ISSN: 2339-2541
JURNAL GAUSSIAN, Volume 8, Nomor 1, Tahun 2019, Halaman 24 - 34
Online di: http://ejournal-s1.undip.ac.id/index.php/gaussian

PEMODELAN REGRESI *RIDGE ROBUST-MM* DALAM PENANGANAN MULTIKOLINIERITAS DAN PENCILAN (Studi Kasus : Faktor-Faktor yang Mempengaruhi AKB di Jawa Tengah Tahun 2017)

Eka Destiyani[1], Rita Rahmawati[2], Suparti[3]
[1,2,3] Departemen Statistika FSM Universitas Diponegoro
Email: ritarahmawati@gmail.com

ABSTRACT

The Ordinary Least Squares (OLS) is one of the most commonly used method to estimate linear regression parameters. If multicollinearity is exist within predictor variables especially coupled with the outliers, then regression analysis with OLS is no longer used. One method that can be used to solve a multicollinearity and outliers problems is Ridge Robust-MM Regression. Ridge Robust-MM Regression is a modification of the Ridge Regression method based on the MM-estimator of Robust Regression. The case study in this research is AKB in Central Java 2017 influenced by population dencity, the precentage of households behaving in a clean and healthy life, the number of low-weighted baby born, the number of babies who are given exclusive breastfeeding, the number of babies that receiving a neonatal visit once, and the number of babies who get health services. The result of estimation using OLS show that there is violation of multicollinearity and also the presence of outliers. Applied ridge robust-MM regression to case study proves ridge robust regression can improve parameter estimation. Based on t test at 5% significance level most of predictor variables have significant effect to variable AKB. The influence value of predictor variables to AKB is 47.68% and MSE value is 0.01538.

Keywords: Ordinary Least Squares (OLS), Multicollinearity, Outliers, Ridge Regression, Robust Regression, AKB.

1. PENDAHULUAN

Angka Kematian Bayi (AKB) dapat didefinisikan sebagai jumlah kematian bayi (0-11 bulan) per 1.000 kelahiran hidup dalam kurun waktu satu tahun. Angka kematian bayi di Provinsi Jawa Tengah tahun 2017 mencapai 8,9 per 1.000 kelahiran hidup (Dinkes Jateng, 2018). Provinsi Jawa Tengah merupakan salah satu provinsi yang menjadi target pelaksanaan Program EMAS (*Exanding Maternal and Newborn Survival*) yaitu program peningkatan kesehatan ibu dan anak baru lahir. Angka Kematian Bayi (AKB) menggambarkan tingkat permasalahan kesehatan masyarakat yang berkaitan dengan faktor-faktor penyebab kematian bayi. Dengan mengetahui faktor mana yang memiliki pengaruh terhadap tingginya AKB, akan dapat dilakukan pengendalian terhadap faktor-faktor tersebut yang kemudian dapat ditentukan upaya apa saja yang dapat dilakukan untuk menurunkan AKB.

Analisis regresi merupakan salah satu metode yang dapat digunakan untuk mengetahui faktor-faktor yang mempengaruhi AKB. Salah satu cara untuk mendapatkan koefisien regresi adalah melalui Metode Kuadrat Terkecil (MKT). Sebuah model regresi dikatakan baik apabila memenuhi asumsi klasik yaitu *error* menyebar normal, ragam dari *error* bersifat homogen, *error* tidak mengalami autokolerasi, dan tidak terjadi multikolinieritas antar variabel bebas.. Jika terdapat pelanggaran terhadap asumsi, estimator yang diperoleh bersifat bias dan tidak efisien sehingga model regresi yang diperoleh tidak cocok terhadap data yang dimodelkan (Hoerl dan Kennard, 1970).

Pelanggaran asumsi model regresi linier klasik yang sering terjadi adalah adanya multikolinieritas yang berarti adanya hubungan linier di antara variabel bebas dalam model regresi. Salah satu metode yang dapat digunakan untuk mengatasi adanya multikolinieritas adalah regresi *ridge* (Montgomery dan Peck, 1992). Penyebab utama pelanggaran asumsi tersebut adalah terdapatnya beberapa pencilan, sebuah pencilan merupakan data yang tidak mengikuti pola umum pada model atau yang keluar pada model. Metode yang dapat digunakan untuk menganalisis data berpencilan adalah regresi r*obust* (Sembiring, 2003).

