PERBANDINGAN PEMBOBOT *BISQUARE TUKEY* DAN PEMBOBOT *ANDREW* DALAM REGRESI *ROBUST M-ESTIMATOR*

(Skripsi)

Oleh

Ni Wayan Mega Pratiwi

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2022

ABSTRAK

PERBANDINGAN PEMBOBOT *BISQUARE TUKEY* DAN PEMBOBOT *ANDREW* DALAM REGRESI *ROBUST M-ESTIMATOR*

Oleh

NI WAYAN MEGA PRATIWI

Analisis regresi *robust* merupakan salah satu teknik analisis regresi yang digunakan untuk mengatasi data yang mengandung pencilan. Pencilan dapat menyebabkan metode kuadrat terkecil tidak dapat digunakan karena standar *error* menjadi tinggi sehingga tidak efektif dalam memodelkan data. Salah satu metode estimasi pada regresi *robust* adalah *M-Estimator*. Metode ini memiliki fungsi pembobot antara lain pembobot *Bisquare Tukey* dan pembobot *Andrew*. Penelitian ini mengaplikasikan kedua pembobot pada data tanpa pencilan dan data dengan pencilan sebesar 5%, 10%, 15%, dan 20%. Ukuran data yang digunakan sebesar 30, 60, 120, 180, 240, dan 300. Berdasarkan kriteria MSE (*Mean Square Error*), diperoleh bahwa pada saat data tidak mengandung pencilan, metode kuadrat terkecil lebih baik digunakan dalam memodelkan data daripada regresi *robust* estimasi M. Namun pada saat data dikontaminasi pencilan 5%, 10%, dan 15%, pembobot *Bisquare Tukey* lebih baik dalam memodelkan data daripada pembobot *Andrew*. Sedangkan pada saat data dikotaminasi pencilan 20%, pembobot *Andrew* lebih baik dalam memodelkan data daripada pembobot *Bisquare Tukey*.

Kata kunci : Regresi *Robust*, Pencilan, *M-estimator*, Pembobot *Bisquare Tukey*, Pembobot *Andrew*.

ABSTRACT

COMPARISON OF TUKEY BISQUARE AND ANDREW ON ROBUST REGRESSION M-ESTIMATOR

By

NI WAYAN MEGA PRATIWI

Robust regression analysis is one of the regression analysis techniques used to deal with data containing outliers. Outliers can cause the ordinary least square to not be used because the standar error becomes high, making it ineffective in modeling data. One of the estimation methods in robust regression is the M-estimator, This method has a weighting function including Tukey Bisquare weighting and Andrew weighting. This study applies both weight to simulated data contaminated with outliers of 5%, 10%, 15%, and 20%. The data sizes used were 30, 60, 120, 180, 240, and 300. Based on the MSE (Mean Square Error) criteria, it was found that when the data did not contain outliers, the ordinary least squares is better used in modeling data than the robust regression estimation M. however, when the data is contaminated with 5%, 10%, and 15% outliers, the Bisquare Tukey weighted is better than Andrew's weight. Meanwhile, when the data is contaminated with 20% outliers, Andrew's weight is better at modeling data than Bisquare Tukey's weight.

Keywords : Robust Regression, Outliers, M-estimator, Tukey Bisquare weights, Andrew weights.

Perbandingan Pembobot *Bisquare Tukey* dan Pembobot *Andrew* dalam Regresi *Robust M-Estimator*

Oleh

NI WAYAN MEGA PRATIWI

Skripsi

Sebagai Salah Satu Syarat Untuk Mencapai Gelar
SARJANA MATEMATIKA

Pada

Jurusan Matematika

Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung

JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2022

Judul Skripsi : **PERBANDINGAN PEMBOBOT *BISQUARE TUKEY* DAN PEMBOBOT *ANDREW* DALAM REGRESI *ROBUST M ESTIMATOR***

Nama Mahasiswa : **Ni Wayan Mega Pratiwi**

Nomor Pokok Mahasiswa : **1817031025**

Progrm Studi : **Matematika**

Fakultas : **Matematika dan Ilmu Pengetahuan Alam**

MENYETUJUI

1. Komisi Pembimbing

Widiarti, S.Si., M.Si.	**Dr. Aang Nuryaman, S.Si., M.Si.**
NIP. 198005022005012003	NIP. 197403162005011001

2. Ketua Jurusan Matematika

Dr. Aang Nuryaman, S.Si., M.Si.
NIP. 197403162005011001

MENGESAHKAN

1. Tim Penguji

Ketua	: **Widiarti, S.Si., M.Si.**	....
Sekretaris	: **Dr. Aang Nuryaman, S.Si., M.Si.**	....
Penguji Bukan Pembimbing	: **Ir. Warsono, M.S. Ph.D**	....

2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam

Dr. Eng. Suripto Dwi Yuwono, S.Si., M.T
NIP: 197407052000031001

Tanggal Lulus Ujian Skripsi : **15 Desember 2022**

PERYATAAN SKRIPSI MAHASISWA

Yang bertanda tangan di bawah ini :

Nama	: **Ni Wayan Mega Pratiwi**
Nomor Pokok Mahasiswa	: **1817031025**
Jurusan	: **Matematika**
Judul Skripsi	: **Perbandingan Pembobot *Bisquare Tukey* dan Pembobot *Andrew* dalam Regresi *Robust M-Estimator***

Dengan ini menyatakan bahwa penelitian ini adalah hasil pekerjaan saya sendiri dan apabila dikemudian hari terbukti bahwa skripsi ini merupakan hasil salinan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku.

Bandar Lampung, 22 Desember 2022
Yang menyatakan,

METERAI TEMPEL
4EA48AKX169735687

Ni Wayan Mega Pratiwi
NPM. 1817031025

RIWAYAT HIDUP

Penulis bernama lengkap Ni Wayan Mega Pratiwi, lahir di Desa Rama Indra pada tanggal 20 November 2000. Merupakan anak pertama dari dua bersaudara, dari pasangan Bapak I Wayan Sadiyo dan Ibu Ni Wayan Sukerti.

Riwayat pendidikan TK Tunas Bangsa lulusan tahun 2007, Sekolah Dasar di SD Negeri 02 Rama Indra lulusan tahun 2013. SMP Negeri 01 Seputih Raman lulusan tahun 2016. SMA Negeri 01 Kotagajah lulusan tahun 2018.

