

**PERBANDINGAN PEMBOBOT *WELSC* DAN *TUKEY BISQUARE* PADA
REGRESI *ROBUST S-ESTIMATOR***

(Skripsi)

Oleh

FIFI NURHAFIFAH

NPM 1717031005



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2021**

ABSTRACT

COMPARISON OF WELSCH AND TUKEY BISQUARE ON ROBUST REGRESSION S-ESTIMATOR

By

FIFI NURHAFIFAH

Robust regression is a method developed to have good performance when the analyzed data deviates from the underlying assumptions, for example, there are outliers that can cause errors to be not normally distributed. One of the estimation methods in robust regression is the S-estimator, this method has weighting functions, including the Welsch and Tukey Bisquare weights. In this study, we compared the weights in the S-estimator method on data sizes: 30, 60, 100 and 200 which were given outlier contamination of: 5%, 10%, 15%, 20%, 25% and 30%. Based on the simulation results, it is found that the two weights produce similar MSE (Mean Square Error) and bias values. So it can be concluded that the two weights provide appropriate and equally good results in the S-estimator regression.

Key word: robust regression, s-estimator, welsch, tukey bisquare.

ABSTRAK

PERBANDINGAN PEMBOBOT WELSCH DAN TUKEY BISQUARE PADA REGRESI ROBUST S-ESTIMATOR

Oleh

FIFI NURHAFIFAH

Regresi *robust* merupakan sebuah metode yang dikembangkan untuk memiliki kinerja yang baik ketika data yang dianalisis menyimpang dari asumsi yang mendasari, misalnya terdapat pencilan yang dapat menyebabkan galat menjadi tidak berdistribusi normal. Salah satu metode estimasi pada regresi *robust* adalah *S-estimator*, metode ini memiliki fungsi pembobot antara lain pembobot *Welsch* dan *Tukey Bisquare*. Pada penelitian ini, kami membandingkan bobot-bobot pada metode *S-estimator* pada data berukuran: 30, 60, 100 dan 200 yang diberikan kontaminasi pencilan sebesar: 5%, 10%, 15%, 20%, 25% dan 30%. Berdasarkan hasil simulasi diperoleh bahwa kedua pembobot menghasilkan nilai MSE (*Mean Square Error*) dan bias yang serupa. Sehingga dapat disimpulkan bahwa kedua pembobot memberikan hasil yang sesuai dan sama baiknya pada regresi *S-estimator*.

Kata kunci: regresi *robust*, *s-estimator*, *welsch*, *tukey bisquare*.

**PERBANDINGAN PEMBOBOT *WELSC* DAN *TUKEY BISQUARE* PADA
REGRESI *ROBUST S-ESTIMATOR***

Oleh

FIFI NURHAFIFAH

Skripsi

**Sebagai Salah Satu Syarat untuk Mencapai Gelar
SARJANA MATEMATIKA**

Pada

**Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2021**

Judul Skripsi : **PERBANDINGAN PEMBOBOT *WELSH*
DAN *TUKEY BISQUARE* PADA REGRESI
*ROBUST S-ESTIMATOR***

Nama Mahasiswa : **Fifi Nurhafifah**

Nomor Pokok Mahasiswa : 1717031005

Jurusan : Matematika

Fakultas : Matematika dan Ilmu Pengetahuan Alam

MENYETUJUI

1. Komisi Pembimbing



Dr. Khoirin Nisa, M.Si.
NIP 19740726 200003 2 001



Amanto, S.Si., M.Si.
NIP 19730314 200012 1 002

2. Ketua Jurusan Matematika



Prof. Dra. Wamilliana, M.A., Ph.D.
NIP 19631108 198902 2 001

MENGESAHKAN

1. Tim Penguji

Ketua

: **Dr. Khoirin Nisa, M.Si.**


.....

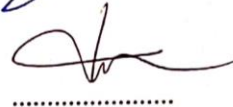
Sekretaris

: **Amanto, S.Si., M.Si.**


.....

Penguji

Bukan Pembimbing : **Drs. Nusyirwan, M.Si.**


.....

2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



Dr. Eng. Supto Dwi Yuwono, M.T.
NIP 19740705 200003 1 001

Tanggal Lulus Ujian Skripsi : **28 April 2021**

PERNYATAAN SKRIPSI MAHASISWA

Yang bertanda tangan di bawah ini:

Nama : Fifi Nurhafifah
Nomor Pokok Mahasiswa : 1717031005
Jurusan : Matematika
Judul Skripsi : Perbandingan Pembobot *Welsch* dan *Tukey*
Bisquare pada Regresi *Robust S-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, 28 April 2021
Yang menyatakan,



Fifi Nurhafifah

RIWAYAT HIDUP

Penulis lahir di Desa Rawa Selapan, pada 16 Desember tahun 1999. Merupakan anak kedua dari dua bersaudara, dari pasangan Bapak Fahruroji dan Ibu Siti Umi Murtofingah.

