ISSN: 2339-2541
JURNAL GAUSSIAN, Volume 8, Nomor 3, Tahun 2019, Halaman 377-3
Online di: http://ejournal3.undip.ac.id/index.php/gaussian

REGRESI *ROBUST* ESTIMASI-M DENGAN PEMBOBOT *ANDREW*, PEMBOBOT *RAMSAY* DAN PEMBOBOT *WELSCH* MENGGUNAKAN *SOFTWARE R*

Aulia Desy Deria[1], Abdul Hoyyi[2], Mustafid[3]
[1,2,3] Departemen Statistika FSM Universitas Diponegoro
hoyyistat@gmail.com

ABSTRACT

Robust regression is one of the regression methods that robust from effect of outliers. For the regression with the parameter estimation used Ordinary Least Squares (OLS), outliers can caused assumption violation, so the estimator obtained became bias and inefficient. As a solution, robust regression M-estimation with Andrew, Ramsay and Welsch weight function can be used to overcome the presence of outliers. The aim of this study was to develop a model for case study of poverty in Central Java 2017 influenced by the number of unemployment, population, school participation rate, Human Development Index (HDI), and inflation. The result of estimation using OLS show that there is violation of heteroskedasticity caused by the presence outliers. Applied robust regression to case study proves robust regression can solve outliers and improve parameter estimation. The best robust regression model is robust regression M-estimation with Andrew weight function. The influence value of predictor variables to poverty is 92,7714% and MSE value is 370,8817.

Keywords: Outliers, Robust Regression, M-Estimator, Andrew, Ramsay, Welsch

1. PENDAHULUAN

Kemiskinan masih menjadi musuh terbesar bangsa Indonesia. Menurut BPS (2018), penduduk yang memiliki pengeluaran per kapita setiap bulannya dibawah garis kemiskinan, maka digolongkan sebagai penduduk miskin. Faktor-faktor yang diduga mempengaruhi kemiskinan di Jawa Tengah adalah pengangguran terbuka, jumlah penduduk, pendidikan, Indeks Pembangunan Manusia (IPM), dan inflasi.

Dari hasil studi awal penelitian menggunakan regresi linier berganda dengan estimasi parameter Metode Kuadrat Terkecil (MKT), diketahui pada data tentang kemiskinan terjadi pelanggaran asumsi heteroskedastisitas. Terjadinya heteroskedastisitas dalam model regresi kadangkala diikuti dengan munculnya beberapa pencilan dalam data. Menurut Sembiring (2003), adanya pencilan dalam data dapat mengakibatkan estimasi koefisien regresi yang diperoleh tidak tepat. Menurut Draper dan Smith (1992) penghapusan begitu saja terhadap data pencilan bukanlah langkah yang bijaksana. Adakalanya pencilan memberikan informasi yang tidak bisa diberikan oleh titik data lainnya.

Metode yang dapat digunakan untuk menganalisis data mengandung pencilan adalah regresi *robust*. Regresi *robust* terdiri dari 5 metode estimasi, yaitu *M-estimator*, *Least Median Square (LMS)-estimator*, *Least Trimmed Square (LTS)-estimator*, *S-estimator*, dan *MM-estimator* (Chen, 2002). Estimasi-M merupakan estimasi yang sering digunakan dan dapat dianalisis dengan mudah secara teoritis maupun komputer. Penelitian ini membahas analisis data kemiskinan tiap kabupaten/kota di Jawa Tengah menggunakan regresi *robust M-estimator* dengan menggunakan fungsi pembobot *Andrew*, pembobot *Ramsay* dan pembobot *Welsch*. Dari hasil analisis, kemudian akan dilakukan penentuan pemodelan regresi *robust* terbaik. *Software* yang digunakan untuk menganalisis penelitian ini adalah *software R*.

2. TINJAUAN PUSTAKA

2.1. Kemiskinan

Kemiskinan secara umum dapat diartikan sebagai kondisi individu yang tidak mampu memenuhi kebutuhan hidup dasarnya secara layak. Menurut BPS (2018), penduduk yang memiliki pengeluaran per kapita setiap bulannya dibawah garis kemiskinan, maka digolongkan sebagai penduduk miskin.

2.2. Faktor-Faktor yang Mempengaruhi Kemiskinan

2.2.1. Pengangguran

Menurut Sukirno (2001), pengangguran merupakan seseorang yang telah masuk dalam golongan angkatan kerja, yang secara aktif sedang mencari pekerjaan pada suatu tingkat upah tertentu, tetapi tidak dapat memperoleh pekerjaan yang diinginkan.

