ESTIMASI: Journal of Statistics and Its Application
Vol. 5, No. 2, Juli, 2024, Hal. 169-183
DOI: https://doi.org/10.20956/ejsa.v5i2.25520

Estimasi Parameter Regresi Ridge Robust pada Data Profil Kesehatan Sulawesi Selatan

Hendriete Tiur Marowi Waibusi[1*], Georgina Maria Tinungki[2], Sitti Sahriman[3]
[123]Departemen Statistika, Fakultas MIPA, Universitas Hasanuddin, Makassar, 90245, Indonesia
***Corresponding author, email: hendrietewaibusi@gmail.com**

Abstract

Multicollinearity is one of the assumption violations in regression analysis. The existence of multicollinearity causes the standard error to increase. Ridge regression is one of the regression analysis approaches that can overcome this problem. Besides multicollinearity, another problem that often occurs is outliers. The existence of outliers causes the data not to be normally distributed. Ridge Robust Least Trimmed Square Regression is a method that can be used to overcome multicollinearity and outlier problems in the data simultaneously in the regression analysis model. The purpose of this study was to obtain the estimation results of the least trimmed square ridge robust regression model on the Health Profile data of South Sulawesi in 2017. From the results and discussion it was found that the least trimmed square ridge robust regression method has an Rsquare value or $R2$ which is 88% and an MSE value 1.96, thus indicating that the ridge robust least trimmed square model fits better in dealing with data containing multicollinearity and outliers.

***Keywords**: Robust Ridge Regression, Least Trimmed Square, Multicollinearity, Outlier, Infant Mortality Rate.*

Abstrak

Multikolineritas adalah salah satu pelanggaran asumsi dalam analisis regresi. Keberadaan multikolineritas menyebabkan standar eror meningkat. Regresi ridge adalah salah satu pendekatan analisis regresi yang dapat mengatasi masalah tersebut. Selain multikolineritas, masalah lain yang sering terjadi adalah terdapat pencilan. Keberadaan pencilan menyebabkan data tidak berdistribusi normal. Regresi Ridge Robust Least Trimmed Square merupakan sebuah metode yang dapat digunakan untuk mengatasi masalah multikolinieritas dan pencilan pada data secara bersamaan dalam model analisis regresi. Tujuan dari penelitian ini adalah untuk mendapatkan hasil estimasi parameter model regresi ridge robust least trimmed square pada data Profil Kesehatan Sulawesi Selatan Tahun 2017. Dari hasil dan pembahasan diperoleh bahwa metode regresi ridge robust least trimmed square memiliki nilai Rsquare atau $R2$ yaitu 88% dan nilai MSE 1,96 sehingga menunjukkan kecocokan model ridge robust least trimmed square lebih baik dalam mengatasi data yang mengandung multikolinieritas dan pencilan.

Kata Kunci: Regresi Ridge Robust, Least Trimmed Square, Multikolinieritas, Pencilan, Angka Kematian Bayi.

1. Pendahuluan

Angka Kematian Bayi (AKB) merupakan indikator yang digunakan untuk melihat status kesehatan anak. Kematian bayi adalah kejadian kematian yang terjadi pada periode sejak bayi lahir sampai bayi belum berusia tepat satu tahun. Analisis regresi merupakan salah satu metode yang dapat digunakan untuk mengetahui faktor-faktor yang mempengaruhi AKB. Dalam analisis regresi, AKB digunakan sebagai respon (Y) dengan variabel prediktor (X) adalah faktor-faktor yang mempengaruhi kematian bayi. Metode kuadrat terkecil (MKT) adalah salah satu teknik estimasi parameter regresi yang umumnya digunakan yaitu untuk mengestimasi parameter model regresi Untuk mendapatkan bentuk model yang baik, data yang digunakan harus memenuhi asumsi-asumsi regresi. Adapun asumsi-asumsi regresi yang harus dipenuhi antara lain adalah data berdistribusi normal, homoskedastisitas, tidak terdapat autokorelasi dan tidak terjadi multikolinieritas [1].

Multikolinearitas merupakan masalah yang timbul berkaitan dengan adanya hubungan linear diantara variabel-variabel prediktor dalam model regresi. Multikolineritas dapat diatasi dengan menggunakan metode analisis regresi ridge. Regresi ridge diajukan sebagai suatu cara untuk mengatasi multikolineritas dengan menentukan penduga yang bias tetapi mempunyai variansi minimum. Masalah lain yang muncul dalam analisis regresi adalah terdapat pencilan pada data. Pencilan adalah salah satu masalah dalam analisis regresi linier, untuk mengatasi masalah tersebut, digunakan metode robust. Regresi robust adalah salah satu metode regresi yang digunakan untuk mengatasi pencilan yang mengakibatkan data menjadi tidak normal [2].