Riwayat pendidikan Sekolah Dasar di SD N 9 Muntok Kab. Bangka Barat lulusan tahun 2011. Sekolah Menengah Pertama di SMP N 4 Muntok Kab. Bangka Barat lulusan tahun 2014, dan Sekolah Menengah Atas di SMA N 1 Pemali Kab. Bangka Kelas Beasiswa PT. Timah, Tbk, lulusan tahun 2017.

Penulis, melanjutkan jenjang pendidikan di Universitas Lampung jalur SNMPTN serta Beasiswa Bidikmisi tahun 2017 dan terdaftar sebagai mahasiswa jurusan Matematika FMIPA. Pada tahun 2020 Penulis melakukan Kuliah Kerja Nyata (KKN) di Desa Tanjung Siom Kec. Limau Kab. Tanggamus, dan Kuliah Praktik (KP) di Badan Pusat Statistik (BPS) Kab. Bangka Barat. Riwayat organisasi yakni Himatika (2018), UKM F Natural (2018), Ikadubas Lampung (2019), Isbabel Lampung (2019) dan Alumni Bujang Dayang Bangka Barat (2020). Prestasi yang pernah diraih selama kuliah yaitu: menjadi Finalis Duta Bahasa Provinsi Lampung (2018), Juara 3 tingkat Nasional kategori Artikel Blog HII (2018), peserta 2nd *youth camp* BRG Republik Indonesia (2019), Juara 1 Dayang Bangka Barat (2020) dan peraih medali perak kategori poster PKM-PE PIMNAS ke-33 (2020).

KATA INSPIRASI

“Tak masalah meski harus jatuh dan gagal berulang kali, jadikan itu sebagai catatan agar tidak salah mengambil langkah ketika mencoba lagi.”

“Tak mengapa jika masa sulitmu tidak ada yang ingin mengetahuinya, simpan baik-baik untuk diceritakan ketika sukses. Cerita yang pahit akan menjadi motivasi bagi orang yang mengambil pelajaran.”

“Jika tidak bisa menjadi orang yang pintar, jadilah orang yang rajin. Jika tak mampu menjadi keduanya maka perbanyak doa, meski peluangnya kecil. Menjadi orang beruntung dapat mengalahkan orang pintar dan rajin.”

“Ketika kamu merasa beruntung, itu artinya Tuhan sedang mengabulkan salah satu doa Ibumu.”

- Fifi Nurhafifah -

PERSEMBAHAN

Puji dan syukur saya ucapkan kepada Allah SWT atas segala rahmat dan karunia-Nya. Tak lupa, selawat beserta salam senantiasa tercurahkan kepada baginda besar Nabi Muhammad SAW yang merupakan suri tauladan terbaik bagi seluruh umat.

Kupersembahkan sebuah karya sederhana ini kepada:

Kedua Orang Tua Tercinta

Terima kasih atas segala hal yang telah kalian berikan, baik dari segi materi, waktu luang, maupun doa yang tiada terhingga. Semoga kebaikan Bapak dan Ibu dibalas oleh Allah SWT dan keberkahan hidup selalu menyertai kalian.

Kakak Tersayang

Semoga karya kecil ini dapat membuat Mas bangga memiliki Adik seperti saya.

Srikandi Terbaik

Terima kasih atas seluruh dukungan yang kalian berikan, semoga hubungan baik kita selalu terjaga tidak hanya di dunia, tetapi juga di akhirat.

Almamater Kebanggaan Universitas Lampung.

SANWACANA

Puji syukur penulis ucapkan kepada Allah SWT atas segala ridha, rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi ini dengan tepat waktu. Skripsi dengan judul “Perbandingan Pembobot *Welsch* dan *Tukey Bisquare* pada Regresi *Robust S-Estimator*” disusun untuk memenuhi salah satu syarat memperoleh gelar sarjana Matematika di Universitas Lampung.

Pada kesempatan kali ini penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Ibu Dr. Khoirin Nisa, M.Si. selaku pembimbing I yang telah memberikan masukan, saran serta bimbingan kepada penulis selama proses pembuatan skripsi ini hingga selesai.
2. Bapak Amanto, S.Si., M.Si. selaku pembimbing II serta pembimbing akademik yang telah memberikan bimbingan, arahan selama perkuliahan.
3. Bapak Drs. Nusyirwan, M.Si. selaku dosen pembahas yang telah memberikan evaluasi, arahan serta saran hingga terselesaikannya skripsi ini.
4. Ibu Prof. Dra. Wamiliana, M.A., Ph.D. selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
5. Bapak Dr. Eng. Suropto Dwi Yuwono, M.T. selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
6. Seluruh civitas akademik, dosen dan staf Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
7. Ibu, Bapak dan Anggota Keluarga yang selalu memberikan dukungan dan doa terbaik agar penulis diberikan kelancaran serta kemudahan dalam menyelesaikan skripsi ini.