2.2.2. Jumlah Penduduk

Menurut BPS (2018), penduduk merupakan semua orang yang berdomisili di wilayah teritorial Indonesia selama 6 (enam) bulan atau lebih dan atau mereka yang berdomisili kurang dari 6 (enam) bulan tetapi bertujuan untuk menetap.

2.2.3. Pendidikan

Angka Partisipasi Sekolah (APS) pada penduduk umur 16-18 tahun digunakan untuk mewakili variabel pendidikan. Menurut BPS (2018), APS merupakan proporsi dari semua anak yang masih sekolah pada suatu kelompok umur tertentu terhadap penduduk dengan kelompok umur yang sesuai.

2.2.4. Indeks Pembangunan Manusia

Menurut BPS (2018), Indeks Pembangunan Manusia (IPM) menjelaskan bagaimana penduduk dapat mengakses hasil pembangunan dalam memperoleh pendapatan, kesehatan, pendidikan dan sebagainya. IPM dibentuk oleh 3 (tiga) dimensi dasar, yaitu umur panjang serta hidup sehat, pengetahuan dan standar hidup layak.

2.2.5. Inflasi

Secara umum, inflasi dapat diartikan sebagai meningkatnya harga-harga secara umum dan terus-menerus.

2.3. Analisis Regresi Linier

2.3.1. Regresi Linier Berganda

Menurut Montgomery dan Peck (2012), Model regresi berganda dinyatakan dalam bentuk persamaan :

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \varepsilon_i \quad (1)$$

Jika ditulis dalam bentuk matriks, persamaan (1) dapat ditulis sebagai berikut :

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \quad (2)$$

$$\mathbf{y} = \begin{bmatrix} \mathrm{y}_1 \\ y_2 \\ \vdots \\ \mathrm{y}_\mathrm{n} \end{bmatrix} \; ; \boldsymbol{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} \; ; \; \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \; ; \; \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_\mathrm{n} \end{bmatrix}$$

2.3.2. Estimasi Parameter Model Regresi Linier Berganda

Metode yang digunakan untuk mengestimasi parameter model regresi adalah dengan meminimumkan jumlah kuadrat residual atau yang sering dikenal dengan Metode Kuadrat Terkecil (MKT) (Gujarati, 1997).

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

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

$$\frac{\partial J}{\partial \boldsymbol{\beta}}|_{\boldsymbol{\beta}=\widehat{\boldsymbol{\beta}}} = -\boldsymbol{2X'y} + \boldsymbol{2X'X\widehat{\beta}} \tag{3}$$

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

$$\begin{aligned} -\boldsymbol{2X'y} + \boldsymbol{2X'X\widehat{\beta}} &= 0 \\ \widehat{\boldsymbol{\beta}} &= (\boldsymbol{X'X})^{-1}\,\boldsymbol{X'y} \end{aligned} \tag{4}$$

2.3.3. Uji Hipotesis dalam Regresi Linier Berganda

a. Uji F

Uji signifikansi regresi digunakan untuk menguji apakah ada hubungan linier antara variabel respon *y* dan variabel prediktor $x_1, x_2, \dots, x_k$ secara bersama-sama (Montgomery dan Runger, 2003). Berikut langkah-langkahnya :

1. Hipotesis
   $H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$
   H_1 : terdapat $\beta_j \neq 0$, dengan j = 1,2,…,k
2. Statistik uji
   $$F_{hitung} = \frac{SS_R/k}{SS_E/(n-k-1)} = \frac{\text{MSR}}{\text{MSE}} \tag{5}$$
3. Kriteria uji
   H_0 ditolak jika $F_{hitung} > F_{(\alpha,k,n-k-1)}$ atau *p-value* $< \alpha$

b. Uji Koefisien Regresi secara Individual (Uji t)

Uji koefisien regresi secara individual digunakan untuk menguji ada tidaknya pengaruh yang signifikan antara masing – masing variabel prediktor terhadap model regresi linier. Menurut Montgomery dan Runger (2003), berikut langkah-langkah pengujiannya :

1. Hipotesis
   $H_0 : \beta_j = 0$
   $H_1 : \beta_j \neq 0$
2. Statistik uji
   $$t_{hitung} = \frac{\hat{\beta}_j}{Se(\hat{\beta}_j)} \;;\; \text{dengan } Se(\hat{\beta}_j) = \sqrt{var(\hat{\beta}_j)} \tag{6}$$
3. Kriteria uji
   H_0 ditolak jika $|t_{hitung}| > t_{(\alpha/2,n-k-1)}$ atau *p-value* $< \alpha$