Penulis melanjutkan pendidikan di Universitas Lampung jalur SNMPTN pada tahun 2018, diterima sebagai mahasiswa S1 Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam. Penulis mendapatkan beasiswa Bank Indonesia tahun 2021. Riwayat organisasi yakni Biro Kesekretariatan Himatika (2019), anggota bidang Penelitian dan Pengembangan UKM Hindu Unila (2019), sekretaris bidang Penelitian dan Pengembangan UKM Hindu Unila (2020), Bendahara Umum UKM Hindu Unila (2021), anggota Kesatuan Mahasiswa Hindu Dharma Indonesia (2019-2022), anggota divisi Kesehatan Masyarakat GenBI Unila (2021).

Pada tahun 2021 penulis melaksanakan Kuliah Praktik (KP) di Badan Pusat Statistik Kabupaten Lampung Timur selama 30 hari dan Kuliah Kerja Nyata (KKN) selama 40 hari di Fajar Mataram, Kecamatan Seputih Mataram, Kabupaten Lampung Tengah. Penulis juga mengikuti program kampus merdeka yaitu Kampus Mengajar di SD Negeri 4 Kotagajah selama kurang lebih 3 bulan.

KATA INSPIRASI

"Jika kamu selalu mengingat Aku, dengan kasih karunia-Ku kamu akan mengatasi semua rintangan dan kesulitan"
(Bhagawad Gita. 18.58)

"Rahasia untuk maju adalah memulai"
(Mark Twain)

"Apa yang rusak bisa diperbaiki, apa yang sakit bisa disembuhkan. Tidak peduli betapapun sakitnya, matahari akan terbit kembali"

(unknown)

PERSEMBAHAN

Puji dan syukur kehadirat Tuhan Yang Maha Esa atas segala berkat nikmat dan kasih-Nya skripsi ini dapat diselesaikan

Kupersembahkan karya sederhana ini kepada:

Keluarga Tercinta

Terimakasih kepada Mama, Bapak, Adik tercinta Made Oka Mahendra atas limpahan kasih sayang, pengorbanan, semangat, serta doa yang selalu tercurahkan menunggu keberhasilan dengan sabar dan penuh pengertian.

Dosen Pembimbing dan Penguji

Yang telah senantiasa mengarahkan dan memberikan motivasi kepada penulis

Almamater Kebanggaan

Universitas Lampung

SANWACANA

Puji syukur penulis ucapkan kepada Tuhan Yang Maha Esa atas berkat dan kasih-Nya sehingga penulis dapat menyelesaikan skripsi ini yang berjudul "Perbandingan Pembobot *Bisquare Tukey* dan Pembobot *Andrew* dalam Regresi *Robust M-Estimator*" disusun sebagai salah satu syarat memperoleh gelar Sarjana Matematika di Universitas Lampung.

Pada kesempatan kali ini penulis ingin menyampaikan terimakasih yang sebesar-besarnya kepada:

1. Ibu Widiarti, S.Si.,M.Si. selaku pembimbing I yang selalu bersedia memberikan bimbingan, masukan, serta arahan selama penulisan skripsi ini.
2. Bapak Dr. Aang Nuryaman, M.Si. selaku pembimbing II yang telah memberikan bimbingan, masukan, serta arahan selama penulisan skripsi ini.
3. Bapak Ir.Warsono, M.S., Ph.D selaku dosen pembahas yang telah memberi evaluasi, masukan, serta arahan selama penulisan skripsi ini.
4. Ibu Widiarti, S.Si., M.Si. selaku pembimbing akademik yang memberikan bimbingan dan arahan selama perkuliahan.
5. Bapak Dr. Aang Nuryaman, S.Si., M.Si. selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
6. Bapak Dr. Eng. Suripto Dwi Yuwono, S.Si., M.T selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
7. Seluruh dosen dan staf Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
8. Mama, Bapak, Adik yang selalu memberi dukungan semangat, motivasi dan perhatian serta selalu mendoakan yang terbaik bagi penulis.

9. Bli Aryudi, Mbak Nopi, Raja, Putri, Danis, serta seluruh *crew* di Kantor Wisma Artha yang telah memberi dukungan dan semangat selama perkuliahan.
10. The Netral: Inka, Rika, Lela, Oktin, Rendi, Zainal yang selalu menemani dan membantu dalam masa-masa perkuliahan.
11. Teman seperjuangan sekaligus sahabat terbaik penulis Kadek Fani Sugiyanti, Inka Krysti Meina Br Peranginangin, dan Ela Desiawati yang selalu memberikan semangat selama perkuliahan, dan selalu menjadi pendengar yang baik mengenai keluh kesah penulis.
12. Teman-teman seperbimbingan skripsi Anis, Inka, Yuni, Hilda, Ria, dan Dora. Terimakasih atas segala dukungan dan kerja samanya selama penyusunan skripsi ini.
13. Seluruh teman-teman mahasiswa Jurusan Matematika angkatan 2018.
14. Kepada seluruh pihak yang tidak dapat penulis sebutkan satu persatu.
15. Diri penulis sendiri, terimakasih sudah selalu berjuang sampai titik ini, semangat untuk proses selanjutnya.

Penulis berharap semoga skripsi ini bermanfaat, penulis juga menyadari dalam penyusunan skripsi ini masih memiliki banyak kekurangan. Oleh karena itu, kritik dan saran yang membangun sangat dibutuhkan, guna penyempurnaan skripsi ini.

Bandar Lampung, 15 Desember 2022
Penulis,

Ni Wayan Mega Pratiwi

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

I. PENDAHULUAN

1.1 Latar Belakang dan Masalah

Analisis regresi merupakan analisis yang digunakan untuk mengetahui hubungan antara variabel terikat (*dependent*) dengan satu atau lebih variabel bebas (*independent*). Metode yang biasa digunakan untuk mengestimasi parameter regresi adalah Metode Kuadrat Terkecil (MKT). Menurut Myers, dkk. (2010), estimator yang dihasilkan MKT akan bersifat tak bias dan efisien (*Best Linear Unbiased Estimator*/BLUE) jika komponen sisaan atau galat memenuhi semua asumsi dalam regresi. Asumsi-asumsi tersebut yaitu asumsi normalitas, asumsi homoskedastisitas, asumsi non-autokorelasi, dan asumsi non-multikolinieritas. Pada beberapa kasus dijumpai bahwa terdapat asumsi yang tidak terpenuhi, sehingga mengakibatkan hasil estimasi pada MKT kurang baik. Salah satu asumsi yang tidak terpenuhi adalah asumsi normalitas. Asumsi normalitas tidak terpenuhi dapat dikarenakan adanya pencilan atau *outlier* pada data pengamatan.