Studi awal penelitian menggunakan regresi linier berganda dengan estimasi parameter Metode Kuadrat Terkecil (MKT), diketahui pada data AKB terjadi pelanggaran asumsi multikolinieritas dan terdapat beberapa pencilan. Oleh karena itu diperlukan suatu metode untuk mendapatkan nilai dugaan parameter yang resisten terhadap pencilan dan menangani masalah multikolinieritas, metode tersebut adalah metode Regresi *Ridge Robust-MM*. Metode *Regresi Ridge Robust-MM* diawali dengan mencari pembobot dan nilai dugaan parameter pada regresi *robus*t dengan *MM-estimator*, kemudian nilai dugaan parameter yang diperoleh digunakan untuk mencari nilai dugaan parameter dari metode Regresi *Ridge Robust-MM*. Dari analisis kemudian dilakukan pembandingan model regresi MKT dan *Ridge Robust-MM* untuk mendapatkan model regresi terbaik. *Software* yang digunakan untuk analisis dalam penelitian ini adalah *software R.*

2. TINJAUAN PUSTAKA

2.1. Analisis Regresi

Menurut Montgomery dan Peck (1992), analisis regresi merupakan salah satu teknik statistika yang sering digunakan untuk mengetahui hubungan antara variabel bebas dengan variabel terikat, model regresi berganda dinyatakan dengan k variabel bebas dan n pengamatan dapat dituliskan sebagai berikut:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \varepsilon_i \ ; \ i = 1, 2, \dots, n$$

Estimasi parameter model regresi linier berganda didapat dengan meminimumkan jumlah kuadrat error atau sering dikenal Metode Kuadrat Terkecil (MKT) Montgomery dan Peck (1992).

$$\boldsymbol{L} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{\varepsilon_i}^{\boldsymbol{2}} = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = (\boldsymbol{Y} - \boldsymbol{X\beta})'(\boldsymbol{Y} - \boldsymbol{X\beta})$$

estimator kuadrat terkecil $(\widehat{\boldsymbol{\beta}})$ yang meminimumkan $\boldsymbol{L}$ disyaratkan bahwa $\frac{\partial \boldsymbol{L}}{\partial \boldsymbol{\beta}}|_{\boldsymbol{\beta}=\widehat{\boldsymbol{\beta}}} = 0$. Turunan pertama dari $\boldsymbol{L}$ terhadap $\widehat{\boldsymbol{\beta}}$ adalah :

$$\frac{\partial \boldsymbol{L}}{\partial \boldsymbol{\beta}}|_{\boldsymbol{\beta}=\widehat{\boldsymbol{\beta}}} = -\boldsymbol{2X'Y} + \boldsymbol{2X'X}\widehat{\boldsymbol{\beta}}$$

karena, $\frac{\partial \boldsymbol{L}}{\partial \boldsymbol{\beta}}|_{\boldsymbol{\beta}=\widehat{\boldsymbol{\beta}}} = 0$, maka

$$-\boldsymbol{2X'Y} + \boldsymbol{2X'X}\widehat{\boldsymbol{\beta}} = 0$$

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

Uji signifikansi parameter regresi linear berganda berguna untuk mengukur ketepatan model, antara lain sebagai berikut:

a. Uji F

Uji F digunakan untuk menguji signifikansi regresi jika ada hubungan linier antara variabel terikat dan variabel bebas secara bersama-sama (Montgomery dan Peck, 1992). Berikut langkah-langkahnya :

1. Hipotesis

   H_0 : $\beta_1 = \beta_2 = \ldots = \beta_k = 0$ (tidak terdapat hubungan antara variabel terikat dengan variabel bebas x_j secara bersama-sama)

H_1 : terdapat $\beta_j \neq 0$, dengan j = 1,2,…,k (terdapat hubungan antara variabel terikat dengan variabel bebas x_j secara bersama-sama)

2. Statistik uji

$$F_{hitung} = \frac{SS_R/k}{SS_E/(n-k-1)} = \frac{\text{MSR}}{\text{MSE}}$$

3. Kriteria uji

H_0 ditolak jika $F_{hitung} > F_{(\alpha,k,n-k-1)}$ atau *p-value* < α

b. Uji Koefisien Regresi secara Individual (Uji t)

Pengujian ini digunakan untuk menguji ada tidaknya pengaruh yang signifikan antara masing – masing variabel bebas terhadap model regresi linier. Langkahnya :

1. Hipotesis

H_0 : $\beta_j = 0$ (koefisien parameter variabel x_j tidak signifikan terhadap y)

H_1 : $\beta_j \neq 0$, (koefisien parameter variabel x_j signifikan terhadap y)

2. Statistik uji

$$t = \frac{\hat{\beta}_j}{Se(\hat{\beta}_j)} \;;\; \text{dengan } Se(\hat{\beta}_j) = \sqrt{var(\hat{\beta}_j)}$$

3. Kriteria uji

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

Goodness of fit digunakan untuk mengetahui seberapa baik kecocokan data observasi dengan model regresi

a. Koefisien Determinasi yang disesuaikan (*Adjusted* R^2)

Pada model regresi berganda, R^2 dikenal untuk mengetahui proporsi variasi y yang dijelaskan beberapa variabel x secara bersama-sama.

$$R^2{}_{Adj,k} = 1 - \frac{SSE/(n-k-1)}{SST/(n-1)} = 1 - \frac{(n-1)}{(n-k-1)}(1 - R^2)$$

b. MSE (*Mean Square Error*)