2.3.4. *Goodness Of Fit* (Ukuran Kecocokan Model)

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

Menurut Ghozali (2011), koefisien determinasi (R^2) merupakan alat untuk mengukur seberapa jauh kemampuan model dalam menerangkan variasi variabel respon.

$$R^2{}_{Adj,k} = 1 - \frac{SS_E/_{(n-k-1)}}{SS_T/_{(n-1)}} \tag{7}$$

b. MSE (Mean Square Error)

Menurut Montgomery dan Peck (2012), dalam analisis regresi rumus MSE adalah

$$\text{MSE} = \frac{SS_E}{n-k-1} \tag{8}$$

2.3.5. Uji Asumsi dalam Regresi Linier Berganda

a. Uji Normalitas

Uji normalitas digunakan untuk mengetahui apakah residual dalam model regresi berdistribusi normal atau tidak. Gujarati (1997) menjelaskan bahwa pada regresi linear klasik diasumsikan bahwa tiap ε_i didistribusikan normal dengan $\varepsilon_i \sim N(0, \sigma^2)$. Salah satu uji formal yang dapat digunakan untuk menguji asumsi normalitas adalah uji statistik Kolmogorov-Smirnov.

$$D = \sup_{e} |F_0(e) - S(e)| \tag{9}$$

b. Multikolinieritas

Menurut Montgomery dan Peck (2012), multikolinearitas terjadi karena terdapat korelasi yang cukup tinggi di antara variabel independen. Salah satu cara untuk mengukur besar multikolinearitas adalah menggunakan *Variance Inflation Factor* (VIF) yang didefinisikan sebagai berikut :

$$VIF_j = \frac{1}{(1-R_j^2)}, \quad j = 1,2,...,k \tag{10}$$

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

c. Uji Autokorelasi

Pada konteks regresi, model regresi linear klasik mengasumsikan bahwa autokorelasi seperti itu tidak terdapat dalam *error* ε_i (Gujarati, 1997). Lambang non-autokorelasi adalah

$$E(\varepsilon_i \varepsilon_j) = 0 \qquad i \neq j$$

Secara sederhana dapat dikatakan model klasik mengasumsikan bahwa error yang berhubungan dengan pengamatan tidak dipengaruhi oleh error yang berhubungan dengan pengamatan lain yang manapun. Salah satu cara yang dapat digunakan untuk mendeteksi autokorelasi adalah dengan uji Durbin Watson.

$$d = \frac{\sum_{i=2}^{i=n} (e_i - e_{i-1})^2}{\sum_{i=1}^{i=n} e_i^2} \tag{11}$$

d. Uji Heteroskedastisitas

Menurut Ghozali (2011) uji heteroskedastisitas bertujuan untuk menguji apakah dalam model regresi terjadi ketidaksamaan varian dari residual satu pengamatan ke pengamatan yang lain. Lambang homoskedastisitas : $E(\varepsilon_i^2) = \sigma^2 \qquad i = 1,2,\ldots,n$

Heteroskedastisitas dapat diperiksa dengan menggunakan uji Glejser. Uji Glejser dilakukan dengan meregresikan nilai absolut residual $|e_i|$ terhadap variabel-variabel prediktor lainnya (Ghozali, 2011).

$$|e_i| = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + v \tag{12}$$

2.4. Pencilan

Montgomery dan Peck (2012) menyatakan bahwa pencilan adalah titik-titik data yang tidak setipe dengan titik data yang lainnya. Salah satu metode yang dapat digunakan untuk mendeteksi pencilan adalah *DFFITS (Difference in Fit Standardized).* Rumus $DFFITS_i$ didefinisikan :

$$(DFFITS_i) = t_i \left(\frac{h_{ii}}{1-h_{ii}}\right)^{\frac{1}{2}} \tag{13}$$

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

$$t_i = e_i\sqrt{\frac{n-p-1}{SS_E(1-h_{ii})-e_i{}^2}} \tag{14}$$

Suatu data disebut pencilan jika nilai $|DFFITS| > 1$ untuk gugus data yang berukuran kecil sampai sedang 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 *Robust M-Estimator*

Regresi *robust* merupakan metode yang digunakan untuk mengatasi masalah dari metode kuadrat terkecil yang disebabkan oleh data yang terkontaminasi oleh pencilan. Menurut Chen (2002), regresi *robust* memiliki lima metode estimasi yaitu *M-estimator, LMS-estimator, LTS-estimator*, *S-estimator* dan *MM-estimator*. Namun pada penelitian ini hanya digunakan *M-estimator*.