Menurut Montgomery, dkk. (2012), pencilan atau *outlier* adalah data yang polanya menyimpang dari pola umum data. Tindakan untuk membuang begitu saja suatu pencilan atau *outlier* bukanlah suatu tindakan yang bijaksana, karena pencilan dapat memberikan informasi yang cukup berarti. Pendeteksian outlier dapat dilakukan dengan beberapa cara, salah satunya adalah dengan menghitung nilai DFFITS (*Difference in Fit Standardized*). Pencilan menyebabkan metode MKT tidak dapat digunakan karena memiliki standar *error* yang tinggi sehingga memiliki sifat yang bias dan tidak efisien dalam memodelkan data.

Oleh karena itu, diperlukan sebuah estimasi yang *robust* atau tahan terhadap pencilan. Regresi tersebut biasa disebut dengan regresi *robust*. Beberapa metode dalam regresi *robust* diantaranya *Least Median of Square* (LMS), *Least Trimmed Squares* (LTS), *M-estimation, S-Estimation* dan *MM-estimation* (Rousseeuw dan Leroy, 1987).

Salah satu dari metode *robust* yang sering digunakan adalah estimasi M, karena hasilnya lebih teliti (Pradewi dan Sudarno, 2012). Penelitian oleh Sinay dan Talakua (2015) mengkaji regresi *robust* model *M-Estimator* dengan estimator *Huber* dan *Bisquare Tukey* dapat mengatasi masalah *outlier* pada OLS. Model regresi linear *robust M-estimator* dengan pembobot *Huber* dan *Bisquare Tukey* lebih baik dibandingkan model OLS. Lebih jauh lagi, berdasarkan nilai kebaikan model yaitu ditinjau dari nilai R^2 dan MSE diperoleh bahwa model *Bisquare Tukey* lebih baik dibandingkan model *Huber*.

Pada penelitian lainnya, Deria, dkk. (2019) mengkaji regresi robust M-estimator dengan fungsi bobot *Andrew*, *Ramsay*, dan *Welsch* yang dapat digunakan untuk mengatasi keberadaan *outlier* dengan studi kasus kemiskinan di Jawa Tengah tahun 2017. Kemiskinan di Jawa Tengah dipengaruhi oleh jumlah pengangguran, jumlah penduduk, tingkat partisipasi sekolah, Indeks Pembangunan Manusia (IPM), dan inflasi. Penelitian ini menyimpulkan bahwa model regresi *robust* terbaik adalah regresi *robust* M-estimation dengan fungsi bobot *Andrew*.

Berdasarkan uraian tersebut, penulis tertarik untuk melakukan perbandingan regresi *robust* estimasi M dengan menggunakan pembobot *Bisquare Tukey* dan pembobot *Andrew*. Perbandingan pembobot akan disimulasikan pada data simulasi dengan parameter yang diperoleh dari data kemiskinan di Indonesia tahun 2021, dengan berbagai ukuran sampel dan persentase pencilan yang bervariasi, kemudian akan dilakukan perbandingan keefektifan yang ditinjau dari MSE (*Mean Square Error*) sehingga dapat memperoleh pembobot terbaik.

1.2 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Mengkaji regresi *robust* dengan estimasi M menggunakan pembobot *Bisquare Tukey* dan *Andrew*.
2. Menentukan pembobot terbaik dalam pemodelan regresi *robust* pada estimasi-M yang ditinjau dari MSE (*Mean Square Error*).

1.3 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah menambah wawasan dan pengetahuan tentang regresi *robust* estimasi-M dengan pembobot *Bisquare Tukey* dan pembobot *Andrew*. Serta diharapkan dapat menjadi bahan bacaan dan referensi bagi pembaca dalam menentukan pembobot terbaik pada regresi *robust* estimasi M.

II. TINJAUAN PUSTAKA

2.1 Analisis Regresi Linier

Analisis regresi merupakan teknik statistika yang digunakan untuk memodelkan hubungan antara variabel terikat (variabel respon) dengan satu atau lebih variabel bebas (variabel prediktor). Menurut Freund, dkk. (2006), model analisis regresi dapat digunakan untuk membuat kesimpulan tentang hubungan variabel respon dengan variabel bebas, hubungan yang dihasilkan kemudian dapat digunakan untuk memprediksi atau menjelaskan variabel respon. Menurut Chatterjee dan Hadi (2006), variabel respon dinotasikan sebagai Y dan variabel prediktor dinotasikan dengan $X_1, X_2, \cdots, X_p$ dimana p untuk menunjukkan jumlah dari variabel prediktor.

Terdapat dua jenis analisis regresi linear, yaitu regresi linear sederhana dan regresi linear berganda. Analisis regresi linier sederhana merupakan analisis regresi yang melibatkan satu variabel bebas dan satu variabel tak bebas. Model regresi linier sederhana dapat dinyatakan sebagai berikut (Montgomery,dkk., 2012):

$$Y_i = \beta_o + \beta_1 X_i + e_i, \qquad i = 1,2,\cdots,n \tag{2.1}$$

dengan:

Y_i	: variabel tak bebas untuk pengamatan ke-i,
X_i	: variabel bebas untuk pengamatan ke-i,
β_0, β_1	: parameter regresi,
e_i	: galat untuk pengamatan ke-i.

Analisis regresi linier sederhana ini dilakukan pendugaan yaitu β_0 dan β_1, persamaan linier pada pendugaan garis regresi linier ditulis dalam bentuk:

$$\hat{y}_i = b_0 + b_1 x_i \tag{2.2}$$

dengan:

$\hat{y}_i$: nilai dugaan variabel tak bebas pengamatan ke-i,

b_0 : titik potong garis regresi pada sumbu y atau nilai dugaan $\hat{y}$ bila $x = 0$,

b_1 : gradien garis regresi atau perubahan nilai dugaan $\hat{y}$ persatuan perubahan dari nilai x,

x_i : nilai variabel bebas pada pengamatan ke-i.

Sedangkan analisis regresi linear berganda merupakan metode regresi yang melibatkan satu variabel tak bebas dan beberapa variabel bebas. Model regresi linier berganda dapat dinyatakan sebagai berikut:

$$Y_i = \beta_o + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + e_i \tag{2.3}$$

dengan:

$i = 1,2,\cdots,n$

$k = 1,2,\cdots,n$

Y_i : variabel tak bebas untuk pengamatan ke-i,

$X_{i1,}\cdots,X_{ik}$: variabel bebas ke-k untuk pengamatan ke-i,

β_k : parameter regresi ke-k,

e_i : galat untuk pengamatan ke-i.