Menurut Montgomery dan Peck (1992), dalam analisis regresi rumus MSE adalah

$$\text{MSE} = \frac{SSE}{n-k-1}$$

2.2. Ukuran Pemusatan dan Penskalaan (*Centering and Scaling*)

Menurut Kutner *et al.* (2005), perbedaan unit satuan pada model regresi perlu dilakukan standarisasi dengan rumus berikut:

$$Y_i^* = \frac{y_i}{\sqrt{n-1}\,(S_Y)} = \frac{1}{\sqrt{n-1}}\left(\frac{Y_i - \bar{Y}}{S_Y}\right) \quad \text{dimana } S_Y = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n-1}}$$

$$Z_{ij} = \frac{x_{ij}}{\sqrt{n-1}\,(S_j)} = \frac{1}{\sqrt{n-1}}\left(\frac{X_{ij} - \bar{X}_j}{S_j}\right) \qquad S_j = \sqrt{\frac{\sum_{i=1}^{n}(X_{ij} - \bar{X}_j)^2}{n-1}}$$

untuk $i = 1, 2, \dots, n \,;\, j = 1, 2, \dots, k$

sehingga diperoleh model regres standar sebagai berikut:

$$Y_i^* = \beta_1^* Z_{i1} + \beta_2^* Z_{i2} + \cdots + \beta_k^* Z_{ik} + \varepsilon_i$$

hubungan estimator regresi bentuk standar dengan bentuk asli adalah:

$$\beta_j = \left(\frac{S_Y}{S_j}\right)\beta_j^*,\; j = 1, 2, \dots, k \text{ dan } \beta_0 = \bar{Y} - \beta_1\bar{X}_1 - \beta_2\bar{X}_2 \dots - \beta_k\bar{X}_k$$

2.3. Multikolinieritas

Deteksi multikolinieritas bertujuan untuk menguji apakah model regresi ditemukan adanya korelasi antar variabel bebas. Model regresi yang baik seharusnya tidak terjadi korelasi diantara variabel bebas. Salah satu cara untuk mengetahui adanya multikolinieritas, yaitu harga Faktor Inflasi Varian (VIF)

$$VIF(x_j) = \frac{1}{(1-R_j^2)}, \text{ i = 1,2,...,n}$$

Nilai VIF >10 menunjukkan multikolinearitas yang kuat (Montgomery dan Peck, 1992).

2.4. Pencilan

Pencilan adalah suatu pengamatan yang ekstrim. Pencilan adalah titik-titik data yang tidak setipe dengan titik data yang lainnya (Montgomery dan Peck, 1992). Salah satu metode yang dapat digunakan untuk mendeteksi pencilan adalah uji *DFFITS (Difference in Fit Standardized)*. Rumus $DFFITS_i$ didefinisikan :

$$(DFFITS_i) = t_i \left(\frac{h_{ii}}{1-h_{ii}}\right)^{\frac{1}{2}}$$

Dimana t_i adalah *studentized deleted residual* untuk kasus ke-*i* dengan rumus:

$$t_i = e_i \sqrt{\frac{n-p-1}{SS_E(1-h_{ii})-e_i{}^2}}$$

Data disebut pencilan jika $|DFFITS| > 1$ untuk data yang berukuran kecil (n < 15) dan nilai $|DFFITS| > 2\sqrt{p/n}$ untuk gugus data yang berukuran besar, dengan $p = k+1$, dan n adalah banyaknya observasi (Neter, 1997).

2.5. Regresi *Ridge*

Menurut Dereny dan Rashwan (2011), model persamaan *ridge* didasarkan pada penambahan tetapan K pada diagonal utama matriks $\boldsymbol{X'X}$ sehingga model persamaan *ridge* menjadi:

$$\boldsymbol{Y} = \boldsymbol{X\beta_R} + \boldsymbol{\varepsilon}$$

Dalam mengestimasi parameter model menurut Tsutsumi et al. (1997) estimasi regresi *ridge* diperoleh dengan meminimumkan jumlah kuadrat *error*:

$$\sum_{i=1}^{n} \varepsilon_i^2 = \boldsymbol{\varepsilon'\varepsilon} = (\boldsymbol{Y} - \boldsymbol{X\beta_R})'(\boldsymbol{Y} - \boldsymbol{X\beta_R})$$

dengan kendala $\boldsymbol{\beta'_R\beta_R} = \boldsymbol{c^2}$. Bila $\boldsymbol{c}$ tetapan positif yang berhingga, dengan menggunakan metode pengali *langrange* maka diperoleh:

$$L(\boldsymbol{\beta_R}, K) = (\boldsymbol{Y} - \boldsymbol{X\beta_R})'(\boldsymbol{Y} - \boldsymbol{X\beta_R}) + K(\boldsymbol{\beta'_R\beta_R} - \boldsymbol{c^2})$$

sehingga diperoleh esrimator regresi *ridge* yaitu:

$$\boldsymbol{\beta_R} = (\boldsymbol{X'X} + K\boldsymbol{I})^{-1}\boldsymbol{X'Y}$$

2.6. Regresi *Robust*

Berdasarkan Draper dan Smith (1998), regresi *robust* adalah metode regresi yang digunakan ada beberapa pencilan yang mempengaruhi model. Metode ini adalah alat yang penting untuk menganalisis data yang terinfeksi oleh pencilan, dengan begitu hasil model menjadi resisten terhadap pencilan.