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

$$\min_{\beta} \sum_{i=1}^{n} \rho(u_i) = \min_{\beta} \sum_{i=1}^{n} \rho\left(\frac{e_i}{s}\right) = \min_{\beta} \sum_{i=1}^{n} \rho\left(\frac{y_i - \sum_{j=0}^{k} x_{ij}\beta_j}{s}\right) \tag{15}$$

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

$$s = \frac{MAD}{0{,}6745} = \frac{median|e_i - median(e_i)|}{0{,}6745} \tag{16}$$

dan $u_i = \frac{e_i}{s}$

Untuk meminimumkan persamaan (15), akan digunakan turunan parsial pertama fungsi ρ terhadap β_j $(j = 0,1, \dots, k)$ sama dengan 0, sehingga diperoleh :

$$\sum_{i=1}^{n} x_{ij}\psi\left(\frac{y_i - \sum_{j=0}^{k} x_{ij}\beta_j}{s}\right) = 0, \qquad j = 0,1, \dots, k \tag{17}$$

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

$$w(u_i) = \frac{\psi\left(\frac{y_i - \sum_{j=0}^{k} x_{ij}\beta_j}{s}\right)}{\left(\frac{y_i - \sum_{j=0}^{k} x_{ij}\beta_j}{s}\right)} \tag{18}$$

dan $w_i = w(u_i)$. Kemudian estimasi persamaan (17) dapat ditulis:

$$\sum_{i=1}^{n} x_{ij} w_i \left(y_i - \sum_{j=0}^{k} x_{ij}\beta_j\right) = 0 \qquad , j = 0, 1, \dots, k \tag{19}$$

Persamaan (19) diselesaikan dengan *Iteratively Reweighted Least Square* (IRLS). Pada notasi matrik, persamaan (19) dapat ditulis:

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{j}} = (\boldsymbol{X}'\boldsymbol{W}\boldsymbol{X})^{-\mathbf{1}}\boldsymbol{X}'\boldsymbol{W}\mathbf{y} \tag{20}$$

Iterasi akan berhenti jika $\widehat{\boldsymbol{\beta}}_{\boldsymbol{j}}$ konvergen yaitu selisih nilai $\widehat{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\boldsymbol{m}+\mathbf{1})}$ dan $\widehat{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\boldsymbol{m})}$ mendekati 0.

2.6. Fungsi Obyektif

Fungsi yang digunakan untuk mencari fungsi pembobot pada regresi robust adalah fungsi obyektif (Fox, 2002). Fungsi obyektif yang diturunkan terhadap u_i akan menghasilkan suatu fungsi yang disebut sebagai fungsi pengaruh $(\psi(u_i))$. Fungsi pengaruh yang dibagi dengan u_i akan menghasilkan fungsi pembobot yang digunakan dalam perhitungan regresi *robust*.

Tabel 1. Fungsi Obyektif dan Fungsi Pembobot untuk Estimasi Kuadrat Terkecil, *Andrew*, *Ramsay* dan *Welsch*

Pembobot	Fungsi Obyektif	Fungsi Pembobot	Interval
Kuadrat Terkecil	$\rho(u_i) = \frac{1}{2}u_i^2$	$w(u_i) = 1$	$\lvert u_i \rvert < \infty$
Andrew	$\rho(u_i) = \begin{cases} c_1\left(1 - \cos\left(\frac{u_i}{c_1}\right)\right) \\ 2c_1 \end{cases}$	$w(u_i) = \begin{cases} \frac{\sin\left(\frac{u_i}{c_1}\right)}{\frac{u_i}{c_1}} \\ 0 \end{cases}$	$\lvert u_i \rvert \le c_1\pi$ $\lvert u_i \rvert > c_1\pi$
Ramsay	$\rho(u_i) = \frac{1 - (exp(-c_2\lvert u_i \rvert))(1 + c_2\lvert u_i \rvert)}{(c_2)^2}$	$w(u_i) = exp(-c_2\lvert u_i \rvert)$	$\lvert u_i \rvert < \infty$
Welsch	$\rho(u_i) = \frac{c_3}{2}\left[1 - exp\left(-\left(\frac{u_i}{c_3}\right)^2\right)\right]$	$w(u_i) = exp\left(-\left(\frac{u_i}{c_3}\right)^2\right)$	$\lvert u_i \rvert < \infty$

Dimana nilai *c* pada pembobot *Andrew*, pembobot *Ramsay* dan pembobot *Welsch* disebut *tunning constant*. Menurut Montgomery dan Peck (2012), *tunning constant* untuk pembobot *Andrew* adalah $c_1 = 1{,}339$, untuk pembobot *Ramsay* adalah $c_2 = 0{,}3$. Menurut Setiarini dan Listyani (2017), *tunning constant* untuk pembobot *Welsch* adalah $c_3 = 2{,}9846$.