Untuk menangani masalah multikolineritas dan pencilan secara bersamaan maka dilakukan penggabungan metode regresi ridge dan regresi robust sehingga menjadi metode regresi ridge robust LTS. Kemudian diaplikasikan pada data Profil Kesehatan Sulawesi Selatan Tahun 2017.

2. Material dan Metode

2.1. Regresi Linier Berganda

Secara umum persamaan regresi linier berganda dengan d variabel bebas dinyatakan dengan :

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i1} + \cdots + \beta_j x_{ij} + \cdots + \beta_p x_{ip} + e_i, \qquad i = 1,2,\cdots,n.$$

atau dalam bentuk matriks dinyatakan sebagai berikut:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{e} \tag{1}$$

2.2. Metode Kuadrat Terkecil (MKT)

Metode kuadrat terkecil (MKT) merupakan salah satu metode untuk mengestimasi parameter regresi dengan cara meminimumkan jumlah kuadrat eror. Untuk mengestimasi parameter dapat dilakukan dengan pendekatan matriks sebagai berikut [3]:

$$\boldsymbol{e} = \boldsymbol{Y} - \boldsymbol{X\beta} \tag{2}$$

2.3. Uji Asumsi Klasik

Uji asumsi klasik adalah persyaratan statistik yang harus dipenuhi pada suatu analisis regresi linier berganda yang menggunakan estimasi MKT. Uji asumsi klasik tersebut antara lain, terpenuhinya distribusi normal, tidak terdapat multikolinieritas, varian konstan, dan homoskedastisitas serta tidak terdapat autokorelasi. Asumsi multikolinieritas adalah asumsi klasik yang dilanggar oleh data.

2.3.1. Uji Normalitas

Menurut [4] banyak jenis uji statistik normalitas yang dapat digunakan diantaranya Kolmogorov Smirnov, Lilliefors, Shapiro Wilk, dan Jarque Bera. Pada penelitian ini akan digunakan uji normalitas dengan Kolmogorov Smirnov. Uji Kolmogorov smirnov adalah metode statistik yang digunakan untuk menguji hipotesis komparatif dari dua sampel prediktor dengan bentuk data ordinal yang disusun pada tabel distribusi frekuensi kumulatif dengan sistem interval kelas. Hipotesis Kolmogorov Smirnov:

H_0 : data berdistribusi normal
H_1 :data tidak berdistribusi normal

Pada uji ini, data berdistribusi tidak normal jika p < 0,05 (H_0 ditolak) dan data berdistribusi normal jika p > 0,05 (H_0 diterima).

2.3.2. Uji Multikolineritas

Menurut Montgomery, salah satu ukuran yang dapat digunakan untuk menguji adanya multikolinieritas pada regresi linier berganda adalah Variance Inflation Factors (VIF). Adanya multikolinieritas dinilai dari nilai VIF yang dihasilkan. Variance Inflation Factors (VIF) dapat dihitung dengan menggunakan rumus sebagai berikut:

$$VIF = \frac{1}{1 - R_j^2} \tag{3}$$

2.4. Pencilan

Dalam statistika, data pencilan harus dilihat terhadap posisi dan sebaran data lainnya sehingga akan dievaluasi apakah data pencilan atau tidak. Salah satu metode untuk mengidentifikasi adanya pencilan dalam suatu analisis adalah Metode DfFITS (Difference in Fit Standardized). Rumus DfFITS didefinisikan sebagai berikut [5]:

$$(DfFITS) = \frac{\widehat{Y_\iota} - \hat{Y}_{i-1}}{S_{i-1}^2 - \sqrt{h_{ii}}} \tag{4}$$

2.5. Regresi Robust

Regresi robust adalah salah satu metode regresi yang digunakan ketika terdapat pencilan pada data yang mempengaruhi kenormalan data. Metode ini digunakan untuk menganalisis data yang mengandung pencilan agar model yang dihasilkan stabil terhadap pencilan. Terdapat beberapa estimasi dalam regresi robust diantaranya estimasi-M, estimasi-S, estimasi-LMS, estimasi- LTS dan estimasi-MM. [6].