8. Teman-teman Srikandi; Rosita Sari, Dwi Oktavia, dan Siti Munawaroh yang selalu memberikan dukungan, pelajaran hidup, waktu luang dan nasihat selama perkuliahan.
9. Mba Chandra Pratiwi, Mba Venna Tristya Nafulani, dan Lilik Anggraini yang telah memberikan dukungan moral dalam mengerjakan skripsi.
10. Seluruh teman-teman mahasiswa jurusan matematika angkatan 2017.

Penulis menyadari bahwa dalam penyusunan skripsi ini masih banyak kekurangan. Oleh karena itu, kritik dan saran yang membangun sangat dibutuhkan, guna penyempurnakan skripsi ini.

Bandar Lampung, 28 April 2021
Penulis,

Fifi Nurhafifah

DAFTAR ISI

Halaman

DAFTAR TABEL	xv
--------------------	----

DAFTAR GAMBAR.....	xviii
--------------------	-------

I. PENDAHULUAN	1
1.1 Latar Belakang dan Masalah.....	1
1.2 Tujuan Penelitian	3
1.3 Manfaat Penelitian	4
II. TINJAUAN PUSTAKA.....	5
2.1 Analisis Regresi Linier	5
2.1.1 Uji Asumsi Klasik Regresi Linear Berganda	6
2.1.2 Estimasi Parameter Model Regresi Linear Berganda.....	8
2.1.3 Uji Hipotesis dalam Regresi Linear Berganda	8
2.1.4 Ukuran Kecocokan Model (<i>Goodnes of Fit</i>)	10
2.2 Pencilan.....	10
2.2.1 Uji DFFITS (<i>Difference in Fit Standardized</i>).....	11
2.3 Regresi <i>Robust</i>	12
2.4 Pembobotan <i>Welsch</i>	13
2.5 Pembobotan <i>Tukey Bisquare</i>	14
2.6 Metode <i>S-Estimator</i>	15
III. METODOLOGI PENELITIAN.....	18
3.1 Waktu dan Tempat Penelitian	18
3.2 Data Penelitian	18
3.3 Metode Penelitian	18
IV. HASIL DAN PEMBAHASAN.....	21
4.1 Hasil dan Pembahasan	21
4.2 Hasil Simulasi Data Berukuran 30	26
4.3 Hasil Simulasi Data Berukuran 60	29
4.4 Hasil Simulasi Data Berukuran 100	31
4.5 Hasil Simulasi Data Berukuran 200	33

4.6 Perbandingan Nilai MSE dan Bias untuk Seluruh Jumlah Data	36
4.7 Pengujian Hipotesis Terhadap Rata-rata Penduga.....	38
V. KESIMPULAN	43
DAFTAR PUSTAKA	44
LAMPIRAN.....	47
Syntax R.....	47
Gambar 14-34.....	50
Tabel 17-55	57

DAFTAR TABEL

Tabel	Halaman
1. Nilai <i>Tuning Constant</i> untuk Setiap <i>Breakdown Point</i>	14
2. Contoh Data Simulasi $n=30$ dengan Pencilan Sebesar 10%	22
3. Nilai Sisaan Galat dari Analisis Regresi	23
4. Nilai Skala Sisaan pada Pengamatan Ke- i	24
5. Nilai Pembobot <i>Welsch</i> dan <i>Tukey Bisquare</i>	25
6. Nilai MSE dari Pembobot <i>Welsch</i> dan <i>Tukey Bisquare</i> Metode <i>S-Estimator</i> dengan $n=30$	27
7. Nilai Bias dari Pembobot <i>Welsch</i> dan <i>Tukey Bisquare</i> Metode <i>S-Estimator</i> dengan $n=30$	28
8. Nilai MSE dari Pembobot <i>Welsch</i> dan <i>Tukey Bisquare</i> Metode <i>S-Estimator</i> dengan $n=60$	29
9. Nilai Bias dari Pembobot <i>Welsch</i> dan <i>Tukey Bisquare</i> Metode <i>S-Estimator</i> dengan $n=60$	30
10. Nilai MSE dari Pembobot <i>Welsch</i> dan <i>Tukey Bisquare</i> Metode <i>S-Estimator</i> dengan $n=100$	31
11. Nilai Bias dari Pembobot <i>Welsch</i> dan <i>Tukey Bisquare</i> Metode <i>S-Estimator</i> dengan $n=100$	32
12. Nilai MSE dari Pembobot <i>Welsch</i> dan <i>Tukey Bisquare</i> Metode <i>S-Estimator</i> dengan $n=200$	34
13. Nilai Bias dari Pembobot <i>Welsch</i> dan <i>Tukey Bisquare</i> Metode <i>S-Estimator</i> dengan $n=200$	35
14. Uji Hipotesis Rata-rata Penduga untuk $n=30$ dan Pencilan=10%.....	39
15. Uji Hipotesis Rata-rata Penduga untuk $n=30$ dan Pencilan=20%.....	40
16. Uji Hipotesis Rata-rata Penduga untuk $n=30$ dan Pencilan=30%.....	41