3. METODOLOGI PENELITIAN

3.1 Jenis dan Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Badan Pusat Statistik Provinsi Jawa Tengah melalui publikasi Jawa Tengah Dalam Angka 2018. Data yang terdapat dalam penelitian ini adalah data jumlah penduduk miskin, jumlah pengangguran terbuka, jumlah penduduk, Angka Partisipasi Sekolah (APS) untuk umur 16-18 tahun, Indeks Pembangunan Manusia (IPM), dan inflasi setiap kabupaten/kota di Jawa Tengah tahun 2017.

3.2 Variabel Penelitian

Variabel yang digunakan dalam penelitian ini adalah:

y : Jumlah penduduk miskin
x_1 : Jumlah pengangguran terbuka
x_2 : Jumlah penduduk
x_3 : Angka Partisipasi Sekolah untuk umur 16-18 tahun
x_4 : Indeks Pembangunan Manusia
x_5 : Inflasi

3.3 Langkah-Langkah Penelitian

1. Mengumpulkan data dan menentukan variabel respon dan variabel prediktor.
2. Estimasi koefisien regresi menggunakan metode kuadrat terkecil.
3. Melakukan uji F dan uji koefisien regresi secara individual (uji t).
4. Menghitung koefisien determinasi yang disesuaikan (*Adjusted* R^2) dan MSE
5. Melakukan uji asumsi regresi liner berganda.
6. Melakukan pendeteksian pencilan dengan $DFFITS_i$.
7. Estimasi parameter dengan regresi *robust M-Estimator*
8. Melakukan uji F dan uji koefisien regresi *robust* secara individual (uji t).
9. Menentukan model regresi *robust* terbaik dan menganalisisnya.

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:

$$\hat{y} = 496{,}303500 + 0{,}000789x_1 + 0{,}000089x_2 + 1{,}708834x_3 - 9{,}041847x_4 + 13{,}684860x_5$$

4.1.2. Uji Hipotesis

a. Uji F

1. Hipotesis
   $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$
   H_1 : terdapat $\beta_j \neq 0$, j = 1,2,3,4,5
2. Taraf signifikansi
   $\alpha = 5\%$
3. Statistik uji
   $F_{hitung} = 21{,}628240$
4. Kriteria uji
   H_0 ditolak jika $F_{hitung} > (F_{(5\%,5,29)} = 2{,}545386)$
5. Keputusan
   H_0 ditolak karena nilai $F_{hitung} = 21{,}628240 > F_{(5\%,5,29)} = 2{,}545386$
6. Kesimpulan
   Pada taraf signifikansi 5% paling sedikit ada satu variabel prediktor (jumlah pengangguran terbuka, jumlah penduduk, APS, IPM, dan inflasi) yang ada hubungan dengan variabel respon jumlah penduduk miskin 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
   $H_1 : \beta_j \neq 0$, dimana j=1,2,3,4,5
2. Taraf signifikansi
   $\alpha = 5\%$
3. Statistik uji
   Disajikan pada Tabel 2.
4. Kriteria uji
   H_0 ditolak jika $|t_{hitung}| > (t_{(2,5\%,135)} = 1{,}97769)$ atau *p-value* < α (0,05)
5. Keputusan

Tabel 2. Tabel Uji t Regresi Linier Berganda

Variabel	$\lvert t_{hitung} \rvert$	Keputusan
Jumlah Pengangguran Terbuka	0,955834	H_0 diterima
Jumlah Penduduk	2,904232	H_0 ditolak
APS	1,845054	H_0 diterima
IPM	4,469570	H_0 ditolak
Inflasi	0,902133	H_0 diterima

6. Kesimpulan
   Pada taraf signifikansi 5% koefisien parameter variabel yang berpengaruh signifikan terhadap jumlah penduduk miskin (y) adalah jumlah penduduk (x_2) dan IPM (x_4).