2.4.1. Robust Least Trimmed Square (LTS)

Rousseeuw dan Hubert menjelaskan bahwa metode LTS mempunyai prinsip yang sama dengan MKT dalam mengestimasi parameter regresi yaitu meminimumkan jumlah kuadrat eror Fungsi tujuan dari metode LTS dapat dituliskan sebagai berikut [7]:

$$\min \sum_{i=1}^{h} e_{(i)}^{2} = \min \sum_{i=1}^{h} (Y_{(i)} - \hat{Y}_{(i)})^{2} \tag{5}$$

2.6. Regresi Ridge

Penduga regresi *ridge* diperoleh dari penduga parameter MKT ditambahkan dengan faktor koreksi, yaitu konstanta (c) yang dikalikan dengan matriks identitas sehingga $(\boldsymbol{X'X})^{-\mathbf{1}}$ menjadi $(\boldsymbol{X'X} + c\boldsymbol{I})^{-\mathbf{1}}$. Penambahan faktor koreksi tersebut bertujuan agar determinan matriks $\boldsymbol{X'X}$ bisa dihitung. Persamaan regresi *ridge* dapat ditulis sebagai berikut:

$$\hat{\boldsymbol{\beta}}_{Ridge} = (\boldsymbol{X'X} + c\boldsymbol{I})^{-\mathbf{1}} \boldsymbol{X'y} \tag{6}$$

2.7. Regresi Ridge Robust

Metode regresi ridge tidak dapat menangani masalah pencilan dan multikolineritas secara bersamaan maka diberikan rumus untuk menangani masalah tersebut dengan mengkombinasikan regresi ridge dan robust dengan cara menggabungkan kedua sifat. Penggabungan kedua metode tersebut disebut regresi ridge robust. Penduga regresi ridge robust yang dihasilkan akan stabil terhadap pencilan [8].

Rumus hasil penggabungan regresi ridge dan regresi robust menurut [8] adalah sebagai berikut:

$$\hat{\boldsymbol{\beta}}_{RR} = (\boldsymbol{X'X} + c\boldsymbol{I})^{-\mathbf{1}} \boldsymbol{X'X} \hat{\boldsymbol{\beta}}_{LTS} \tag{7}$$

2.8. Koefisien Determinasi

Koefisien determinasi adalah salah satu ukuran tentang kecocokan data dengan model. Rumus koefisien determinasi adalah sebagai berikut:

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{y} - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \hat{y})^2} \tag{8}$$

2.8.1. MSE (Mean Square Eror)

MSE adalah rata-rata kesalahan kuadrat diantara nilai aktual dan nilai peramalan. MSE secara umum digunakan untuk mengukur ketepatan nilai dugaan model yang dinyatakan dalam rata-rata kuadrat dari kesalahan. Secara matematis dapat ditulis pada 2.16 [3]:

$$MSE = \frac{\text{SSE}}{n - p - 1} \tag{9}$$

2.9. Uji Signifikansi Parameter

Uji signifikansi parameter digunakan untuk mengetahui pengaruh dari variabel-variabel prediktor terhadap variabel respon baik secara simultan maupun secara parsial.

2.9.1. Pengujian Simultan

Pengujian simultan parameter dimaksudkan untuk mengetahui apakah terdapat pengaruh secara bersama-sama dari variabel prediktor terhadap variabel respon. Sehingga untuk menunjukkanya dilakukan pengujian hipotesis berikut [9]:

Hipotesis:

H_0 : $\beta_1 = \beta_2 = \cdots = \beta_j = \cdots = \beta_p = 0,$
H_1 : minimal ada satu $\beta_j \neq 0,\ j = 1,2,\ldots,p$

Maka statistik uji F yaitu:

$$F = \frac{MSR}{MSE} = \frac{\sum_{i=1}^{n}(\hat{y} - \bar{y})^2 / k}{\sum_{i=1}^{n}(y_i - \hat{y})^2 / n - p - 1} \tag{10}$$

Kriteria pengambilan keputusan:

Jika $F_{hitung} > F_{tabel}$, maka H_0 ditolak

Jika $F_{hitung} \leq F_{tabel}$, maka H_0 diterima

2.9.2. Pengujian Parsial

Uji parsial digunakan untuk melakukan uji hipotesis pada setiap parameter. Adapun hipotesis dalam uji ini sebagai berikut:

H_0: $\beta_j = 0$, (tidak ada hubungan yang signifikan) , $j = 1,2,\dots,p$.
H_1: $\beta_j \neq 0$, (paling tidak terdapat satu hubungan yang signifikan)