Persamaan (2.3) dapat ditulis dalam bentuk matriks sebagai berikut:

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{e} \tag{2.4}$$

dengan:

$$\boldsymbol{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad ; \quad \boldsymbol{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} \quad ; \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \quad ; \boldsymbol{e} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

2.2 Metode Kuadrat Terkecil

Parameter dalam regresi linier $\beta_1, \beta_2, \cdots, \beta_k$ belum diketahui, oleh karena itu diperlukan metode pendugaan untuk mengestimasi parameter tersebut. Menurut Montgomery, dkk. (2012), metode yang umum digunakan untuk mengestimasi parameter model regresi tersebut adalah dengan meminimumkan jumlah kuadrat residual atau yang dikenal dengan Metode Kuadrat Terkecil (MKT). MKT merupakan penduga yang terbaik karena bersifat tak bias dan konsisten, metode ini menghasilkan ragam yang minimum bagi parameter regresi. Misalkan terdapat *k* parameter dan *n* pengamatan maka akan memperoleh model sebagai berikut:

$$\begin{aligned} Y_1 &= \hat{\beta}_0 + \hat{\beta}_1 X_{11} + \cdots + \hat{\beta}_k X_{1k} + e_1 \\ Y_2 &= \hat{\beta}_0 + \hat{\beta}_1 X_{21} + \cdots + \hat{\beta}_k X_{2k} + e_2 \\ &\vdots \\ Y_n &= \hat{\beta}_0 + \hat{\beta}_1 X_{n1} + \cdots + \hat{\beta}_k X_{nk} + e_n \end{aligned}$$

Yang dapat ditulis secara singkat dalam bentuk matriks yaitu:

$$\boldsymbol{Y} = \boldsymbol{X}\hat{\boldsymbol{\beta}} + \boldsymbol{e} \tag{2.5}$$

Pendugaan MKT untuk kasus n amatan dapat diperoleh dengan meminimumkan

$$\sum_{i=1}^{n} e_i^2 = \sum \left[\left(Y_i - \left(\hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \cdots + \hat{\beta}_p X_{ip-1}\right)\right)\right]^2 \tag{2.6}$$

Dengan $\sum e_i^2$ merupakan jumlah kuadrat galat.

Notasi matriks untuk meminimumkan $\mathbf{e}^{\mathbf{T}}\mathbf{e}$ dari Persamaan (2.4) diperoleh:

$$\boldsymbol{e} = \boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}} \tag{2.7}$$

Oleh karena itu,

$$\begin{aligned} \boldsymbol{e}^{\boldsymbol{T}}\boldsymbol{e} &= \left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\right)^{\boldsymbol{T}}\left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\right) \\ &= \boldsymbol{Y}^{\boldsymbol{T}}\boldsymbol{Y} - \boldsymbol{2}\widehat{\boldsymbol{\beta}}^{\boldsymbol{T}}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{Y} + \widehat{\boldsymbol{\beta}}^{\boldsymbol{T}}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X}\widehat{\boldsymbol{\beta}}. \end{aligned} \tag{2.8}$$

Untuk menaksir parameter $\beta_1, \beta_2, \dots, \beta_k$ maka $\sum_{i=1}^{n} e_i^2$ harus sekecil mungkin. Hal tersebut dapat dicapai dengan menurunkan persamaan $\boldsymbol{e}^{\boldsymbol{T}}\boldsymbol{e}$ terhadap $\hat{\beta}$ dan membuatnya sama dengan nol, sehingga diperoleh:

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X})^{-\boldsymbol{1}}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{Y} \tag{2.9}$$

2.3 Uji Asumsi Analisis Regresi Linier

Uji asumsi merupakan uji yang dilakukan sebelum melakukan analisis data. Apabila seluruh asumsi dapat terpenuhi maka akan menghasilkan model regresi yang memenuhi kriteria BLUE (*Best Liniar Unbiased Estimator*) (Myers,dkk., 2010).

Uji asumsi klasik pada analisis regresi linier berganda, terdiri dari:

2.3.1 Uji Normalitas

Uji normalitas digunakan untuk mengetahui apakah residual dalam model regresi berdistribusi normal atau tidak. 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*.

Uji hipotesis

i. H_0 : residu berdistribusi normal

H_1 : residu tidak berdistribusi normal

ii. Taraf signifikansi α

iii. Menghitung statistik uji

$$D_{hitung} = Max|F_n(x) - F^*(x)| \tag{2.10}$$

dengan $F_n(x)$ adalah fungsi distribusi kumulatif empiris, dan $F^*(x)$ adalah fungsi distribusi kumulatif normal.

iv. Menentukan daerah kritis

H_0 ditolak jika $D_{hitung} > D_{(1-\propto,n)}$ dengan $D_{(1-\propto,n)}$ didapat dari tabel *Kolmogorov-Smirnov.*

v. Menentukan kesimpulan.

2.3.2 Uji Multikolinieritas

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 :

$$VIF_j = \frac{1}{(1-R_j^2)} \tag{2.11}$$

dengan:

j : 1,2,…,k

k : banyaknya peubah prediktor,

R_j^2 : koefisien determinasi dari persaman regresi dengan X_j sebagai peubah respon dan X lainnya sebagai peubah prediktor.

Jika nilai $VIF < 10$ maka asumsi non-multikolinieritas dipenuhi.

2.3.3 Uji Autokorelasi

Model regresi linear klasik mengasumsikan bahwa autokorelasi seperti tidak terdapat dalam *error* ε_i. 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*.

$$d = \frac{\sum (e_i - e_{i-1})^2}{\sum e_i^2} \tag{2.12}$$

Uji hipotesis

i. H_0: $\rho = 0$, artinya tidak terdapat autokorelasi
   H_1: $\rho \neq 0$, artinya terdapat autokorelasi
ii. Taraf signifikasi α
iii. Keputusan Uji
   Jika dapat ditunjukan $d_u < d < 4 - d_u$, dimana
   $d = d$-hitung *Durbin-Watson*
   d_u = nilai kritis untuk batas atas dari tabel *Durbin-Watson*
   Maka dapat disimpulkan bahwa model tidak terjadi auotokorelasi

2.3.4 Uji Homogenitas

Uji homogenitas bertujuan untuk menguji apakah dalam model regresi terjadi ketidaksamaan varian dari residual satu pengamatan ke pengamatan yang lain. Uji homogenitas variansi residu dilakukan dengan uji *Breusch Pagan*.