4.1.3. *Goodness Of Fit*

Dalam penelitian ini, uji kecocokan model yang digunakan adalah nilai koefisien determinasi yang disesuaikan dan nilai MSE. Diperoleh nilai koefisien determinasi yang disesuaikan sebagai berikut:

$$R^2{}_{Adj,k} = 0{,}752080$$

Hal ini mengindikasikan bahwa variasi total jumlah penduduk miskin (y) sebesar 75,2080% diterangkan oleh jumlah pengangguran terbuka (x_1), jumlah penduduk (x_2), APS (x_3), IPM (x_4), dan inflasi (x_5), sedangkan sisanya 24,7920% dipengaruhi faktor yang lain.

Nilai MSE untuk regresi linier berganda menggunakan estimasi Metode Kuadrat Terkecil adalah sebagai berikut:

MSE = 1366,756

4.1.4. Uji Asumsi dalam Regresi Linier Berganda

a. Uji Normalitas

Uji Normalitas menggunakan uji Kolmogorov-Smirnov, sehingga diperoleh nilai D sebesar 0,094193. Nilai D tersebut lebih kecil dari $d_{\left(1-\frac{0,05}{2}\right)} = d_{(0,975)} = 0{,}246$ maka dapat disimpulkan bahwa pada taraf signifikansi 5% residual berdistribusi normal.

b. Multikolinieritas

Tabel 3. Nilai VIF

Variabel	VIF	Keterangan
Jumlah Pengangguran Terbuka	4,266684	VIF < 10
Jumlah Penduduk	4,199171	VIF < 10
APS	2,054416	VIF < 10
IPM	2,043627	VIF < 10
Inflasi	1,164812	VIF < 10

Berdasarkan Tabel 3 dapat disimpulkan bahwa pada taraf signifikansi 5%, tidak terjadi multikolinieritas

c. Uji Autokorelasi

Uji Autokorelasi menggunakan uji Durbin Watson, sehingga diperoleh nilai d sebesar 2,140334. Nilai d tersebut berada diantara d_U dan 4 - d_U sebagai berikut:

d_U(1,802900) ≤ d(2,140334) ≤ 4 - d_U (2,197100)

Pada taraf signifikansi 5%, dapat disimpulkan bahwa tidak ada autokorelasi antar residual.

d. Uji Heteroskedastisitas

Tabel 4. Tabel Uji Glejser

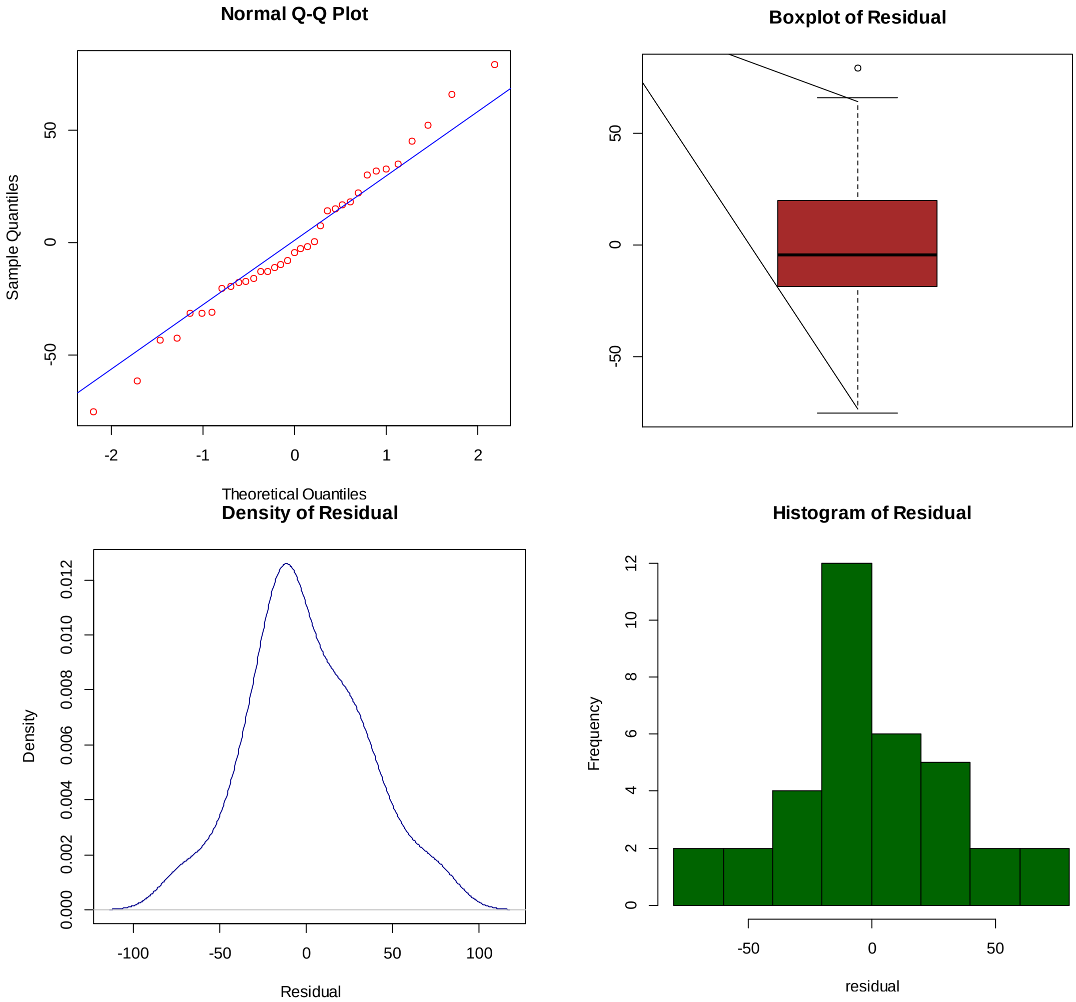
Variabel	$\lvert t_{hitung} \rvert$	Keputusan
Jumlah Pengangguran Terbuka	2,273737	H_0 ditolak
Jumlah Penduduk	0,294061	H_0 diterima
APS	0,742580	H_0 diterima
IPM	1,464981	H_0 diterima
Inflasi	1,005814	H_0 diterima