Dengan statistik uji yang digunakan adalah:

$$t_{hit(j)} = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \tag{11}$$

2.10. Angka Kematian Bayi

Angka Kematian Bayi (AKB) merupakan salah satu indikator penting dalam menentukan tingkat kesehatan masyarakat karena dapat menggambarkan kesehatan penduduk secara umum AKB dapat dihitung dengan menggunakan rumus sebagai berikut:

$$AKB = \frac{D_{0-<1th}}{lahir\ hidup} \times 1000 \tag{12}$$

dengan:

AKB : Angka Kematian Bayi
$D_{0-<1th}$: Jumlah kematian bayi berumur kurang dari 1 tahun pada daerah tertentu

2.11. Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder, bersumber dari Laporan Profil Kesehatan Provinsi Sulawesi Selatan tahun 2017 yang dipublikasikan oleh Dinas Kesehatan Provinsi Sulawesi Selatan. Data tersebut diakses pada www.ppid.sulselprov.go.id.

2.12. Metode Analisis

Adapun tahapan metode analisis yang digunakan pada penelitian ini:

1. Melakukan estimasi koefisien regresi menggunakan MKT.
2. Melakukan analisis deskriptif data.
3. Melakukan uji multikolinearitas dengan melihat nilai VIF. Jika nilai VIF lebih besar dari 10 maka data mengandung multikolineritas.
4. Melakukan pendeteksian pencilan dengan melihat nilai $DfFITS$. Jika nilai DfFITS $> 2\sqrt{\frac{7}{24}}$ maka data mengandung pencilan.
5. Menghitung nilai $\widehat{\boldsymbol{\beta}}_{LTS}$ pada metode regresi *robust LTS*

$$\widehat{\boldsymbol{\beta}}_{LTS} = min_{\beta} \sum_{i}^{h} {e_i}^2$$

Prosedur estimasi dengan menggunakan estimasi LTS adalah sebagai berikut:

a) Mengestimasi koefisien regresi menggunakan MKT.
b) Menentukan n eror ${e_i}^2 = \left(\hat{Y}_i - X_i\, \widehat{\boldsymbol{\beta}}_0\right)^2$ yang bersesuaian dengan $(\widehat{\boldsymbol{\beta}}_0)$, kemudian menghitung jumlah $h_0 = (n + p + 2)/\,2$ pengamatan dengan nilai ${e_i}^2$ terkecil.
c) Menghitung $\sum_{i}^{h_0} {e_i}^2$,
d) Mengestimasi parameter $\widehat{\boldsymbol{\beta}}_{baru}$ dari $\widehat{\boldsymbol{\beta}}_0$ observasi,
e) Ditentukan n kuadrat eror ${e_i}^2 = \left(\hat{Y}_i - X_i\, \widehat{\boldsymbol{\beta}}_0\right)^2$ yang bersesuaian dengan $\widehat{\boldsymbol{\beta}}_{baru}$ kemudian observasi dengan ${e_i}^2$ terkecil,
f) Menghitung $\sum_{i}^{h_{baru}} {e_i}^2$,
g) Melakukan iterasi yaitu tahap (d) sampai (f) untuk mendapatkan fungsi yang konvergen.

6. Mengestimasikan $\boldsymbol{\beta}_R$ dengan menggunakan regresi *ridge*
7.

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{Ridge}} = (\boldsymbol{X}'\boldsymbol{X} + c\boldsymbol{I})^{-1}\, \boldsymbol{X}'\boldsymbol{y}$$

8. Mengestimasikan $\boldsymbol{\beta}_R$ dengan $\widehat{\boldsymbol{\beta}}_{LTS}$

$$\widehat{\boldsymbol{\beta}}_{RR} = (\boldsymbol{X}'\boldsymbol{X} + c\boldsymbol{I})^{-1}\, \boldsymbol{X}'\boldsymbol{y}$$

dengan $\boldsymbol{y} = \boldsymbol{X}\widehat{\boldsymbol{\beta}}_{LTS}$,
Sehingga

$$\widehat{\boldsymbol{\beta}}_{RR} = (\boldsymbol{X}'\boldsymbol{X} + c\boldsymbol{I})^{-1}\, \boldsymbol{X}'\boldsymbol{X}\widehat{\boldsymbol{\beta}}_{LTS}$$

9. Melakukan uji F dan uji t koefisien regresi *robust* dan menghitung R^2 untuk model regresi *ridge robust*-LTS
10. Menganalisi model regresi *ridge robust* terbaik berdasarkan R^2 dan MSE

3. Hasil dan Diskusi

3.1. Estimasi Parameter Regresi Ridge Robust-LTS

Regresi *ridge robust-LTS* merupakan metode kombinasi yang menggabungkan antara metode regresi ridge dengan regresi robust LTS. Askin dan Montgomery (1980) menyatakan bahwa, estimator ridge robust merupakan solusi untuk mengatasi masalah multikolinieritas dan pencilan pada data.