17. Uji Hipotesis Rata-rata Penduga untuk $n=60$ dan Pencilan=5%.....	57
18. Uji Hipotesis Rata-rata Penduga untuk $n=60$ dan Pencilan=10%.....	57
19. Uji Hipotesis Rata-rata Penduga untuk $n=60$ dan Pencilan=15%.....	58
20. Uji Hipotesis Rata-rata Penduga untuk $n=60$ dan Pencilan=20%.....	58
21. Uji Hipotesis Rata-rata Penduga untuk $n=60$ dan Pencilan=25%.....	58
22. Uji Hipotesis Rata-rata Penduga untuk $n=60$ dan Pencilan=30%.....	59
23. Uji Hipotesis Rata-rata Penduga untuk $n=100$ dan Pencilan=5%.....	59
24. Uji Hipotesis Rata-rata Penduga untuk $n=100$ dan Pencilan=10%.....	59
25. Uji Hipotesis Rata-rata Penduga untuk $n=100$ dan Pencilan=15%.....	60
26. Uji Hipotesis Rata-rata Penduga untuk $n=100$ dan Pencilan=20%.....	60
27. Uji Hipotesis Rata-rata Penduga untuk $n=100$ dan Pencilan=25%.....	60
28. Uji Hipotesis Rata-rata Penduga untuk $n=100$ dan Pencilan=30%.....	61
29. Uji Hipotesis Rata-rata Penduga untuk $n=200$ dan Pencilan=5%.....	61
30. Uji Hipotesis Rata-rata Penduga untuk $n=200$ dan Pencilan=10%.....	61
31. Uji Hipotesis Rata-rata Penduga untuk $n=200$ dan Pencilan=15%.....	62
32. Uji Hipotesis Rata-rata Penduga untuk $n=200$ dan Pencilan=20%.....	62
33. Uji Hipotesis Rata-rata Penduga untuk $n=200$ dan Pencilan=25%.....	62
34. Uji Hipotesis Rata-rata Penduga untuk $n=200$ dan Pencilan=30%.....	63
35. Nilai $\hat{\beta}$ untuk $n=30$ dan Pencilan=10%.....	64
36. Nilai $\hat{\beta}$ untuk $n=30$ dan Pencilan=20%.....	66
37. Nilai $\hat{\beta}$ untuk $n=30$ dan Pencilan=30%.....	68
38. Nilai $\hat{\beta}$ untuk $n=60$ dan Pencilan=5%.....	70
39. Nilai $\hat{\beta}$ untuk $n=60$ dan Pencilan=10%.....	73
40. Nilai $\hat{\beta}$ untuk $n=60$ dan Pencilan=15%.....	75
41. Nilai $\hat{\beta}$ untuk $n=60$ dan Pencilan=20%.....	77
42. Nilai $\hat{\beta}$ untuk $n=60$ dan Pencilan=25%.....	79
43. Nilai $\hat{\beta}$ untuk $n=60$ dan Pencilan=30%.....	81
44. Nilai $\hat{\beta}$ untuk $n=100$ dan Pencilan=5%.....	83
45. Nilai $\hat{\beta}$ untuk $n=100$ dan Pencilan=10%.....	85
46. Nilai $\hat{\beta}$ untuk $n=100$ dan Pencilan=15%.....	88
47. Nilai $\hat{\beta}$ untuk $n=100$ dan Pencilan=20%.....	90

48. Nilai $\hat{\beta}$ untuk $n=100$ dan Pencilan=25%.....	92
49. Nilai $\hat{\beta}$ untuk $n=100$ dan Pencilan=30%.....	94
50. Nilai $\hat{\beta}$ untuk $n=200$ dan Pencilan=5%.....	96
51. Nilai $\hat{\beta}$ untuk $n=200$ dan Pencilan=10%.....	98
52. Nilai $\hat{\beta}$ untuk $n=200$ dan Pencilan=15%.....	101
53. Nilai $\hat{\beta}$ untuk $n=200$ dan Pencilan=20%.....	103
54. Nilai $\hat{\beta}$ untuk $n=200$ dan Pencilan=25%.....	105
55. Nilai $\hat{\beta}$ untuk $n=200$ dan Pencilan=30%.....	107

DAFTAR GAMBAR

Gambar	Halaman
1. Diagram Alir Metodologi Penelitian.....	20
2. Grafik MSE untuk Data $n=30$	27
3. Grafik Bias untuk Data $n=30$	28
4. Grafik MSE untuk Data $n=60$	29
5. Grafik Bias untuk Data $n=60$	30
6. Grafik MSE untuk Data $n=100$	32
7. Grafik Bias untuk Data $n=100$	33
8. Grafik MSE untuk Data $n=200$	34
9. Grafik Bias untuk Data $n=200$	35
10. MSE Pembobot <i>Welsch</i>	36
11. MSE Pembobot <i>Tukey Bisquare</i>	37
12. Bias Pembobot <i>Welsch</i>	37
13. Bias Pembobot <i>Tukey Bisquare</i>	38
14. Plot Sebaran Data $n=30$ dan Pencilan=10%	50
15. Plot Sebaran Data $n=30$ dan Pencilan=20%	50
16. Plot Sebaran Data $n=30$ dan Pencilan=30%	50
17. Plot Sebaran Data $n=60$ dan Pencilan=5%	51
18. Plot Sebaran Data $n=60$ dan Pencilan=10%	51
19. Plot Sebaran Data $n=60$ dan Pencilan=15%	52
20. Plot Sebaran Data $n=60$ dan Pencilan=20%	52
21. Plot Sebaran Data $n=60$ dan Pencilan=25%	52
22. Plot Sebaran Data $n=60$ dan Pencilan=30%	52
23. Plot Sebaran Data $n=100$ dan Pencilan=5%	53
24. Plot Sebaran Data $n=100$ dan Pencilan=10%	53