Berdasarkan Tabel 4 dapat disimpulksn bahwa pada taraf signifikansi 5%, terjadi heteroskedastisitas.

4.2. Indentifikasi Pencilan

Dalam kaitannya dengan analisis regresi, adanya pelanggaran asumsi heteroskedastisitas kadangkala diikuti dengan munculnya beberapa pencilan. Berikut cara untuk mengidentifikasi pencilan:

a. Metode Grafis



Gambar 1. Plot Residual dengan Estimasi Parameter MKT

Pada Gambar 1, secara visual terdapat pencilan dalam data residual. Untuk memastikan seberapa banyak pencilan dan data mana saja yang termasuk pencilan, maka dilakukan pendeteksian pencilan secara formal menggunakan *DFFITS*.

b. Identifikasi Pencilan dengan *DFFITS*

Jumlah data (n) = 35 termasuk gugus data kecil atau sedang, maka batas suatu data dikatakan sebagai pencilan jika nilai $|DFFITS| > 1$. Dari hasil *output*, didapatkan bahwa ada sebanyak 4 pencilan terdeteksi, yaitu data ke-1 (Kota Semarang), data ke-16 (Kab. Banyumas), data ke-32 (Kab. Tegal), dan data ke-35 (Kab. Brebes).

4.3. Regresi *Robust*

4.3.1. Regresi Robust *M-Estimator* (Pembobot *Andrew*)

Model regresi *robust M-estimator* dengan pembobot *Andrew* yang didapatkan setelah proses 87 kali iterasi adalah sebagai berikut :

$$\hat{y} = 341{,}6013 + 0{,}003069x_1 + 0{,}000061x_2 + 1{,}064543x_3 - 5{,}089440x_4 - 13{,}976840x_5$$

4.3.2. Regresi *Robust M-Estimator* (Pembobot *Ramsay*)

Model regresi *robust M-estimator* dengan pembobot *Ramsay* yang didapatkan setelah proses 23 kali iterasi adalah sebagai berikut :

$$\hat{y} = 502{,}6822 + 0{,}001153x_1 + 0{,}000080\, x_2 + 1{,}676396x_3 - 8{,}809452x_4 + 7{,}899445\, x_5$$

4.3.3. Regresi *Robust M-Estimator* (Pembobot *Welsch*)

Model regresi *robust M-estimator* dengan pembobot *Welsch* yang didapatkan setelah proses 30 kali iterasi adalah sebagai berikut :

$$\hat{y} = 323{,}516900 + 0{,}002922\, x_1 + 0{,}000065\, x_2 + 1{,}143818\, x_3 - 5{,}142116\, x_4 - 9{,}763728\, x_5$$