Least Trimmed Square (LTS) atau kuadrat terpangkas terkecil adalah salah satu teknik statistik yang digunakan untuk mengestimasi parameter dari model regresi linear yang memberikan alternatif robust ke metode regresi untuk meminimalkan jumlah kuadrat eror. Adapun penduga untuk menentukan nilai estimasi robust-LTS.

$$\min \sum_{i=1}^{h} e_{(i)}^2 = \min \sum_{i=1}^{h} (Y_{(i)} - \hat{Y}_{(i)})^2 \tag{13}$$

Persamaan metode ini dapat dilihat sebagai berikut:

$$\widehat{\boldsymbol{\beta}}_{LTS} = min_{\beta} \sum_{i}^{h} {e_i}^2 \tag{14}$$

Fungsi pembobot diatas pembobotnya jika nilai r = 3 sebagai berikut:

$$W_{ii} = \begin{cases} 1 & jika \left| \frac{e_i}{S_{LTS}} \right| \le 3 \\ 0 & untuk\ lainnya \end{cases} \tag{15}$$

Dapat diselesaikan dengan menggunakan metode *IRLS* (*Iteratively Reweighted Least Squares*), maka:

$$\sum_{i=1}^{n} X_{ij} w_i y_i - \sum_{i=1}^{n} X_{ij} w_i \sum_{j=0}^{d} X_{ij} \beta_j = 0 \tag{16}$$

Persamaan (4.3) dapat dinotasikan dalam bentuk matriks, yaitu:

$$\boldsymbol{X'WY} - \boldsymbol{X'WX}\widehat{\boldsymbol{\beta}} = \boldsymbol{0}$$

$$\boldsymbol{X'WY} = \boldsymbol{X'WX}\widehat{\boldsymbol{\beta}}$$

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X'WX})^{-1}\boldsymbol{X'WX}$$

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{R}} = (\boldsymbol{X'X} + c\boldsymbol{I})^{-1}\boldsymbol{X'y} \tag{17}$$

Regresi *Ridge* merupakan salah satu metode yang digunakan untuk mengatasi masalah multikolinieritas melalui modifikasi terhadap metode kuadrat terkecil. Modifikasi tersebut ditempuh dengan cara menambah tetapan bias c yang relatif kecil pada diagonal utama matriks $\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X}$, sehingga koefisien estimator *Ridge* dipenuhi dengan besarnya tetapan bias c. Sehingga parameter penduganya menjadi:

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{R}} = (\boldsymbol{X}'\boldsymbol{X} + c\boldsymbol{I})^{-\boldsymbol{1}}\boldsymbol{X}'\boldsymbol{y} \tag{18}$$

Ridge Robust-LTS merupakan gabungan dari metode *Robust* dan metode *Ridge*. Untuk mendapatkan estimasi y yaitu $\widehat{\boldsymbol{y}} = \boldsymbol{X}\widehat{\boldsymbol{\beta}}_{\boldsymbol{LTS}}$, sehingga Persamaan diatas menjadi:

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{RR}} = (\boldsymbol{X}'\boldsymbol{X} + c\,\boldsymbol{I})^{-\boldsymbol{1}}\boldsymbol{X}'\boldsymbol{X}\widehat{\boldsymbol{\beta}}_{\boldsymbol{LTS}} \tag{19}$$

Dimana c adalah nilai konstan yang merupakan bilangan positif atau $c \geq 0$ umumnya c terletak antara interval $0 < c < 1$.

