Uji

$$t_{hitung} = \frac{b_j - \beta_j}{s_{bj}} \tag{16}$$

Kriteria Uji

H_0 ditolak jika $|t_{hitung}| > t_{(\alpha/2, n-k-1)}$ atau $p - value < \alpha$

Keakuratan Model

Koefisien Determinasi (R^2)

Gujarati (2012) menjelaskan bahwa koefisien determinasi (R^2) adalah angka yang menunjukkan besarnya derajat kemampuan dalam menjelaskan peubah prediktor terhadap respon. Semakin besar nilai Adj-Square, maka semakin besar variasi variabel dependen dan sebaliknya. Formula dari statistik ini adalah :

$$R^2 = \frac{JKR}{JKT} \tag{17}$$

Dimana:

JKR = Jumlah Kuadrat Regresi

JKT = Jumlah Kuadrat Total

Mean Square Error (MSE)

Ada beberapa cara untuk melihat kebaikan pendugaan parameter regresi yaitu dengan melihat nilai MSE. Model persamaan yang baik adalah model regresi dengan nilai MSE kecil. Semakin kecil nilai MSE yang dihasilkan, maka semakin baik penduga parameter yang dihasilkan. Berikut perhitungan nilai MSE (Sembiring, 2003).

$$MSE = \frac{JKG}{n-k} \tag{18}$$

Dengan

$$JKG = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \tag{19}$$

Dimana:

MSE = Penyimpangan rata-rata kuadrat

JKG = Jumlah kuadrat galat

METODE

Jenis Penelitian

Jenis penelitian yang digunakan dalam penelitian ini adalah penelitian kuantitatif.

Populasi dan Sampel Penelitian

Populasi dalam penelitian ini adalah seluruh Kecamatan di Lamongan. Sedangkan sampel dalam penelitian ini yaitu Kecamatan Sekaran.

Sumber Data

Sumber data pada penelitian ini menggunakan jenis data sekunder yang diperoleh dari instansi Badan Pusat Statistik (BPS) Kabupaten Lamongan Tahun 2021.

Variabel Penelitian

Variabel-variabel yang digunakan dalam penelitian ini adalah produksi padi (Y), luas lahan (x_1), produktivitas (x_2), dan jumlah penduduk (x_3).

Teknik Analisi

Tahap analisis data yang dilakukan adalah sebagai berikut:

1. Mengumpulkan data produksi padi yang diperoleh dari Badan Pusat Statistik (BPS) Kabupaten Lamongan
2. Menentukan variabel dari data produksi padi
3. Mengestimasi parameter dari regresi linier menggunakan Metode Kuadrat Terkecil (MKT) sebagai model awal.
4. Melakukan pengujian asumsi klasik:
 - a. Melakukan uji Normalitas pada residu menggunakan uji Kolmogrov-Smirnov.
 - b. Melakukan uji Heterokedastisitas pada residu menggunakan uji Breusch-Pangan.
 - c. Melakukan uji Multikolinieritas antar variabel terikat menggunakan VIF dan TOL.
 - d. Melakukan uji Autokorelasi antar residu menggunakan uji Durbin-Watson.
5. Mendeteksi adanya pencilan dengan menggunakan metode DfFITS.
6. Regresi robust
 - a. Estimasi Method Of Moment (MM)
 - b. Estimasi Maximum Likelihood Type (M)
 - c. Estimasi Least Trimmed Square (LTS)
7. Validasi data menggunakan uji Simultan dan Parsial
8. Mendapatkan hasil metode terbaik dengan menggunakan Koefisien Determinasi (R^2) dan Mean Square Error (MSE) pada model regresi robust.

HASIL DAN PEMBAHASAN

Analisis Deskriptif

data yang digunakan dalam penelitian ini yaitu data produksi padi di kecamatan Sekaran tahun 2021. Data terdiri dari 21 desa di kecamatan Sekaran.

Tabel 2. Analisis Deskriptif

Variabel	N	Rata-rata
Produksi padi (Y)	21	23.160,72
Luas Lahan (X_1)	21	333,38

Produktivitas (X_2)	21	69,29
Jumlah Penduduk (X_3)	21	2.178,10

Berdasarkan Tabel 4.1 didapatkan hasil bahwa rata-rata produksi padi (y) sebesar 23160,72, rata-rata luas lahan (x_1), produktivitas (x_2), dan jumlah penduduk (x_3) masing-masing sebesar 333,38, 69,29, 2178,10.