25. Plot Sebaran Data $n=100$ dan Pencilan=15%	53
26. Plot Sebaran Data $n=100$ dan Pencilan=20%	54
27. Plot Sebaran Data $n=100$ dan Pencilan=25%	54
28. Plot Sebaran Data $n=100$ dan Pencilan=30%	54
29. Plot Sebaran Data $n=200$ dan Pencilan=5%	55
30. Plot Sebaran Data $n=200$ dan Pencilan=10%	55
31. Plot Sebaran Data $n=200$ dan Pencilan=15%	55
32. Plot Sebaran Data $n=200$ dan Pencilan=20%	56
33. Plot Sebaran Data $n=200$ dan Pencilan=25%	56
34. Plot Sebaran Data $n=200$ dan Pencilan=30%	56

I. PENDAHULUAN

1.1 Latar Belakang dan Masalah

Analisis regresi adalah sebuah analisis statistika yang bertujuan untuk mengestimasi model hubungan antara variabel terikat (y) dengan satu atau lebih variabel bebas (x). Analisis regresi dengan satu variabel bebas biasa dikenal dengan analisis regresi linear sederhana. Sedangkan analisis regresi dengan dua atau lebih variabel bebas dikenal sebagai analisis regresi linear berganda. Dengan kata lain, analisis regresi akan menghasilkan model yang mampu mengestimasi rata-rata variabel terikat berdasarkan nilai variabel bebas yang diketahui.

Salah satu metode yang sering digunakan untuk mengestimasi parameter model regresi adalah Metode Kuadrat Terkecil (MKT). Cara kerja MKT adalah dengan meminimumkan jumlah kuadrat galat sehingga menghasilkan estimasi linier yang tak bias serta ragam yang minimum (efisien). Metode Kuadrat Terkecil dikenal sebagai metode penduga terbaik diantara semua penduga tak bias linier^[1]. Untuk memperoleh sebuah model persamaan regresi yang baik, diperlukan beberapa asumsi klasik yang harus dipenuhi yakni galat berdistribusi normal, ragam galat homogen, tidak terjadi autokorelasi serta tidak mengandung multikolinearitas. Setelah semua asumsi dipenuhi, maka dapat dikatakan bahwa suatu model dalam analisis regresi tersebut bersifat BLUE (*Best Linear Unbiased Estimator*) atau merupakan penduga terbaik^[2].

Metode Kuadrat Terkecil (MKT) sangat peka terhadap adanya penyimpangan asumsi. Jika terdapat asumsi yang dilanggar, maka MKT menjadi tidak efisien. Salah satu asumsi yang sering dilanggar yakni normalitas. Asumsi galat berdistribusi normal menjadi tidak terpenuhi ketika terdapat pencilan pada data. Pencilan diartikan sebagai sebuah pengamatan yang tidak mengikuti sebagian besar pola data, sehingga keberadaannya akan melanggar asumsi normalitas. Mengingat, letaknya yang jauh dari pusat data, keberadaan pencilan dapat menyebabkan distribusi data menjadi sangat menjulur (*Heavy Tailed Distribution*). Membuang pencilan bukan merupakan keputusan yang tepat, karena tidak menutup kemungkinan pencilan mengandung informasi yang tidak mampu diberikan oleh data yang lainnya. Sehingga ketika sebuah data mengandung pencilan, diperlukan satu metode yang resisten terhadap adanya pencilan seperti metode regresi *robust*. Pada dasarnya analisis regresi *robust* tidak menjadikan galat berdistribusi normal akan tetapi model yang dihasilkan oleh metode ini memiliki tingkat keakuratan yang lebih tinggi dibandingkan MKT^[3].

Regresi *robust* merupakan metode yang dirancang agar memiliki kinerja yang baik ketika bentuk model yang dihasilkan menyimpang dari asumsi yang mendasarinya misalnya terdapat sebuah pencilan^[4]. Regresi *robust* terdiri dari lima metode estimasi yakni *Maximum Likelihood Type (M)-estimator*, *Least Median Square (LMS)-estimator*, *Least Trimmed Square (LTS)-estimator*, *Scale (S)-estimator* dan *Method of Moment (MM)-estimator*. Dengan *breakdown point* yang sama, metode *S-estimator* memiliki efisiensi statistik yang lebih tinggi jika dibandingkan dengan metode *LTS-estimator*^[5].