Cara untuk memilih kontanta bias (c) yaitu dengan menggunakan rumus HKB yang diperkenalkan oleh Hoerl, Kennard dan Balwin :

$$c = \frac{p + 1(\sigma_{LTS}^2)}{\beta_{LTS}'\,\beta_{LTS}}$$

Dimana $\hat{\beta}$ dan $\hat{\sigma}$ diperoleh dari *Robust*-LTS, dengan p adalah banyaknya parameter regresi yang diestimasi dan $\hat{\sigma}^2$ adalah:

$$\sigma_{LTS}^2 = \frac{\left(Y - X\widehat{\boldsymbol{\beta}}_{\boldsymbol{LTS}}\right)'(Y - X\widehat{\boldsymbol{\beta}}_{\boldsymbol{LTS}})}{n - (p + 1)}$$

3.2. Statistik Deskriptif

Statistik deskriptif digunakan untuk mengetahui rata-rata (*mean*), variansi *(variance)* dan standar deviasi *(std. deviation)* pada penelitian ini. Statistik deskriptif terhadap variabel respon (Y) dan variabel prediktor (X):

Table 1. Hasil Deskriptif

Variabel	Rata-rata	Std. Deviasi	Variansi
Y	8,375	4,424	19,574
X_1	709,605	1897,540	3600695,842
X_2	55,966	14,319	205,054
X_3	227,751	154,566	23890,741
X_4	2830,985	3690,353	13618706,561
X_5	6234,916	5037,141	25372786,781
X_6	5725,542	4850,427	23526643,741

Sumber: Data diolah 2022

Berdasarkan Tabel 4.1 nilai rata-rata tertinggi pada variabel (X_5). Nilai standar deviasi tertinggi pada variabel (X_5). Nilai varinasi tertinggi pada variabel (X_5).

3.3. Uji Normalitas

Model yang baik adalah model yang berdistribusi normal. Uji normalitas yang digunakan dalam penelitian ini menggunakan uji *kolmogorov smirnov* dengan hipotesis sebagai berikut:

H_0 : data berdistribusi normal
H_1 : data tidak berdistribusi normal

Table 2. Uji Normalitas

D	p-value
0,0855	0,921

Sumber: Data diolah 2022

Berdasarkan Tabel 4.2 nilai *p-value* sebesar 0,921 $\geq$ 0,05 sehingga H_0 diterima artinya data berdistribusi normal.

3.4. Uji Multikolinearitas

Uji multikolinearitas dimaksudkan untuk menguji apakah dalam model regresi ditemukan adanya korelasi antar variabel prediktor. Model regresi yang baik seharusnya tidak terjadi korelasi diantara variabel prediktor. Untuk mengetahui adanya multikolinieritas maka dilakukan metode VIF (Varian Inflation Factor). Tabel 4.3 merupakan hasil VIF variabel prediktor.

Table 3. Uji VIF

Variabel	Nilai VIF
X_1	5,4671
X_2	1,3672
X_3	7,5533
X_4	15,8415
X_5	36,9474
X_6	26,9423

Sumber: Data diolah 2022

Berdasarkan Tabel 4.3 dapat dilihat bahwa variabel X_4 sampai X_6 memiliki nilai *VIF* yang lebih besar dari 10 yaitu X_4 15,8415 $\geq$ 10, X_5 36,9474 $\geq$ 10, X_6 26,9423 $\geq$ 10, sehingga dapat dikatakan bahwa data mengandung masalah multikolinearitas.

3.5. Uji DfFITS

Uji *DfFITS* dimaksudkan untuk menguji apakah dalam data tersebut ditemukan adanya pencilan. Keberadaan data pencilan akan mengganggu proses analisa data sehingga harus dihindari. Data dapat diindikasikan sebagai pencilan jika nilai *DfFITS* lebih dari $2\sqrt{\frac{7}{24}}$. Berdasarkan hasil perhitungan, data yang terdeteksi pencila yaitu data ke 11 dan 23.

3.6. Estimasi Parameter Model Regresi Robust Dengan Estimasi-LTS

Robust least trimmed square mempunyai prinsip yang sama dengan MKT dalam mengestimasi parameter regresi yaitu meminimumkan jumlah kuadrat eror. Adapun algoritma untuk mencari nilai dari estimasi robust least trimmed square:

1. Sebelum mencari nilai dari estimasi *robust least trimmed square*, terlebih dahulu menaksir parameter regresi (β) menggunakan metode MKT. Berikut merupakan parameter (β) dari metode MKT.