Analisis Regresi Metode Kuadrat Terkecil (MKT)

Analisis regresi dengan MKT dilakukan dengan bantuan software SPSS 16 didapatkan hasil estimasi parameter produksi padi di kecamatan Sekaran sebagai berikut:

$$\hat{Y} = -21595,642 + 68,703x_1 + 311,675x_2 + 0,092x_3$$

Uji Asumsi Klasik

1. Uji Normalitas
Setelah dilakukan pengecekan menggunakan software SPSS didapatkan output yang telah disajikan dalam tabel di bawah ini:

Tabel 3. Hasil Uji Normalitas

Statistik Uji	Nilai kritis	P-value	Keputusan
0,510	0,287	0,957	Terima H_0

Berdasarkan tabel 1 dari uji Kolmogorov Smirnov diperoleh hasil nilai signifikan untuk uji distribusi normal yaitu $0,957 > 0,05$

Karena nilai signifikansi untuk uji distribusi normal lebih dari 0,05 maka dapat disimpulkan bahwa data berdistribusi normal.

2. Uji Heterokedastisitas

Tabel 4. Hasil Uji Heterokedastisitas

Model	Signifikan		Keputusan
Constant	0,423		Tolak H_0
x_1	0,423	$> 0,05$	Tolak H_0
x_2	0,433		Tolak H_0
x_3	0,231		Tolak H_0

Berdasarkan tabel 4 dapat disimpulkan bahwa asumsi terpenuhi yaitu ragam galat tidak terjadi gejala heterokedastisitas.

3. Uji Multikolinieritas

Tabel 5. Hasil Uji Multikolinieritas

Variabel Terikat	Tolerance	VIF
x_1	0,287	3,486

x_2	0,958	1,044
x_3	0,283	3,532

Dari tabel 5 dapat diambil kesimpulan bahwa semua variabel prediktor mempunyai nilai tolerance lebih dari 0,1 dan memiliki nilai VIF kurang dari 10 sehingga hasil dari uji VIF menunjukkan bahwa pengujian asumsi multikolinearitas dapat terpenuhi.

4. Uji Autokorelasi

Tabel 6. Hasil Uji Autokorelasi

Statistik Uji	dL	dU
1,636	1,0262	1,6694

Berdasarkan kriteria pengambilan keputusan pada tabel 4 nilai Durbin-Watson memenuhi kriteria $dL < d < dU$ yaitu $1,0262 < 1,636 < 1,6694$, maka dapat disimpulkan pengujian tidak menyakinkan sehingga tidak dapat disimpulkan ada atau tidaknya autokorelasi antar residual.

Pendeteksi Pencilan (Outlier)

Metode pendugaan parameter regresi dengan Metode Kuadrat Terkecil (MKT) sangat sensitif terhadap kehadiran pencilan. Pada hasil pengolahan data menggunakan DfFITS dengan ketentuan dimana suatu data dikatakan outlier ketika $|DfFITS| > 2\sqrt{\frac{p}{n}} = 2\sqrt{\frac{3}{21}} = 0,75592894603$. Nilai DfFITS diperoleh dengan bantuan SPSS 16.0 dapat disimpulkan bahwa data ke-3, ke-15, ke-16, dan ke-17 merupakan data pencilan.

Regresi Robust Estimasi MM

Karena hasil uji menyatakan ada data pencilan maka dilanjutkan dengan estimasi MM. Berdasarkan hasil estimasi menggunakan metode ini dengan dibantu software SPSS, didapatkan model regresi sebagai berikut:

$$\hat{y} = -21634,759 + 68,708x_1 + 312,179x_2 + 0,092x_3$$

Besarnya koefisien determinasi model sebesar 1,000 yang artinya sebesar 100% hasil produksi padi di kecamatan Sekaran dijelaskan oleh variabel luas lahan (x_1), produktivitas (x_2), dan jumlah penduduk (x_3). Sedangkan metode estimasi MM mempunyai nilai Mean Square Error (MSE) sebesar 6500,473 artinya kesalahan dalam memprediksi (\hat{y}) sebesar 80,626.

Regresi Robust Estimasi M

Berdasarkan hasil estimasi menggunakan pembobot Huber, didapatkan model persamaan regresi robust estimasi M sebagai berikut:

$$\hat{y} = -21733,411 + 68,719x_1 + 313,571x_2 + 0,091x_3$$

Metode estimasi M mempunyai nilai Mean Square Error (MSE) sebesar 76,481 artinya kesalahan dalam memprediksi \hat{y} sebesar 76,481. Sedangkan metode estimasi M mempunyai nilai Koefisien Determinasi (R^2) sebesar 1,000 artinya 100% hasil produksi padi di kecamatan Sekaran dijelaskan oleh variabel luas lahan (x_1), produktivitas (x_2), dan jumlah penduduk (x_3).