Menurut Chen (2002), regresi *robust* memiliki beberapa metode estimasi, diantaranya *M-estimator, S-estimator* dan *MM-estimator.*

a. *Robust M-Estimator*

Menurut Draper dan Smith (1998), *M-estimator* meminimumkan fungsi obyektif :

$$\min_{\beta} \sum_{i=1}^{n} \rho(u_i) = \min_{\beta} \sum_{i=1}^{n} \rho\left(\frac{e_i}{s}\right) = \min_{\beta} \sum_{i=1}^{n} \rho\left(\frac{y_i - \sum_{j=0}^{k} x_{ij}\beta_j}{s}\right)$$

dimana *s* adalah skala estimasi robust. Estimasi *s* yang digunakan adalah:

$$s = \frac{MAD}{0{,}6475} = \frac{median|e_i - median(e_i)|}{0{,}6475} \text{ dan } u_i = \frac{e_i}{s}$$

Untuk mendapatkan estimasi parameter dengan mencari turunan parsial pertama dari ρ terhadap β_j $(j = 0{,}1, \dots, k)$ disamakan dengan 0

$$\sum_{i=1}^{n} x_{ij}\psi\left(\frac{y_i-\sum_{j=0}^{k} x_{ij}\beta_j}{s}\right) = 0, \qquad j = 0,1,\dots,k$$

Draper dan Smith (1998) memberikan solusi dengan mendefinisikan fungsi pembobot :

$$w(u_i) = \frac{\psi\left(\frac{y_i-\sum_{j=0}^{k} x_{ij}\beta_j}{s}\right)}{\left(\frac{y_i-\sum_{j=0}^{k} x_{ij}\beta_j}{s}\right)}$$

dan $w_i = w(u_i)$. Kemudian estimasi persamaan dapat ditulis:

$$\sum_{i=1}^{n} x_{ij}w_i\left(y_i - \sum_{j=0}^{k} x_{ij}\beta_j\right) = 0 \qquad , j = 0, 1, \dots, k$$

Kemudian diselesaikan dengan *Iteratively Reweighted Least Square* (IRLS). Iterasi akan berhenti jika $\widehat{\boldsymbol{\beta}}_{\boldsymbol{j}}$ konvergen yaitu selisih nilai $\widehat{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\boldsymbol{m+1})}$ dan $\widehat{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\boldsymbol{m})}$ mendekati 0.

b. ***Robust S-Estimator***

Estimasi regresi yang memiliki *breakdown point* tinggi salah satunya adalah *S-estimator* yang diperkenalkan oleh Rousseeuw dan Yohai. *S-estimator* didefinisikan:

$$\hat{\beta} = \min_{\beta} s(\, e_1(\boldsymbol{\beta}), e_2(\boldsymbol{\beta}) \dots, e_n(\boldsymbol{\beta}))$$

e_i merupakan residual ke-i dari β dan $s(e_1, \dots, e_n)$ didefinisikan sebagai solusi dari:

$$\frac{1}{n}\sum_{i=1}^{n} \rho\left(\frac{e_i}{s}\right) = \frac{1}{1}\sum_{i=1}^{n} \rho\left(\frac{y_i-\sum_{j=0}^{k} x_{ij}\beta_j}{s}\right) = \delta$$

Agar *breakdown point* 50%, maka $\delta = E_{\Phi}\rho(u_i)$ = 0.1995, dengan skala *robust* (s) yang digunakan adalah:

$$s = \sqrt{\frac{1}{n\delta}\sum_{i=1}^{n} w_i {e_i}^2}$$

pembobot $w_i = w(u_i) = \frac{\rho(u_i)}{{u_i}^2}$ untuk iterasi berikutnya. s untuk iterasi pertama:

$$s = \frac{MAD}{0{,}6475} = \frac{median|e_i - median(e_i)|}{0{,}6475}$$

Dengan menurunkan parsial pertama ρ terhadap β_j seperti pada *MM-Estimator* sehingga didapatkan estimasi persamaan:

$$\sum_{i=1}^{n} x_{ij}w_i\left(y_i - \sum_{j=0}^{k} x_{ij}\beta_j\right) = 0 \qquad , j = 0, 1, \dots, k$$

Kemudian diselesaikan dengan *Iteratively Reweighted Least Square* (IRLS). Iterasi akan berhenti jika $\widehat{\boldsymbol{\beta}}_{\boldsymbol{j}}$ konvergen yaitu selisih nilai $\widehat{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\boldsymbol{m+1})}$ dan $\widehat{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\boldsymbol{m})}$ mendekati 0.