Table 4. Parameter $(\boldsymbol{\beta})$ dari metode MKT

Parameter	Penduga
β_0	13,15227
β_1	-0,00125
β_2	-0,07170
β_3	0,01661
β_4	0,00132
β_5	-0,00208
β_6	0,00097

Sumber: Data diolah, 2022

2. Menentukan eror yang bersesuaian dengan $(\widehat{\boldsymbol{\beta}}_0)$ dari metode MKT kemudian menentukan $h_0 = (24 + 6 + 2)/\ 2$ pengamatan dengan nilai $e_i{}^2$ terkecil
3. Menghitung $\sum_i^{h_0} e_i{}^2$,
4. Mengestimasi parameter $\widehat{\boldsymbol{\beta}}_{baru}$ dari $\widehat{\boldsymbol{\beta}}_0$ observasi
5. Ditentukan kuadrat eror $e_i{}^2 = \left(\hat{Y}_i - X_i\,\widehat{\boldsymbol{\beta}}_0\right)^2$ yang bersesuaian dengan $\widehat{\boldsymbol{\beta}}_{baru}$ kemudian observasi dengan $e_i{}^2$ terkecil,
6. Menghitung $\sum_i^{h_{baru}} e_i{}^2$,
7. Melakukan iterasi yaitu tahap (4) sampai (6) untuk mendapatkan fungsi yang konvergen.

Berdasarkan perhitungan, maka didapatkan hasilestimasi parameter-parameter yang ditunjukkan seperti pada Tabel 4.5 berikut:

Table 5. Iterasi Parameter *Robust*

Parameter	$\hat{\boldsymbol{\beta}}_{LTS}$
$\hat{\beta}_0$	13,15227
$\hat{\beta}_1$	-0,00125
$\hat{\beta}_2$	-0,07170
$\hat{\beta}_3$	0,01661
$\hat{\beta}_4$	0,00132
$\hat{\beta}_5$	-0,00208
$\hat{\beta}_6$	0,00098

Sumber: Data diolah, 2022

Setelah didapatkan bentuk model persamaan dengan menggunakan estimasi parameter regresi *robust least trimmed square* maka selanjutnya dilakukan estimasi parameter regresi *ridge robust*-LTS.

Hasil estimasi parameter pada Tabel 4.5 digunakan untuk menentukan variansi untuk memperoleh nilai tetapan bias dan matriks identitas ***(I)*** berukuran $n \times n$, dimana n = 24. Berdasarkan hasil dan perhitungan, nilai tetapan bias adalah c = 0.609.

Table 6. Parameter Regresi *Ridge Robust Least Trimmed Square*

Parameter	Penduga
$\hat{\beta}_0$	8,39963
$\hat{\beta}_1$	-0,00184
$\hat{\beta}_2$	0,00028
$\hat{\beta}_3$	0,01521
$\hat{\beta}_4$	0,00161
$\hat{\beta}_5$	-0,00232
$\hat{\beta}_6$	0,00130

Sumber: Data diolah, 2022

Berdasarkan Tabel 4.6 dapat diketahui bahwa model regresi yang dapat dibentuk untuk estimator regresi *ridge robust* LTS yaitu:

$$\hat{y} = 8{,}39963 - 0{,}00184X_1 + 0{,}00028X_2 + 0{,}01521X_3 + 0{,}00161X_4 - 0{,}00232X_5 + 0{,}00130X_6$$

3.7. Uji Signifikansi Parameter

3.7.1. Uji Simultan

Table 7. Uji Simultan

F_{hitung}	F_{tabel}
39	2,23

Berdasarkan Tabel 4.13 dapat dilihat bahwa F_{hitung} lebih besar dari F_{tabel} maka H_0 ditolak, seluruh variabel prediktor berpengaruh signifikan terhadap variabel respon.

3.7.2. Uji Parsial

Pengaruh variabel prediktor secara parsial atau individu dapat dihitung dengan melakukan uji t. Hasil uji pengaruh masing-masing variabel prediktor secara parsial dengan $\alpha = 0{,}05$ dapat dilihat pada Tabel 4.14 sebagai berikut:

Table 8. Uji Parsial Satu-Satu

Variabel	t_{hitung}	t_{tabel}	Keputusan
X_1	6,68496	2,000	Signifikan
X_2	1,07153		Tidak Signifikan
X_3	1,92859		Tidak Signifikan
X_4	10,43877		Signifikan
X_5	9,76970		Signifikan
X_6	5,03784		Signifikan

Sumber: Data diolah, 2022

Berdasarkan Tabel 4.14 nilai t_{hitung} dari (X_1), (X_4), (X_5), (X_6) lebih besar dari nilai t_{tabel} sehingga dapat disimpulkan secara parsial berpengaruh signifikan terhadap angka kematian bayi (Y).