Uji hipotesis

i. H_0: model regresi bersifat homoskedastisitas
   H_1: model regresi tidak bersifat homoskedastisitas
ii. Taraf signifikasi α
iii. Keputusan Uji: menolak H_0 apabila $p_{value} < \alpha$

2.4 Uji Hipotesis dalam Regresi Linier Berganda

Uji hipotesis merupakan proses yang dilakukan dalam rangka mengambil keputusan dari dua hipotesis yang berlawanan. Kedua hipotesis dirumuskan sedemikian rupa sehingga masing-masing hipotesis merupakan negasi dari hipotesis yang lainnya. Rumusan hipotesis mengakibatkan salah satu akan selalu bernilai benar dan hipotesis lainnya akan selalu bernilai salah. Kedua hipotesis tersebut dinamakan hipotesis nol dan hipotesis alternatif. Menentukan hipotesis nol dan hipotesis alternatif, merupakan langkah yang sangat penting.

Pada uji hipotesis dalam regresi linier berganda dilakukan beberapa uji yaitu:

2.4.1 Uji F

Uji F merupakan uji signifikasi kecocokan model yang bertujuan untuk melihat ada atau tidaknya hubungan linier antar variabel secara keseluruhan. Berikut ini merupakan langkah-langkah uji F (Montgomery dan Runger, 2003):

a. Hipotesis

$H_0: \beta_1 = \beta_2 = \cdots = \beta_k = 0$

H_i:terdapat $\beta_j \neq 0$, dengan $j = 1,2, ...,$k

b. Statistik Uji

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

c. Kriteria Uji

H_0ditolak jika $F_{hitung} > F_{(\alpha,k,n-k-1)}$ atau $p - value < \alpha$

2.4.2 Uji Parsial (t)

Uji Parsial merupakan uji koefisien regresi secara individual digunakan untuk menguji ada atau tidaknya pengaruh yang signifikan antara masing-masing variabel predictor terhadap terhadap model regresi linier. Berikut ini merupakan langkah-langkah pengujiannya:

a. Hipotesis

$H_0: \beta_j = 0$

$H_1: \beta_j \neq 0$ dengan $j = 1,2, ...,$k

b. Statistik Uji

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

c. Kriteria Uji

H_0ditolak jika $|t_{hitung}| > t_{(\frac{\alpha}{2},n-k-1)}$atau $p - value < \alpha$

2.5 Ukuran Kecocokan Model (*Goodness Of Fit*)

Ukuran kecocokan model (*Goodness Of Fit*) digunakan untuk memperoleh gambaran apakah model yang dipilih sesuai atau tidak.

2.5.1 Koefisien Determinasi yang disesuaikan (*R^2*)

Menurut Chatterjee dan Hadi (2006), koefisien determinasi (R^2) merupakan alat untuk mengukur bagaimana kemampuan variabel prediktor *X* memperhitungkan atau menentukan respons variabel *Y*.

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \qquad (2.13)$$

2.5.2 MSE (*Mean Square Error*)

Mean square error merupakan salah satu yang digunakan untuk melihat kebaikan pendugaan parameter dari regresi. MSE menghitung nilai kesalahan dari rata-rata penduga. Model persamaan yang baik adalah model yang memiliki nilai MSE yang kecil, semakin kecil nilai dari MSE maka pendugaan parameter regresi semakin baik. Nilai MSE diperoleh dari nilai jumlah kuadrat galat dibagi dengan db jumlah kuadrat sisaan. Nilai MSE dapat dilakukan dengan perhitungan berikut ini:

$$MSE = \frac{\sum(y_i - \hat{y}_i)^2}{n-2} \qquad (2.14)$$

(Chatterjee dan Hadi, 2006)

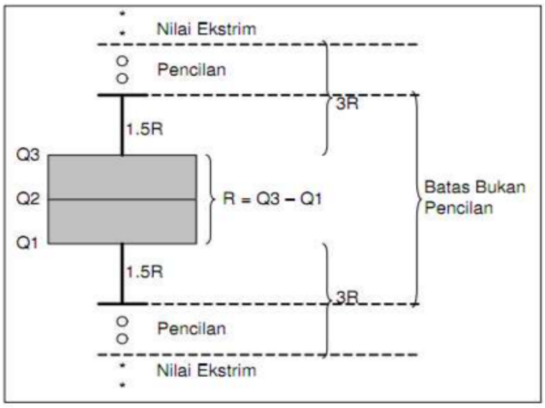
2.6 Pencilan

Menurut Neter, dkk. (1997), pencilan merupakan data amatan yang berada jauh dari amatan lainnya. Pencilan dalam data akan mengganggu proses analisis data sehingga dapat memberikan pengaruh yang besar terhadap pendugaan parameter regresi. Keberadaan pencilan akan menimbulkan suatu masalah dalam metode kuadrat terkecil.

Pada umumnya pencilan atau *outlier* mempunyai sisaan atau *error* berjarak tiga simpangan baku. Pencilan dapat menyebabkan munculnya nilai rata-rata dan simpangan baku yang tidak konsisten terhadap mayoritas data. Selain itu, estimasi koefisien garis regresi yang diperoleh tidak tepat, dan pada beberapa analisis *inferensia* dapat menyebabkan kesalahan dalam pengambilan keputusan dan kesimpulan (Sembiring, 2003).

2.6.1 Identifikasi Pencilan dengan *Boxplot*

Pendeteksian pencilan dapat dilakukan dengan menggunakan beberapa metode, salah satunya adalah dengan metode *boxplot*. Menurut Febrianto, dkk (2018), *boxplot* merupakan metode yang mempergunakan nilai kuartil dan jangkauan untuk mendeteksi adanya *outlier* (pencilan), sehingga pada metode ini dapat mengetahui adanya *outlier* untuk masing-masing variabel. *Boxplot* mempergunakan nilai kuartil dan jangkauan. Kuartil 1, 2, dan 3 akan membagi sebuah urutan data menjadi empat bagian. Jangkauan (IQR, *Interquartile Range*) didefinisikan $IQR = Q3 - Q1$ sebagai selisih kuartil 1 terhadap kuartil 3. Adapun pencilan dapat ditentukan dari nilai yang kurang dari $1{,}5 * IQR$ terhadap kuartil 1 dan nilai yang lebih dari $1{,}5 * IQR$ terhadap kuartil 3 (Dawson, 2011).



Gambar 1. Skema Identifikasi Pencilan dengan Metode *Boxplot*.