c. ***Robust MM-Estimator***

Menurut Chen (2002), *MM-estimator* merupakan kombinasi antara estimator yang memiliki *breakdownt point* tinggi dan *M-estimator*. Yohai (1987) mendiskripsikan tiga tahapan *MM-estimator* :

1. Menentukan estimasi awal yang ditunjukkan $\widehat{\boldsymbol{\beta}}$ dengan estimator yang memiliki *breakdown point* tinggi, kemungkinan 50% , biasanya *S-estimator*
2. Menghitung residual dan skala *M-estimator* yang didefinisikan oleh persamaan (17) dengan 50% *breakdown point*. $s(e_1(\widehat{\boldsymbol{\beta}}), e_2(\widehat{\boldsymbol{\beta}}) \dots, e_n(\widehat{\boldsymbol{\beta}}))$ dinotasikan s_n.
3. Menghitung estimasi parameter akhir dengan M-*estimator*, menggunakan turunan fungsi pengaruh $\psi(u_i)$ dan skala s_n didapatkan dari langkah kedua.

2.7. Regresi *Ridge Robust-MM*

Regresi *ridge robust-MM* merupakan penggabungan dari metode regresi ridge dan regresi ro*bust* dengan *MM-Estimator*, bedanya dengan regresi *ridge* yang biasa adalah nilai

penduga parmeter yang digunakan. Untuk regresi *ridge* biasa parmeter yang digunakan adalah $\hat{\boldsymbol{\beta}}^{\boldsymbol{MKT}}$, sedangkan *ridge robust-MM* menggunakan $\hat{\boldsymbol{\beta}}^{(\boldsymbol{mm})}$
Rumus penduga parameter regresi *ridge robust* adalah sebagai berikut:

$$\hat{\boldsymbol{\beta}}_{\boldsymbol{RR}} = (\boldsymbol{X}'\boldsymbol{X} + K\boldsymbol{I})^{-\boldsymbol{1}}\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}}^{(\boldsymbol{mm})}$$

3. METODE PENELITIAN

3.1. Sumber Data dan Variabel Penelitian

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

3.2. Variabel Penelitian

Y : Angka Kematian Bayi (AKB)
X_1 : Kepadatan Penduduk
X_2 : Persentase Rumah Tangga Berperilaku Hidup Bersih dan Sehat
X_3 : Jumlah Bayi Berat Badan Lahir Rendah (BBLR)
X_4 : Jumlah Bayi yang Diberi ASI Eksklusif
X_5 : Jumlah Bayi yang Mendapat Kunjungan Neonatal 1 Kali
X_6 : Jumlah Bayi yang Mendapat Pelayanan Kesehatan

3.3. Tahapan Analisis Data

1. Melakukan standarisasi data dengan Transfomasi Korelasi
2. Mengestimasi koefisien regresi $\hat{\boldsymbol{\beta}}$ menggunakan Metode Kuadrat Terkecil.
3. Melakukan Uji F dan uji koefisien secara individual (uji t) dan menghitung *Adjusted* R^2 dan MSE.
4. Melakukan uji multikolinearitas dengan melihat nilai VIF.
5. Melakukan pendeteksian pencilan dengan $DFFITS_i$.
6. Mengestimasi nilai $\hat{\boldsymbol{\beta}}^{(\boldsymbol{mm})}$) dengan menggunakan metode Regresi Robust-MM dengan pembobot tukey's-bisquare
7. Menghitung nilai $\hat{\boldsymbol{\beta}}_{\boldsymbol{RR}}$ pada metode Regresi *Ridge Robust-MM* dengan memasukkan nilai $\boldsymbol{\beta}^{(\boldsymbol{mm})}$ ke dalam rumus penduga tetapan K pada regresi *ridge*.
8. Melakukan uji F dan uji koefisien regresi secara individual (uji t) untuk model regresi *ridge robust-MM* dan menghitung *Adjusted* R^2 dan MSE untuk model regresi *ridge robust-MM.*
9. Membandingkan nilai *Adjusted* R^2 dan MSE model Regresi *Ridge Robust-MM* dengan *Adjusted* R^2 dan MSE metode kuadrat terkecil untuk memilih metode terbaik.

4. HASIL DAN PEMBAHASAN

4.1. Regresi Linier Berganda

Pada awal penelitian dilakukan analisis regresi linier berganda dengan estimasi parameter menggunakan Metode Kuadrat Terkecil.