Regresi Robust Estimasi LTS

Berdasarkan estimasi menggunakan metode ini, didapatkan model regresi sebagai berikut:

$$\hat{y} = -19022,639 + 68,672x_1 + 274,284x_2 + 0,098x_3$$

Metode estimasi LTS mempunyai nilai Mean Square Error (MSE) sebesar 1,736 artinya kesalahan dalam memprediksi \hat{y} sebesar 1,736. Sedangkan metode estimasi LTS mempunyai nilai Koefisien Determinasi (R^2) sebesar 1,000 artinya 100% hasil produksi padi di kecamatan Sekaran dijelaskan oleh variabel luas lahan (x_1), produktivitas (x_2), dan jumlah penduduk (x_3).

Signifikansi Model

Uji Simultan

Pengujian simultan bertujuan untuk menguji semua variabel bebas model tersebut mempunyai pengaruh secara bersama terhadap variabel terikat.

Tabel 7. Hasil Uji Simultan

No	Model	F-hitung
1	Estimasi MM	2,658E5
2	Estimasi M	2,880E5
3	Estimasi LTS	2,289E5

Berdasarkan Tabel 7 $|F_{hitung}| > F_{tabel} = 3,20$ sehingga dapat disimpulkan bahwa setiap variabel bebas berpengaruh signifikan terhadap model.

Uji Parsial

Pengujian parsial berfungsi untuk mengetahui pengaruh masing-masing variabel bebas terhadap variabel terikat. Hasil perhitungan nilai t_{hitung} setiap variabel model disajikan pada Tabel 6.

Tabel 8. Hasil Uji Parsial

Parameter	Nilai Hitung		
	Estimasi MM	Estimasi M	Estimasi LTS
Constant	-9,654	-10,081	-162,116
x_1	474,252	496,899	1,269E4
x_2	9,669	10,100	163,456
x_3	3,484	3,578	82,001

Keputusan yang diambil pada uji parsial ini adalah jika $|t_{hitung}| > (t_{(2,5\%;17)} = 2,10982)$ sehingga dapat disimpulkan bahwa variabel bebas luas lahan (x_1), produktivitas (x_2), dan jumlah penduduk (x_3) berpengaruh secara parsial terhadap model.

Pemilihan Model Estimasi Terbaik

Setelah mengetahui model regresi dari masing-masing parameter, dilanjutkan dengan membandingkan nilai Koefisien Determinasi (R^2) dan nilai Mean Square Error (MSE) untuk mendapatkan model terbaik.

Tabel 9. Hasil Pemilihan Model Estimasi Terbaik

Parameter	Estimasi MM	Estimasi M	Estimasi LTS
Intercept	-21634,759	-21733,411	-19022,638
Luas lahan (x_1)	68,708	68,718	68,672
Produktivitas (x_2)	312,179	313,571	274,284
Jumlah Penduduk (x_3)	0,092	0,091	0,098
R^2	1,000	1,000	1,000
MSE	80,626	76,481	1,736

Metode terbaik adalah metode yang memiliki nilai Mean Square Error (MSE) terkecil serta nilai Koefisien Determinasi (R^2) paling besar. Berdasarkan Tabel 7 dapat dilihat bahwa nilai Koefisien Determinasi (R^2) untuk semua metode sama besar, sedangkan nilai Mean Square Error (MSE) terkecil dimiliki oleh estimasi LTS, maka dapat disimpulkan bahwa metode estimasi LTS merupakan metode estimasi robust yang paling baik digunakan dalam mengestimasi parameter regresi untuk data produksi padi di kecamatan Sekaran.

PENUTUP

SIMPULAN

Berdasarkan hasil analisis dan pembahasan maka dapat disimpulkan metode regresi robust estimasi LTS merupakan metode yang menghasilkan model terbaik dalam kasus data produksi padi di kecamatan Sekaran, karena metode estimasi LTS memiliki nilai Koefisien Determinasi (R^2) sebesar 1,000 dan nilai Mean Square Error (MSE) sebesar 1,736 yang lebih kecil dibandingkan dengan nilai MSE metode estimasi MM sebesar 80,626 dan metode estimasi M sebesar 76,481.

SARAN

Penelitian ini membahas tiga metode robust, dan metode estimasi LTS merupakan metode terbaik dibandingkan dengan metode estimasi lainnya. Sehingga harapan hasil penelitian ini dapat dijadikan acuan penelitian selanjutnya dalam melakukan analisis data yang mengandung outlier menggunakan metode regresi robust estimasi LTS.

Setelah dilakukan analisis diketahui bahwa faktor-faktor yang mempengaruhi besarnya produksi padi adalah luas lahan, produktivitas, dan jumlah penduduk. Sehingga harapan peneliti selanjutnya dapat menerapkan dan mengembangkan metode regresi robust diberbagai studi kasus yang lebih luas.