S-estimator memiliki beberapa fungsi pembobot antara lain pembobot *Welsch* dan *Tukey Bisquare*. Penggunaan fungsi pembobot *Welsch* dan *Tukey Bisquare* pada *S-estimator* adalah untuk menghasilkan nilai skala pembobot dengan cara melakukan iterasi hingga *estimator* yang diperoleh konvergen. Berbagai penelitian terkait dengan analisis regresi *robust S-estimator* untuk mengatasi pencilan, diantaranya berjudul “Analisis Regresi *Robust* Estimasi-S Menggunakan

Pembobot *Welsch* dan *Tukey Bisquare*”. Pada penelitian ini, metode estimasi-S diterapkan pada data IPM (Indeks Pembangunan Manusia) menurut provinsi tahun 2015 yang diolah dengan menggunakan perangkat lunak R. Berdasarkan nilai RSE (*Residual Standar Error*) dan *R-square* yang dihasilkan oleh kedua pembobot, diperoleh kesimpulan bahwa pembobot *Welsch* menghasilkan estimasi yang lebih baik jika diterapkan pada estimasi-S karena nilai RSE yang kecil serta nilai *R-square* yang besar^[6]. Penelitian selanjutnya berjudul “Metode Estimasi-S pada Analisis Regresi *Robust* dengan Pembobotan *Tukey Bisquare*”. Pada penelitian ini akan dilihat pengaruh IPM, angka partisipasi sekolah usia (APS) 16-18 tahun dan konsumsi terhadap kemiskinan di Indonesia pada tahun 2016. Setelah teridentifikasi adanya pencilan pada data maka metode regresi *robust* estimasi-S dengan pembobot *Tukey Bisquare* dipilih untuk mengatasinya, setelah dianalisis lebih lanjut ditarik suatu kesimpulan bahwa regresi *robust* estimasi-S dengan pembobot *Tukey Bisquare* menghasilkan model yang lebih baik jika dibandingkan dengan MKT (Metode Kuadrat Terkecil)^[7].

Berdasarkan uraian di atas, penelitian ini akan membahas mengenai estimasi regresi *robust S-estimator* dengan pembobotan *Welsch* dan *Tukey Bisquare*. Selanjutnya dengan berbagai ukuran sampel, akan dilakukan perbandingan keefektifan ditinjau dari nilai MSE (*Mean Square Error*) dan bias agar diperoleh pembobot terbaik.

1.2 Tujuan Penelitian

Adapun tujuan penelitian ini adalah:

1. Mengkaji regresi *robust S-estimator* dengan pembobot *Welsch* dan *Tukey Bisquare*.
2. Membandingkan analisis regresi *robust S-estimator* dengan pembobot *Welsch* dan *Tukey Bisquare* ditinjau dari nilai MSE dan bias.

1.3 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah:

1. Untuk menambah wawasan dan pengetahuan bagi penulis mengenai regresi *robust S-estimator* dengan pembobot *Welsch* dan *Tukey Bisquare*. Serta diharapkan dapat memberikan alternatif bagi pembaca, dalam menentukan pembobot terbaik pada regresi *robust S-estimator*.

II. TINJAUAN PUSTAKA

2.1 Analisis Regresi Linier

Analisis regresi adalah sebuah metode statistika yang digunakan untuk mengestimasi model hubungan antara variabel terikat (y) dengan satu atau lebih variabel bebas (x). Analisis regresi dengan satu variabel bebas biasa dikenal dengan analisis regresi linier sederhana, sedangkan analisis regresi dengan dua atau lebih variabel bebas dikenal dengan analisis regresi linier berganda. Salah satu dari tujuan penggunaan analisis regresi dalam statistika adalah untuk mengestimasi rata-rata populasi atau nilai rata-rata variabel terikat berdasarkan nilai variabel bebas yang diketahui. Hasil dari analisis regresi berupa koefisien regresi, untuk masing-masing variabel bebas^[6].

Model regresi adalah model yang memberikan gambaran mengenai hubungan antara variabel bebas dengan variabel terikat. Model regresi berganda dinyatakan dalam bentuk persamaan^[8]:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \varepsilon_i. \quad (2.1)$$

dengan:

Y_i : variabel terikat
 $\beta_0, \beta_1, \beta_2, \dots, \beta_k$: koefisien regresi
 $X_{i1}, X_{i2}, \dots, X_{ik}$: variabel bebas
 ε_i : galat.

Jika ditulis dalam bentuk matriks, persamaan (2.1) dapat ditulis sebagai berikut:

$$\mathbf{Y}_{(n \times 1)} = \mathbf{X}_{(n \times (k+1))} \boldsymbol{\beta}_{((k+1) \times 1)} + \boldsymbol{\varepsilon}_{(n \times 1)}. \quad (2.2)$$

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & X_{11} & X_{12} & \dots & X_{1k} \\ 1 & X_{21} & X_{22} & \dots & X_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \dots & X_{nk} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}.$$

2.1.1 Uji Asumsi Klasik Analisis Regresi Linier Berganda

Uji asumsi klasik dilakukan untuk melihat apakah model yang dihasilkan memenuhi seluruh asumsi analisis regresi linier berganda. Jika semua asumsi terpenuhi, maka model yang dihasilkan akan bersifat BLUE (*Best Linier Unbiased Estimator*)^[2].