4.1.1. Estimasi Parameter Regresi

Dari hasil *output* didapatkan model regresi linier berganda sebagai berikut:

$$\hat{y} = -0{,}377\, Z_1 + 0{,}273\, Z_2 + 0{,}995\, Z_3 - 0{,}259\, Z_4 - 0{,}244\, Z_5 - 1{,}013\, Z_6$$

4.1.2. Uji Signifikansi Parameter

a. Uji F

1. Hipotesis
   $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0$
   H_1 : terdapat $\beta_j \neq 0$, j = 1,2,3,4,5,6
2. Taraf signifikansi
   α = 5% = 0,05
3. Statistik uji
   F_{hitung} =5,66 dan *p-value* = 0,000599
4. Kriteria uji
   H_0 ditolak jika $F_{hitung} > (F_{(5\%,;6;28)} = 2{,}45)$ atau *p-value* < α (0,05)
5. Keputusan
   H_0 ditolak karena $F_{hitung} = 5{,}655063 > (F_{(5\%;6;28)} = 2{,}45)$ atau *p-value* (0,000599) < α (0,05)
6. Kesimpulan
   Pada taraf signifikansi 5% terdapat hubungan antara variabel terikat AKB dengan variabel bebas (kepadatan penduduk, persentase rumah tangga yang menggunakan sumber air minum bersih, jumlah bayi berat badan lahir rendah, jumlah bayi yang diberi asi eksklusif, jumlah bayi yang mendapat kunjungan neonatal 1 kali, dan jumlah bayi yang mendapat pelayanan kesehatan) secara bersama-sama.

b. Uji Koefisien Regresi secara Individual (Uji t)

1. Hipotesis
   $H_0 : \beta_j = 0$, dimana j=1,2,3,4,5,6
   $H_1 : \beta_j \neq 0$, dimana j=1,2,3,4,5,6
2. Taraf signifikansi
   α = 5% = 0,05
3. Statistik uji
   Disajikan pada Tabel 1.
4. Kriteria uji
   H_0 ditolak jika $|t_{hitung}| > (t_{(2,5\%;28)} = 2{,}048)$ atau *p-value* < α (0,05)
5. Keputusan

Tabel 1. Tabel Uji t Regresi Linier Berganda dengan MKT

Variabel	t_{hitung}	*p-value*	Keputusan
Z_1	-0,377304	0,03303	H_0 ditolak
Z_2	0,273480	0,10815	H_0 diterima
Z_3	0,994908	0,00116	H_0 ditolak
Z_4	-0,258653	0,07430	H_0 diterima
Z_5	-0,241188	0,68193	H_0 diterima
Z_6	-1,013138	0,11785	H_0 diterima

6. Kesimpulan
   Pada taraf signifikansi 5% koefisien parameter kepadatan penduduk (Z_1) dan jumlah bayi berat badan lahir rendah (Z_3) berpengaruh signifikan terhadap AKB (y). Sedangkan koefisien parameter persentase rumah tangga yang menggunakan sumber air minum bersih (Z_2), jumlah bayi yang diberi asi eksklusif (Z_4), jumlah bayi yang mendapat kunjungan neonatal 1 kali (Z_5), dan jumlah bayi yang mendapat pelayanan kesehatan (Z_6) tidak berpengaruh signifikan terhadap AKB (y).

4.1.3. *Goodness Of Fit*

a. *Adjusted R^2*

$R^2{}_{Adj,k}$ =0,4509964

Artinya AKB dipengaruhi oleh kepadatan penduduk, persentase rumah tangga yang menggunakan sumber air minum bersih, jumlah bayi berat badan lahir rendah, jumlah bayi yang diberi asi eksklusif, jumlah bayi yang mendapat kunjungan neonatal 1 kali, dan jumlah bayi yang mendapat pelayanan kesehatan sebanyak 45,1% dan sisanya 54,9% dipengaruhi faktor yang lain.

b. MSE

MSE = 0,01614716

4.2. Pendeteksian Multikolinieritas dan Pencilan

a. Pendeteksian Multikolinieritas

Tabel 2. Nilai VIF

Variabel	VIF	Keterangan
Z_1	1,75328	VIF < 10
Z_2	1,68168	VIF < 10
Z_3	4,68604	VIF < 10
Z_4	1,20530	VIF < 10
Z_5	21,00526	VIF > 10
Z_6	24,41789	VIF > 10

Kesimpulan : Pada taraf signifikansi 5% terjadi multikolinieritas pada variabel jumlah bayi yang mendapat kunjungan neonatal 1 kali (Z_5), dan jumlah bayi yang mendapat pelayanan kesehatan (Z_6)

b. Pendeteksian Pencilan

Dikarenakan jumlah data (n) = 35 termasuk gugus data besar, maka batas suatu data dikatakan sebagai pencilan jika nilai $|DFFITS| > 2\sqrt{(k+1)/n}$ dengan k = 6, sehingga $|DFFITS| > 0{,}89443$. Dari hasil *output*, didapatkan bahwa ada sebanyak 5 pencilan terdeteksi, yaitu data ke 12, 14, 29, 31, dan 33.