DAFTAR PUSTAKA

- Badan Pusat Statistik Lamongan. 2022. Kecamatan Sekaran Dalam Angka 2022. Lamongan: Badan Pusat Statistik.
- Amstrong, Scott J. 2012. Illusion in Regression Analysis. *International Journal Forecasting* 28: 689-693.
- Chen, C. 2002. Robust Regression and Detection With the Robustreg Procedure. *Statistics and Data Analysis*. SAS Institute: Cary, NC.
- Nurchayadi, H. 2010. Analisis Regresi pada Data Outlier dengan Metode Regresi Robust Pembobot Welsch pada Data yang Mengandung Outlier. *Jurnal Matematika UNAND* 2: 18-26.
- Aji Pangesti, A. P., Sugito, & Yasin, H. 2021. Pemodelan Regresi Ridge Robust S, M, MM-Estimator Dalam Penanganan Multikolinieritas dan Pencilan. *Jurnal Gaussian* 10: 402-412.
- Atamia, N. A., Susanti, Y., & Handajani, S. S. 2021. Perbandingan Analisis Regresi Robust

- Estimasi-S dan Estimasi-M dengan Pembobot Huber dalam Mengatasi Outlier. PRISMA, Prosiding Seminar Nasional Matematika (UNNES) 4: 673-679.
- Pemerintah Kabupaten Lamonga. 2022. Lamongan Masih Salah Satu Penyumbang Produksi Padi Terbesar di Jawa Timur. Lamongan: Diskominfo.
- Walpole, Ronald E. 1995. Pengantar Statistika. Edisi Ketiga. Jakarta: PT. Gramedia Pustaka Utama.
- Supranto, J. 2001. Statistik Teori dan Aplikasi. Edisi Keenam. Jakarta: Erlanga.
- Suryowati, K. 2016. Analisis Pseudoinvers dan Aplikasinya pada Regresi Linear Berganda. Prosiding Seminar Aplikasi Sains & Teknologi (SNAST) Periode IV. Yogyakarta.
- Semar, A., Virgantari, F., & Wijayanti, H. 2020. Perbandingan Estimasi S dan Estimasi MM pada Model Regresi Robust dengan Data Pencilan Statmat: Jurnal Statistika dan Matematika 2: 12.
- Situmorang, M. H. S., & Susanti, Y. 2020. Pemodelan Indeks Keparahan Kemiskinan di Indonesia Menggunakan Analisis Regresi Robust. Indonesia Journal Of Applied Statistics. 3: 51.
- Gujarati, D. N. 2004. Basic Econometrics Fourth edition. New York: McGraw-Hill.
- L. S. Febrianto, N. K. Dwijayanti dan P. Hendikawati. 2018. Perbandingan Metode Robust Least Median Of Square (LMS) dan Penduga S untuk Menangani Outlier pada Regresi Linier Berganda. UNNES Journal Mathematic (UJM) 3: 18-26.
- Azizah, R. J., & Wachidah, L. 2022. Regresi Robust Estimasi-M dengan Pembobot Huber dan Tukey Bisquare pada Data Tingkat Pengangguran di Indonesia Menurut Provinsi Tahun 2020. Jurnal Statistik Bandung. 2.: 18-26.
- Soemartini. 2007. Outlier (Pencilan). Bandung: UNPAD
- Semar, A., Virgantari, F., dan Wijayanti, H. 2020. Perbandingan Estimasi S dan Estimasi MM pada Model Regresi Robust dengan Data Pencilan Statmat: Jurnal Statistika dan Matematika. 2: 12.
- Azizah, R. J., & Wachidah, L. 2022. Regresi Robust Estimasi-M dengan Pembobot Huber dan Tukey Bisquare pada Data Tingkat Pengangguran di Indonesia Menurut Provinsi Tahun 2020. Bandung Conference Series: Statistics. 2: 18-26.
- Wardani, I. K., Susanti, Y., Subanti, S., Statistika, P. S., & Maret, U. S. 2021. Pemodelan Indeks Kedalaman Kemiskinan di Indonesia Menggunakan Analisis Regresi Robust. Prosiding Seminar Nasional Aplikasi Sains & Teknologi (SNAST) 2021, 15-23.
- Atamia, N. A., Susanti, Y., & Handajani, S. S. (2021). Perbandingan Analisis Regresi Robust Estimasi-S dan Estimasi-M dengan Pembobot Huber dalam Mengatasi Outlier. Prosiding Seminar Nasional Matematika. 4: 673-679
- Sembiring, R.K. 2003. Analisis Regresi. Edisi kedua. Bandung: ITB