2.6.2 Identifikasi Pencilan dengan Diagram Pencar (*Scatter Plot*)

Metode lain yang digunakan untuk mengidentifikasi adanya pencilan (*outlier)* yaitu metode grafis atau gambar, seperti diagram pencar (*scatter plot)*. Metode ini tanpa melibatkan perhitungan yang rumit. Kelemahan metode ini adalah keputusan yang memperlihatkan data tersebut merupakan pencilan atau tidak bergantung pada kebijakan dari peneliti, karena hanya mengandalkan visualisasi gambar.

Metode *scatterplot* dilakukan dengan cara memplot data dengan observasi ke-i $(i = 1,2,\cdots,n)$. Selain itu, jika sudah mendapatkan model regresi maka dapat dilakukan dengan cara memplot antara residual (e) dengan nilai prediksi $(\hat{Y})$. Jika terdapat data yang berada jauh dari pola kumpulan data keseluruhan maka hal ini mengidentifikasi sebagai pencilan.

2.6.3 Identifikasi Pencilan dengan Uji DFFITS (*Difference in Fit Standardized*)

Metode lain yang dapat digunakan untuk mendeteksi pencilan adalah uji DFFITS (*Difference in Fit Standardized*). Rumus DFFITS didefinisikan sebagai berikut:

$$(DFFITS_i) = t_i \left(\frac{h_{ii}}{1-h_{ii}}\right)^{\frac{1}{2}} \tag{2.15}$$

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

$$t_i = \frac{e_i}{s_i\sqrt{(1-h_{ii})}}, i = 1,2,\cdots,n$$

$$s_i = \sqrt{\frac{(n-p)s^2 - e_i^2/(1-h_{ii})}{n-p-1}}$$

$$s^2 = \sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)}{n - p}$$

Dengan $e_i$adalah residual ke-i, h_{ii} adalah elemen-elemen diagonal dari matriks **H. H** adalah matriks n x n

$$\boldsymbol{H} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$$

Suatu data disebut pencilan jika nilai $|\text{DFFITS}| > 1$ untuk gugus data yang berukuran kecil sampai sedang ($n \leq 30$) dan nilai $|\text{DFFITS}| > 2\sqrt{p/n}$ untuk gugus data yang berukuran besar ($n > 30$), dengan $p = k + 1$, dan n adalah banyaknya observasi (Pitselis, 2013).

2.7 Regresi *Robust*

Regresi *robust* merupakan metode regresi yang digunakan ketika dalam menduga parameter regresi menggunakan metode kuadrat terkecil distribusi dari galat tidak normal dan atau adanya beberapa pencilan yang berpengaruh pada model. Metode kuadrat terkecil memberikan bobot yang sama pada setiap pengamatan untuk mendapatkan estimasi parameter. Sedangkan regresi *robust* memungkinkan pengamatan dibobot secara tidak sama. Pada dasarnya, pengamatan yang menghasilkan residu besar dibobotkan dengan metode estimasi regresi *robust*. Sejumlah metode tersedia, secara umum, metode regresi *robust* memerlukan komputasi yang jauh lebih banyak daripada metode kuadrat terkecil (Draper dan Smith, 1998).

Metode dalam analisis regresi *robust* yang dapat digunakan untuk menangani data pencilan yaitu metode M , *Least Trimmed Square* (LTS), *Scale* (S), *Method of Moment* (MM) serta *Least Median of Squares* (LMS) atau metode Kuadrat Median Terkecil (Rousseeuw dan Leroy, 1987). Penelitian ini akan menggunakan regresi *robust* metode-M.

2.8 Pembobot *Bisquare Tukey*

Data yang mengandung pencilan akan menghasilkan galat atau *error* yang besar dan untuk mengatasi hal tersebut umumnya akan diberikan bobot. Regresi robust estimasi M mempunyai beberapa fungsi pembobotan, diantaranya yaitu pembobotan *Bisquare Tukey* dan pembobotan *Andrew*. Penggunaan fungsi pembobot *Bisquare Tukey* dan *Andrew* adalah untuk menghasilkan nilai skala pembobot dengan cara melakukan iterasi hingga estimator yang diperoleh konvergen.

Menurut Maronna,dkk. (2019), fungsi objektif untuk pembobot *Bisquare Tukey* yaitu:

$$\rho(u_i) = \begin{cases} \frac{k^2}{6}\left\{1 - \left[1 - \left(\frac{u_i}{k}\right)^2\right]^3\right\} & \text{untuk } |u_i| \leq k \\ \frac{k^2}{6}, & \text{untuk } |u_i| > k \end{cases} \tag{2.16}$$

Dengan fungsi pembobot yaitu:

$$w(u_i) = \begin{cases} \left[1 - \left(\frac{u_i}{k}\right)^2\right]^2 & \text{untuk } |u_i| \leq k \\ 0, & \text{untuk } |u_i| > k \end{cases} \tag{2.17}$$

Nilai k untuk estimator *Bisquare Tukey* disebut *tuning constans.* Nilai k yang optimal untuk *Bisquare Tukey* adalah 4,685.

2.9 Pembobot *Andrew*

Menurut Draper dan Smith (1998), fungsi objektif untuk pembobot *Andrew* adalah sebagai berikut:

$$\rho(u_i) = \begin{cases} k\left(1 - \cos\left(\frac{u_1}{k}\right)\right) & \text{untuk } |u_i| \leq k\pi \\ 2k\ , & \text{untuk } |u_i| > k\pi \end{cases} \tag{2.18}$$

Fungsi pembobot *Andrew* adalah sebagai berikut:

$$w(u_i) = \begin{cases} \frac{sin\left(\frac{u_i}{k}\right)}{\frac{u_i}{k}}, & \text{untuk } |u_i| \leq k\pi \\ 0, & \text{untuk } |u_i| > k\pi \end{cases} \tag{2.19}$$

Dimana nilai k pada pembobot *Andrew* disebut *tunning constant* dengan nilai $k = 1{,}339$.