4.4. Pemilihan Model Regresi *Robust* Terbaik

Tabel 5. Tabel Pembanding Regresi (MKT) dan Regresi *Robust M-Estimator*

	$Adjusted\ R^2$	MSE	Variabel yang Signifikan
Pembobot Kuadrat Terkecil	0,752080	1366,7560	x_2, x_4
Pembobot *Andrew*	0,927714	370,8817	x_1, x_2, x_4
Pembobot *Ramsay*	0,791707	817,1030	x_2, x_3, x_4
Pembobot *Welsch*	0,914254	396,2673	x_1, x_2, x_3, x_4

Kriteria pemilihan model regresi *robust* terbaik yaitu mempunyai $R^2{}_{Adj,k}$ terbesar dan nilai MSE terkecil. Dari Tabel 5, disimpulkan bahwa model terbaik adalah *M-estimator* dengan pembobot *Andrew*.

$$\hat{y} = 341{,}601300 + 0{,}003069\, x_1 + 0{,}000061\, x_2 + 1{,}064543\, x_3 - 5{,}089440x_4 - 13{,}976840x_5$$

Dengan besar pengaruh jumlah pengangguran terbuka, jumlah penduduk, APS, IPM dan inflasi terhadap jumlah penduduk miskin di Jawa Tengah adalah 92,7714% dan nilai MSE nya 370,8817.

5. KESIMPULAN

Berdasarkan hasil penelitian diketahui bahwa melalui regresi dengan estimasi parameter menggunakan MKT terjadi pelanggaran asumsi heteroskedastisitas dan terdapat 4 pencilan, sehingga estimator yang dihasilkan bersifat bias dan tidak efisien. Regresi *robust M-estimator* dengan pembobot *Andrew*, pembobot *Ramsay* dan pembobot *Welsch* digunakan untuk mengatasi masalah tersebut. Kemudian dibandingkan berdasarkan nilai *Adjusted R-square* dan *MSE*, diperoleh model terbaik adalah pembobot *Andrew* :

$$\hat{y} = 341{,}601300 + 0{,}003069\, x_1 + 0{,}000061\, x_2 + 1{,}064543\, x_3 - 5{,}089440x_4 - 13{,}976840x_5$$

Regresi *robust M-estimator* dengan pembobot *Andrew* memberikan kesimpulan bahwa kemiskinan di Jawa Tengah dipengaruhi oleh jumlah pengangguran terbuka(x_1), jumlah penduduk(x_2), APS(x_3), IPM(x_4), dan inflasi(x_5), sebesar 92,7714% dan nilai *MSE* nya 370,8817. Untuk variabel jumlah pengangguran terbuka, jumlah penduduk dan IPM berpengaruh signifikan terhadap kemiskinan, sedangkan APS dan inflasi tidak berpengaruh signifikan terhadap kemiskinan.

DAFTAR PUSTAKA

[BPS] Badan Pusat Statistik. 2018. *Katalog Provinsi Jawa Tengah dalam Angka 2018*. Jakarta : BPS.

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

Draper, N.R. dan Smith, H. 1992. *Analisis Regresi Terapan*. Edisi Kedua. Diterjemahkan oleh : Bambang Sumantri. Jakarta : Gramedia Pustaka Utama. Terjemahan dari : *Applied Regression Analysis*.

Fox, J. 2002. *Robust Regression : Appendix to An R and S-Plus Companion to Applied Regression*.

Ghozali, I. 2011. *Ekonometrika Teori, Konsep dan Aplikasi dengan SPSS*. Semarang : Badan Penerbit Universitas Diponegoro.

Gujarati, D.N. 1997. *Ekonometri Dasar*. Diterjemahkan oleh : Sumarno Zain. Jakarta : Erlangga. Terjemahan dari : Basic Econometrics.

Montgomery, D.C. dan Peck, E.A. 2012. *Introduction To Linier Regression Analysis. Fifth Edition.* New York : John Wiley and Sons, Inc.

Montgomery, D.C. dan Runger, G.C. 2003. *Applied Statistics and Probability for Engineers. Third Edition*. United States of America : John Wiley and Sons.

Neter, J., Wasserman, W., dan Kutner, M.H. 1988. *Applied Linier Regression Model.* USA : Richard D. Irwin. Inc.

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

Setiarini, Z., Listyani, E. 2017. *Analisis Regresi Robust Estimasi-S Menggunakan Pembobot Welsch dan Tukey Bisquare*. Jurnal Matematika. Vol. 6, No.1.

Sukirno, S. 2001. *Pengantar Teori Makroekonomi*. Jakarta : Raja Grafindo.