4.3. Regresi *Ridge Robust-MM*

4.3.1. Estimasi Parameter Regresi *Robust-MM*

Model regresi *robust MM-estimator* yang didapatkan setelah proses 15 kali iterasi *S-Estimator* dan 12 kali iterasi *M-estimator* adalah sebagai berikut :

$$\hat{y} = -\,0{,}021 - 0{,}544\,Z_1 + 0{,}232\,Z_2 + 0{,}925\,Z_3 - 0{,}332\,Z_4 + 0{,}151\,Z_5 - 1{,}667\,Z_6$$

4.3.2 Estimasi Parameter Regresi *Ridge Robust-MM*

Pada *output* regresi *ridge robust-MM* dengan nilai *K* = 0,012194 didapatkan model Regresi *Ridge Robust-MM* adalah sebagai berikut:

$$\hat{y} = -\,0{,}376\,Z_1 + 0{,}266\,Z_2 + 0{,}907\,Z_3 - 0{,}246\,Z_4 - 0{,}354\,Z_5 - 0{,}824\,Z_6$$

4.3.3 Uji Signifikansi Parameter

a. Uji F

1. Hipotesis
   $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0$
   H_1 : terdapat $\beta_j \neq 0$, j = 1,2,3,4,5,6
2. Taraf signifikansi
   α = 5% = 0,05
3. Statistik uji
   F_{hitung} =6,16 dan *p-value* = 0,000362

4. Kriteria uji
   H_0 ditolak jika $F_{hitung} > (F_{(5\%,;6;28)} = 2{,}45)$ atau *p-value* < α (0,05)
5. Keputusan
   H_0 ditolak karena $F_{hitung} = 6{,}16 > (F_{(5\%;6;28)} = 2{,}45)$ atau *p-value* $(0{,}000362)$ < α (0,05)
6. Kesimpulan
   Pada taraf signifikansi 5% terdapat hubungan antara variabel terikat AKB dengan variabel bebas (kepadatan penduduk, persentase rumah tangga yang menggunakan sumber air minum bersih, jumlah bayi berat badan lahir rendah, jumlah bayi yang diberi asi eksklusif, jumlah bayi yang mendapat kunjungan neonatal 1 kali, dan jumlah bayi yang mendapat pelayanan kesehatan) secara bersama-sama.

b. Uji Koefisien Regresi secara Individual (Uji t)

1. Hipotesis
   $H_0 : \beta_j = 0$, dimana j=1,2,3,4,5,6
   $H_1 : \beta_j \neq 0$, dimana j=1,2,3,4,5,6
2. Taraf signifikansi
   α = 5% = 0,05
3. Statistik uji
   Disajikan pada Tabel 3.
4. Kriteria uji
   H_0 ditolak jika $|t_{hitung}| > (t_{(2,5\%;28)} = 2{,}048)$ atau *p-value* < α (0,05)
5. Keputusan

Tabel 3. Tabel Uji t Regresi *Ridge Robust-MM*

Variabel	t_{hitung}	*p-value*	Keputusan
Z_1	-3,263436	0,002898	H_0 ditolak
Z_2	2,365519	0,025165	H_0 ditolak
Z_3	4,941762	0,000033	H_0 ditolak
Z_4	-2,556092	0,016299	H_0 ditolak
Z_5	-1,068770	0,294301	H_0 diterima
Z_6	-2,320392	0,078270	H_0 ditolak

7. Kesimpulan
   Pada taraf signifikansi 5% koefisien parameter kepadatan penduduk (Z_1), persentase rumah tangga yang menggunakan sumber air minum bersih (Z_2), jumlah bayi berat badan lahir rendah (Z_3), jumlah bayi yang diberi asi eksklusif (Z_4), dan jumlah bayi yang mendapat pelayanan kesehatan (Z_6) berpengaruh signifikan terhadap AKB (y). Sedangkan koefisien parameter jumlah bayi yang mendapat kunjungan neonatal 1 kali (Z_5) tidak berpengaruh signifikan terhadap AKB (y).

4.3.4. *Goodness Of Fit*

a. *Adjusted R^2*

$R^2{}_{Adj,k}$ =0,4768258

Artinya AKB dipengaruhi oleh kepadatan penduduk, persentase rumah tangga yang menggunakan sumber air minum bersih, jumlah bayi berat badan lahir rendah, jumlah bayi yang diberi asi eksklusif, jumlah bayi yang mendapat kunjungan neonatal 1 kali, dan jumlah bayi yang mendapat pelayanan kesehatan sebanyak 47,68% dan sisanya 52,32% dipengaruhi faktor yang lain.

b. MSE

MSE = 0,01538748

4.4. Pemilihan Model Regresi *Robust* Terbaik

Sebelumnya model regresi harusd dikembalikan ke dalam bentuk variabel asli, didapatkan dua model regresi sebagai berikut:

a. Model regresi linier berganda dengan estimasi parameter menggunakan Metode Kuadrat Terkecil yaitu:

$$\hat{y} = 7{,}469 - 0{,}00056\,X_1 + 0{,}10197\,X_2 + 0{,}01063\,X_3 - 0{,}00009\,X_4 - 0{,}00013\,X_5 - 0{,}0053\,X_6$$

b. Model regresi linier berganda dengan estimasi parameter menggunakan Regresi *Ridge Robust-MM* yaitu:

$$\hat{y} = 7{,}663 - 0{,}00056\,X_1 + 0{,}09929\,X_2 + 0{,}00969\,X_3 - 0{,}00009\,X_4 - 0{,}00019\,X_5 - 0{,}0043\,X_6$$

Pemilihan metode terbaik didasarkan pada besarnya nilai $Adjusted\ R^2$ dan MSE, disajikan dalam Tabel 4.