2.10 Regresi *Robust M-Estimator*

Estimasi regresi *robust* yang paling luas digunakan adalah Estimasi-M, yang diperkenalkan oleh Huber dan merupakan pendekatan yang paling sederhana baik secara teori maupun perhitungan. Regresi *robust* estimasi-M pada prinsipnya adalah meminimumkan suatu fungsi objektif ρ.

$$\sum_{i=1}^{n} \rho(u_i) = \sum_{i=1}^{n} \rho\left(\frac{e_i}{\hat{\sigma}}\right) = \sum\nolimits_{i=1}^{n} \rho\left(\frac{y_i - \sum_{j=0}^{k} X_{ij}\beta_j}{\hat{\sigma}}\right) \tag{2.20}$$

Dimana $\rho(u_i)$ adalah fungsi simetris dari residual atau fungsi yang memberikan kontribusi pada masing-masing residual pada fungsi objektif. Pada umumnya suatu estimasi skala *robust* perlu diestimasi dan $\hat{\sigma}$ adalah skala estimasi *robust*. Untuk mencari $\hat{\sigma}$ pada regresi *robust* dengan estimasi-M yang sering digunakan yaitu persamaan:

$$\hat{\sigma} = \frac{median|e_i - median(e_i)|}{0{,}6745} \tag{2.21}$$

Jika $\psi = \rho'$ merupakan turunan dari ρ. Maka untuk meminimumkan persamaan (2.20) diperlukan turunan parsial pertama dari ρ terhadap β_j dimana $j = 0,1,2, \dots, k$ harus disamadengankan nol, sehingga akan menghasilkan:

$$\sum_{i=1}^{n} X_{ij}\psi\left(\frac{e_i}{\hat{\sigma}}\right) = \sum_{i=1}^{n} X_{ij}\psi\left(\frac{y_i - \sum_{j=0}^{k} X_{ij}\beta_j}{\hat{\sigma}}\right) = 0 \tag{2.22}$$

Dimana X_{ij}adalah observasi ke-*i* pada parameter ke-*j*. ψ merupakan fungsi *influence* yang digunakan untuk memperoleh bobot. Fungsi pembobot didefinisikan seperti berikut:

$$W(u_i) = \frac{\psi\left(\frac{e_i}{\hat{\sigma}}\right)}{\left(\frac{e_i}{\hat{\sigma}}\right)}\mathrm{n}$$

Penelitian ini akan menggunakan fungsi pembobot *Bisquare Tukey* pada Persamaan (2.17) dan pembobot *Andrew* pada Persamaan (2.19). Persamaan (2.22) akan menjadi:

$$\sum_{i=1}^{n} X_{ij}\, W\left(\frac{y_i - \sum_{j=0}^{k} X_{ij}\beta_j}{\hat{\sigma}}\right) = 0 \qquad (2.23)$$

Persamaan (2.23) jika dibuat dalam bentuk matriks menjadi:

$$\boldsymbol{X'WX\beta} = \boldsymbol{X'Wy} \qquad (2.24)$$

Persamaan (2.24) tersebut akan meminimumkan persamaan

$$\sum_{i=1}^{n} W_i(y_i - \hat{y}_i)^2$$

Persamaan (2.24) disebut juga sebagai kuadrat terkecil terboboti (*weighted least square*). *Weighted least square* digunakan sebagai alat untuk menghitung estimasi-M. Sehingga parameter penduga menjadi:

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X'WX})^{-\mathbf{1}}\boldsymbol{X'Wy} \qquad (2.25)$$

Menduga parameter diperlukan solusi iterasi yang disebut IRLS (*Iteratively reweighted least squares*). Langkah-langkah IRLS yang dilakukan dalam mengestimasi parameter dengan estimasi-M adalah:

1. Menaksir $\boldsymbol{\beta}$ awal yaitu $\widehat{\boldsymbol{\beta}}_{(0)}$ dengan menggunakan metode kuadrat terkecil.
2. Menghitung nilai residual $e_i = y_i - \hat{y}_i$.

3. Menghitung nilai

$$\hat{\sigma} = \frac{median|e_i - median(e_i)|}{0{,}6745}$$

4. Menghitung nilai $u_i = \frac{e_i}{\hat{\sigma}}$

5. Menghitung nilai bobot (W_i) menggunakan fungsi pembobot:

   a. Pembobot *Bisquare Tukey*

$$w(u_i) = \begin{cases} \left[1 - \left(\frac{u_i}{k}\right)^2\right]^2 & \text{untuk } |u_i| \leq k \\ 0, & \text{untuk } |u_i| > k \end{cases}$$

   b. Pembobot *Andrew*

$$w(u_i) = \begin{cases} \frac{sin\left(\frac{u_i}{k}\right)}{\frac{u_i}{k}}, & \text{untuk } |u_i| \leq k\pi \\ 0, & \text{untuk } |u_i| > k\pi \end{cases}$$

6. Menghitung $\hat{\boldsymbol{\beta}}$ menggunakan kuadrat terkecil terboboti berdasarkan nilai bobot $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{W}\boldsymbol{X})^{-\mathbf{1}}\boldsymbol{X}'\boldsymbol{W}\boldsymbol{y}$.

W adalah matriks diagonal berukuran $n \times n$ dengan elemen-elemen diagonalnya $W_{1,1}, W_{1,2}, W_{1,3} \cdots W_{1,n}$. Dugaan parameter $\hat{\boldsymbol{\beta}}$ dari tahap 6 digunakan menjadi $\hat{\boldsymbol{\beta}}_{(0)}$ dalam tahap 1 lalu mengulangi sampai langkah ke-6. Tahapan ini dilanjutkan sampai diperoleh penduga parameter yang konvergen (Pitselis, 2013).

III. METODOLOGI PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian ini dilakukan pada semester ganjil tahun ajaran 2022/2023, bertempat di Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung.

3.2 Data Penelitian

Data yang digunakan pada penelitian ini merupakan data simulasi yang dibangkitkan dengan parameter awal dari data sekunder yang diperoleh dari laman web resmi Badan Pusat Statistik bps.go.id berupa data IPM sebagai variabel bebas X dan data jumlah penduduk miskin sebagai variabel respon Y. Data yang disimulasikan berukuran 30, 60, 120, 180, 240, dan 300. Pencilan disimulasikan pada variabel respon Y sebesar 10% dan 20% untuk ukuran data 30, sedangkan untuk ukuran data 60, 120, 180, 240, dan 300 disimulasikan dengan pencilan sebesar 5%, 10%, 15%, dan 20% menggunakan perangkat lunak R-1.4.3.