Uji asumsi klasik pada analisis regresi linier berganda terdiri dari^[9]:

1. Uji Normalitas Galat.

Uji normalitas bertujuan untuk menguji apakah dalam sebuah model regresi, variabel galat berdistribusi normal. Uji normalitas data dilakukan dengan uji *Kolmogorov-Smirnov*.

i. Uji Hipotesis:

H_0 : galat berdistribusi normal

H_1 : galat tidak berdistribusi normal.

ii. Taraf Signifikansi α .

iii. Menghitung Statistik Uji

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

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 $D_{hitung} > D_{(1-\alpha, n)}$ dengan $D_{(1-\alpha, n)}$ diperoleh dari tabel *Kolmogorov-Smirnov*.

2. Uji Homogenitas Ragam Galat.

Uji homogenitas bertujuan untuk melihat apakah dalam sebuah regresi memiliki ragam galat yang homogen. Uji homogenitas ragam galat dilakukan dengan uji *Breusch Pagan*.

i. Uji Hipotesis

H_0 : model regresi bersifat homoskedastisitas

H_1 : model regresi tidak bersifat homoskedastisitas.

ii. Taraf Signifikansi α .

iii. Keputusan Uji:

Tolak H_0 ketika $P_{value} < \alpha$.

3. Uji Autokorelasi Galat.

Uji autokorelasi bertujuan untuk menguji apakah dalam suatu model regresi linear terdapat korelasi antar galat pada periode t dengan kesalahan galat pada periode t-1 (sebelumnya).

i. Uji Hipotesis

$H_0: \rho = 0$, artinya tidak terdapat autokorelasi

$H_1: \rho \neq 0$, artinya terdapat autokorelasi.

ii. Taraf Signifikansi α .

iii. Keputusan Uji

jika dapat di tunjukkan $d_u < d < 4 - d_u$, dimana

d = d-hitung *Durbin Watson*

d_u = nilai kritis batas atas dari tabel *Durbin Watson*.

Maka dapat disimpulkan bahwa dalam model tidak terjadi autokorelasi.

4. Uji Multikolinieritas.

Uji multikolinearitas digunakan untuk mengetahui ada atau tidaknya korelasi antar variabel bebas dalam model regresi. Uji multikolinieritas dilakukan dengan cara menentukan nilai VIF (*Variance Inflation Factor*), jika $VIF < 10$ maka asumsi non multikolinieritas terpenuhi.

2.1.2 Estimasi Parameter Model Regresi Linier Berganda

Metode yang sering digunakan untuk mengestimasi parameter model regresi dengan cara meminimumkan jumlah kuadrat galat adalah metode OLS (*Ordinary Least Square*) atau biasa dikenal dengan Metode Kuadrat Terkecil (MKT). Misalkan terdapat p parameter dan n pengamatan maka model yang akan diperoleh sebagai berikut^[10]:

$$\begin{aligned} Y_1 &= \hat{\beta}_0 + \hat{\beta}_1 X_{11} + \cdots + \hat{\beta}_{p-1} X_{1,p-1} + \varepsilon_1 \\ Y_2 &= \hat{\beta}_0 + \hat{\beta}_1 X_{21} + \cdots + \hat{\beta}_{p-1} X_{2,p-1} + \varepsilon_2 \\ &\vdots \\ Y_n &= \hat{\beta}_0 + \hat{\beta}_1 X_{n1} + \cdots + \hat{\beta}_{p-1} X_{n,p-1} + \varepsilon_n. \end{aligned}$$

Persamaan di atas dapat ditulis secara ringkas dalam notasi matriks sebagai:

$$Y = X\hat{\beta} + \varepsilon. \quad (2.3)$$

Estimasi parameter MKT untuk kasus n pengamatan dapat diperoleh dengan cara meminimumkan jumlah kuadrat galat sebagai berikut:

$$\sum_{i=1}^n \varepsilon_i^2 = \sum [(Y_i - (\beta_0 + \hat{\beta}_1 X_{i1} + \cdots + \hat{\beta}_{p-1} X_{i,p-1}))]^2.$$

Notasi matriks untuk meminimumkan $\varepsilon^T \varepsilon$ dari persamaan (2.3) diperoleh:

$$\varepsilon = Y - X\hat{\beta}$$

oleh karena itu,

$$\begin{aligned} \varepsilon^T \varepsilon &= (Y - X\hat{\beta})^T (Y - X\hat{\beta}) \\ &= Y^T Y - 2\hat{\beta}^T X^T Y + \hat{\beta}^T X^T X \hat{\beta}. \end{aligned} \quad (2.4)$$