Berdasarkan hasil pengujian parsial, ada 4 variabel (X) yang signifikan terhadap variabel (Y) diantaranya (X_1), (X_4), (X_5), (X_6). Maka dilakukan pengujian ulang menggunakan metode regresi *ridge robust*-LTS dengan variabel (X) yang signifikan terhadap variabel (Y). Berdasarkan hasil perhitungan, koefisien parameter *ridge robust*-LTS dapat dilihat pada Tabel 4.15:

Table 9. Parameter *Ridge Robust Least Trimmed Square*

Parameter	Penduga
$\hat{\beta}_0$	9,93504

$\hat{\beta}_1$	-0,00236
$\hat{\beta}_4$	0,00309
$\hat{\beta}_5$	-0,00313
$\hat{\beta}_6$	0,00181

Sumber: Data diolah, 2022

Berdasarkan Tabel 4.15 dapat diketahui bahwa model regresi yang dapat dibentuk untuk estimator regresi *ridge robust*-LTS yaitu:

$$\hat{y}_{n=100} = 9{,}93504 - 0{,}00236X_1 + 0{,}00309\,X_4 - 0{,}00313\,X_5 + 0{,}00181X_6$$

3.7.3. Uji Kebaikan Model

Kebaikan suatu model dapat dilihat atau diukur dari nilai koefisien determinasi (R^2) dan nilai MSE suatu data. Semakin baik nilai R^2 berarti semakin baik model prediksi dari model penelitian yang diajukan. MSE merupakan salah satu ukuran yang sering digunakan untuk menguji kualitas suatu model pada regresi, ketepatan nilai dugaan model dinyatakan dalam rata-rata kuadrat dari kesalahan.

Table 10. Uji Kebaikan Model

Model	R^2	MSE
Regresi *Ridge Robust Least Trimmed Square*	88 %	1,96

Sumber: Data diolah, 2022

Tabel 4.16 menjelaskan bahwa metode regresi *ridge robust*-LTS memiliki nilai R^2 88% MSE yang lebih kecil 1,96 terlihat bahwa nilai tersebut sangat baik karena nilai R^2 cukup tinggi mendekati 1, sehingga dapat disimpulkan bahwa metode regresi *ridge robust least trimmed square* sangat baik digunakan dalam mengatasi masalah multikolineritas dan pencilan.

4. Kesimpulan

Berdasarkan penelitan yang telah dilakukan maka didapatkan beberapa kesimpulan sebagai berikut :

1. Model yang diperoleh untuk data angka kematian bayi di Sulawesi Selatan Tahun 2017 adalah:
   $$\hat{y} = 9{,}93504 - 0{,}00236X_1 + 0{,}00309\,X_4 - 0{,}00313\,X_5 + 0{,}00181\,X_6$$
2. Berdasarkan analisis pembahasan secara parsial dapat disimpulkan bahwa faktor-faktor yang mempengaruhi Angka Kematian Bayi di Provinsi Sulawesi Selatan Tahun 2017 adalah (X_1), (X_4), (X_5), (X_6) dengan nilai R^2 88% dan nilai MSE 1,96.

Daftar Pustaka

[1] Supranto, J. *Ekonometri*. Bogor: Ghalia Indonesia. 2005.

[2] Cahyawati, D. Efektifitas Metode Regresi Robust Penduga Welsch dalam Mengatasi Pencilan pada Pemodelan Regresi Linear Berganda. Jurnal Penelitian Sains, 12(1). Unsri, 2009.

[3] Montgomery, D. C., & Peck, E. A. *Introduction to Linear Regression Analysis (2nd ed)*. New York: John Wiley & SonsInc. 1992.

[4] Jarque, C. M., & Bera, A. K. A Test for Normality of Observations and Regression Residuals. *International Statistical Review*, 163-172, 1987.

[5] Soemartini. *Pencilan (Outlier)*. Universitas Padjadjaran, Bandung melalui http://resources.unpad.ac.id/unpad-content/uploads/publikasidosen/. Outlier(Pencilan).pdf. 2007.

[6] Chen, C. *Robust Regression and Outlier Detection with the ROBUSTREG Procedure*. Cary NC: SAS Institute Inc. 2002.

[7] Rousseuw, P. J. & Leroy, A. M. *Robust Regression* and *Outlier Detection*. New York: John Wiley & Sons, Inc. 1987.

[8] Samkar, H. & Alpu, O. *Ridge Regression Based on Some Robust Estimators. Journal of Modern Applied Statistical Methods*. 9, 45-501, 2010.

[9] Gujarati, D. *Ekonometrika Dasar Terjemahan Sumarno Zain*. Jakarta:Erlangga. 2003.