3.3 Metode Penelitian

Penelitian ini dilakukan dengan menggunakan studi pustaka yaitu mempelajari berbagai sumber guna menunjang proses penelitian dan studi simulasi untuk menentukan pembobot terbaik antara pembobot *Bisquare Tukey* dan pembobot

Andrew dalam mengatasi data yang mengandung pencilan. Adapun langkah-langkah yang dilakukan dalam penelitian ini adalah sebagai berikut:

1. Mengumpulkan data berupa data IPM sebagai variabel X dan data jumlah penduduk miskin di Indonesia pada tahun 2021 sebagai variabel Y.
2. Menghitung nilai parameter μ dan σ pada variabel bebas X
3. Membangkitkan galat dari sebaran normal $N(0, 1)$ berukuran 30, 60, 120, 180, 240, dan 300.
4. Membangkitkan variabel bebas X dengan parameter yang telah diperoleh pada langkah 2.
5. Menetapkan nilai dari β_0 dan β_1, nilai ini diperoleh dari hasil estimasi regresi linier menggunakan metode kuadrat terkecil dari data sekunder X dan Y yang sudah dikumpulkan.
6. Dari nilai-nilai yang sudah diketahui maka dicari nilai Y_i dari model regresi yaitu: $Y_i = \beta_0 + \beta_1 X_1 + e_i$
7. Mengkontaminasi variabel respon *Y* dengan pencilan sebesar 10% dan 20% pada ukuran data 30, sedangkan variabel respon *Y* dikontaminasi dengan pencilan sebesar 5%, 10%, 15%, dan 20% pada ukuran data 60, 120, 180, 240, dan 300 dengan 1000 kali pengulangan yang diikuti analisis regresi *robust* estimasi M dengan menggunakan masing-masing pembobot *Bisquare Tukey* dan pembobot *Andrew*. Pencilan diberikan dengan cara menambahkan nilai maksimum dari variabel respon Y dengan lebih dari tiga kali standar deviasi yang bersesuaian.

$$\text{Pencilan}(Y) \geq \max(Y) + 3 * \text{stdev}(Y)$$

8. Melakukan pendeteksian pencilan dengan *boxplot*, *scatterplot* dan *uji DFFITS.*
9. Melakukan pengujian normalitas menggunakan uji *Kolmogorov-Smirnov.*
10. Menghitung nilai estimasi $\hat{\beta}_0$ dan $\hat{\beta}_1$ dengan menggunakan metode regresi *robust* estimasi M dengan pembobot *Bisquare Tukey* dan pembobot *Andrew*.
11. Menghitung nilai MSE (*Mean Square Error)* dari koefisien regresi *robust* dengan masing masing pembobot.
12. Melakukan perbandingkan rata-rata nilai MSE dari analisis regresi *robust* estimasi M menggunakan masing-masing pembobot.

13. Melakukan analisis dari hasil yang diperoleh
14. Membuat kesimpulan untuk memperoleh pembobot terbaik.

.

V. KESIMPULAN

Kesimpulan yang diperoleh dari penelitian ini adalah:

1. Regresi *robust* estimasi M dengan pembobot *Bisquare Tukey* dan pembobot *Andrew* memiliki fungsi objektif yang berbeda, sehingga menghasilkan nilai parameter $\hat{\beta}$ yang berbeda.
2. Berdasarkan hasil dan pembahasan diperoleh bahwa metode kuadrat terkecil lebih baik digunakan daripada regresi *robust* saat data tidak mengandung pencilan. Sedangkan, regresi *robust* estimasi M dengan pembobot *Bisquare Tukey* dan pembobot *Andrew* sama-sama lebih baik dalam memodelkan data daripada metode kuadrat terkecil untuk data yang dikontaminasikan pencilan sebesar 5%, 10%, 15% dan 20%. Ditinjau dari nilai rata-rata MSE dengan 1000 pengulangan maka dapat disimpulkan pembobot *Bisquare Tukey* lebih baik digunakan pada persentase pencilan 5%, 10%, dan 15% sedangkan pembobot *Andrew* lebih baik digunakan pada persentase pencilan 20%. Hasil penelitian ini sejalan dengan penelitian Sinay dan Talakua (2015) dan penelitian Deria, dkk. (2019).

DAFTAR PUSTAKA

Chatterjee, S. dan Hadi, A.S. 2006. *Regression Analysis By Example*. 4 thEdition. John Wiley & Sons, New York.

Dawson, R. 2011. How Significant is a Boxplot Outlier?. *Journal of Statistics Education.* **19**(2).

Deria, A.D., Hoyyi, A., dan Mustafid. 2019. Regresi Robust Estimasi-M dengan Pembobot Andrew, Pembobot Ramsay dan Pembobot Welsch Menggunakan Software R. *Jurnal Gaussian*. **8**(3): 377–388.

Draper, N.R., dan Smith, H. 1998. *Applied Regression Analysis.* 3 th Edition. John Wiley & Sons, New York.

Febrianto, L.S., Dwidayati, N.K., dan Hendikawati, P. 2018. Perbandingan Metode Robust Least Median Of Square (LMS) dan Penduga S untuk Menangani Outlier pada Regresi Linier Berganda. *Journal of Mathematics.* **7**(1).

Freund, R.J., Wilson, W.J., dan Sa, P. 2006. *Regression Analysis.* 2 nd Edition. Elsevier Inc., San Diego.

Maronna, R.A., Martin, R.D., Yohai, V.J., dan Barrera, M.S. 2019. *Robust Statistics.* John Wiley and Sons Ltd, Haboken.

Montgomery, D.C., Peck, E.A., dan Vining, G.G. 2012. *Introduction to Liniar Regression Analysis.* 4 th Edition. John Wiley and Sons, New Jersey.

Montgomery, D.C. dan Runger, G.C. 2003. *Applied Statistics and Probability for Engineer*. 3 rd Edition. John Wiley and Sons, United States of America.

Myers, R.H., Montgomery, D.C., Vining, G.G., dan Robinson, T.J. 2010. *Generalized Liniar Models with Applications in Engineering and the Sciences.* 2 nd Edition. John Wiley and Sons, New Jersey.

Neter, J., Wasserman, W., dan Kutner, M. 1997. *Model Linear Terapan Buku I: Analisi Regresi Linier Sederhana*. ITB, Bandung.

Pitselis, G. 2013. A Review on Robust Estimators Applied to Regression Credibility. *Journal of Computational and Applied Mathematics.* 231-249.

Pradewi, E.D. dan Sudarno. 2012. Kajian Estimasi-M Menggunakan Fungsi Pembobot Huber dan Bisquare Tukey pada Data Ketahanan Pangan di Jawa Tengah. *Media Statistika*. **5**(1): 1–10.

Rousseeuw, P.J., dan Leroy, A.M. 1987. *Robust Regression and Outlier Detection.* John Wiley & Sons, New York.

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

Sinay, L.J. dan Talakua, M.W. 2015. Pemodelan Harga Saham Indeks LQ45 Menggunakan Regresi Linier Robust M-Estimator: Huber dan Bisquare Tukey. *BAREKENG: Jurnal Ilmu Matematika dan Terapan*. **9**(1): 51–61.