Tabel 4. Tabel Pembanding Regresi (MKT) dan Regresi *Ridge Robust*

	$\boldsymbol{Adjusted\ R^2}$	**MSE**
Metode Kuadrat Terkecil	0,45099	0,016147
Ridge Robust-MM	0,47682	0,015387

Kriteria pemilihan model regresi *robust* terbaik yaitu mempunyai $R^2{}_{Adj,k}$ terbesar dan nilai MSE terkecil. Dari Tabel 4, disimpulkan bahwa model terbaik adalah Regresi *Ridge Robust-MM*

$$\hat{y} = 7{,}66279 - 0{,}00056\,X_1 + 0{,}09929\,X_2 + 0{,}00969\,X_3 - 0{,}00009\,X_4 - 0{,}00019\,X_5 - 0{,}0043\,X_6$$

dengan besar pengaruh variabel bebas terhadap AKB di Jawa Tengah adalah 47,68% dan nilai MSE nya 0,015387

5. KESIMPULAN

Berdasarkan hasil dan pembahasan dalam pemodelan SHU koperasi di Jawa Tengah, diperoleh kesimpulan sebagai berikut :

1. Estimasi parameter MKT diperoleh model regresi sebagai berikut:

   $$\hat{y} = 7{,}46977 - 0{,}00056\,X_1 + 0{,}10197\,X_2 + 0{,}01063\,X_3 - 0{,}00009\,X_4 - 0{,}00013\,X_5 - 0{,}0053\,X_6$$

   Terdapat 2 variabel dengan masalah multikolinieritas dan terdapat 5 buah pencilan. Maka perlu dilakukan penanganan menggunakan metode Regresi *Ridge Robust-MM*. Setelah dilakukan estimasi parameter menggunakan Regresi *Ridge Robust-MM* diperoleh model:

   $$\hat{y} = 7{,}66279 - 0{,}00056\,X_1 + 0{,}09929\,X_2 + 0{,}00969\,X_3 - 0{,}00009\,X_4 - 0{,}00019\,X_5 - 0{,}0043\,X_6$$

   kemudian kedua model dibandingkan berdasarkan nilai *Adjusted* R^2 dan MSE, diperoleh metode Regresi *Ridge Robust-MM* sebagai model terbaik.
2. Angka Kematian Bayi Provinsi Jawa Tengah tahun 2017 dipengaruhi oleh kepadatan penduduk, persentase rumah tangga yang menggunakan sumber air minum bersih,

jumlah bayi berat badan lahir rendah, jumlah bayi yang diberi asi eksklusif, dan jumlah bayi yang mendapat pelayanan kesehatan, dengan Regresi *Ridge Robust-MM* yang menghasilkan $Adjusted\ R^2$ sebesar 47,682 % dan MSE sebesar 0,015387

DAFTAR PUSTAKA

Chen, C. 2002. *Robust Regression and Outlier Detection with The ROBUSTREG Procedure*, paper 265-27. SAS Instutute Inc., Lary, NC.

Dereny, M.E., Rashwan, N. I. 2011. *Solving Multicolinierity Problems Using Ridge Regression Models*. Int. J. Contemp. Math. Sciences Vol. 6 No. 12 Hal. 585-600.

[Dinkes Jateng] Dinas Kesehatan Provinsi Jawa Tengah. 2018. *Profil Kesehatan Provinsi Jawa Tengah Tahun 2017*. http://dinkesjatengprov.go.id/ (diunduh 30 Desember 2018).

Draper, N.R., Smith, H. 1998. *Applied Regression Analysis.* Third Edition. New York: John Wiley and Sons.

Hoerl, A.E., Kennard, R.W. 1970. *Ridge Regression: Biased Estimation for Nonorthogonal Problems*. Technometrics Vol. 12, No. 1: Hal. 55-67. Americal Statistical Association and American Society for Quality.

Kutner, M.H., Nachtsheim, C. J., Li, W. 2005. *Applied Linear Statistics Model*, Fifth Edition. New York: The Mc Graw- Hill Companies.

Montgomery, D. C., Peck, E. A. 1992. *Introduction To Linier Regression Analysis.* New York: John Wiley and Sons, Inc.

Neter, J. 1997. *Model Linier Terapan*. Diterjemahkan oleh: B. Sumantri. Bandung. Terjemahan dari: *Applied Linier Regression Model.* Richard D. Irwin. Inc.

Sembiring, R.K. 2003. *Analisis Regresi* Edisi Kedua. Bandung: ITB.

Tsutsumi, M., Matsuba, Y., Shimizu, E. 1997. *A Comparative Study on Counter-Measures of Multicollinearity in Regression Analysis*. Jurnal of the Eastern Asia Society for Transportation Studies Vol. 2 No. 6.