Untuk mengestimasi parameter $\beta_1, \beta_2, \dots, \beta_{p-1}$ maka $\sum_{i=1}^n \varepsilon_i^2$ harus sekecil mungkin. Sehingga diperlukan penurunan persamaan (2.4) terhadap $\hat{\beta}$ dan membuatnya sama dengan nol. Agar diperoleh persamaan sebagai berikut:

$$\hat{\beta} = (X^T X)^{-1} (X^T Y). \quad (2.5)$$

2.1.3 Uji Hipotesis dalam Regresi Linier Berganda

Hipotesis merupakan pernyataan mengenai sifat populasi, sedangkan uji hipotesis diartikan sebagai sebuah prosedur untuk membuktikan kebenaran sifat populasi

berdasarkan data sampel. Dalam statistika hipotesis yang ingin diuji kebenarannya akan dibandingkan dengan negasinya. Hipotesis yang berperan sebagai negasi biasanya dilambangkan dengan hipotesis nol H_0 sedangkan H_1 dilambangkan sebagai hipotesis alternatif. Dalam menguji kebenaran suatu hipotesis, statistika telah mengembangkan uji-t. Uji-t merupakan suatu prosedur yang bertujuan agar suatu sampel dapat digunakan untuk verifikasi kebenaran atau kesalahan H_0 . Keputusan untuk menerima atau menolak H_0 ditentukan berdasarkan uji statistik yang diperoleh dari data^[11].

1. Uji Signifikansi Kecocokan Model Regresi

Uji signifikansi kecocokan model regresi bertujuan untuk melihat ada atau tidaknya hubungan linier antara variabel terikat (y) dan variabel bebas (x) secara keseluruhan^[12].

i. Uji Hipotesis

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

$$H_1: \text{terdapat } \beta_j \neq 0, \text{ dengan } j=1,2,\dots,k.$$

ii. Taraf Signifikansi α .

iii. Statistik Uji

$$F_{hitung} = \frac{JKR/k}{JKG/(n-k-1)} = \frac{MSR}{MSE}.$$

iv. Keputusan Uji

$$\text{Tolak } H_0 \text{ ketika } F_{hitung} > F_{(\alpha,k,n-k-1)} \text{ atau } P_{value} < \alpha.$$

2. Uji Koefisien Regresi Secara Individual

Pengujian koefisien regresi secara individual bertujuan untuk menguji ada atau tidaknya pengaruh yang signifikan antara masing-masing variabel bebas terhadap model regresi linier.

i. Uji Hipotesis

$$H_0: \beta_j = 0$$

$$H_1: \beta_j \neq 0, \text{ dengan } j=1,2,\dots,k.$$

ii. Taraf Signifikansi α .

iii. Statistik Uji

$$t_{hitung} = \frac{\hat{\beta}_j}{Se(\hat{\beta}_j)} ; \text{ dengan } Se(\hat{\beta}_j) = \sqrt{\hat{\sigma}^2 c_{jj}}.$$

iv. Keputusan Uji

Tolak H_0 ketika $|t_{hitung}| > t_{(\frac{\alpha}{2}, n-k-1)}$ atau $P_{value} < \alpha$.

2.1.4 Ukuran Kecocokan Model (*Goodnes of Fit*)

1. Koefisien Determinasi yang Disesuaikan (*Adjusted R^2*)

Nilai koefisien determinasi R^2 menentukan besarnya ragam dari variabel terikat (y) yang dapat dijelaskan oleh variabel bebas (x). Nilai R^2 bernilai antara 0 hingga 1, jika nilai R^2 mendekati 1 menunjukkan tingkat ketepatan model yang semakin baik dalam menerangkan ragam data^[12].

$$R^2_{Adj,k} = 1 - \frac{\frac{JKG}{(n-k-1)}}{\frac{JKT}{(n-1)}}. \quad (2.6)$$

2. RSE (*Residual Standard Error*)

Residual Standard Error (RSE) adalah estimasi simpangan baku (*standard deviation*) dari galat.

$$RSE = \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n-k-1}} = \sqrt{MSE}. \quad (2.7)$$

2.2 Pencilan

Pencilan merupakan sebuah pengamatan yang tidak mengikuti sebagian besar pola dan terletak jauh dari pusat data. Keberadaan pencilan dalam data dapat mengganggu proses analisis, sehingga mengakibatkan ragam menjadi besar dan interval kepercayaan memiliki rentang yang lebar^[13]. Adanya pencilan dalam suatu pengamatan, akan memberikan dampak negatif pada estimasi regresi dan parameter skala yang mempengaruhi pengambilan kesimpulan. Secara khusus,

deteksi pencilan dapat dipengaruhi oleh efek *masking* dan *swamping* yang masing-masing mengacu pada pencilan yang tidak berdiferensiasi dan pencilan yang diidentifikasi secara salah^[14].

Dalam analisis regresi terdapat tiga jenis pencilan yang dapat mempengaruhi pendugaan pada Metode Kuadrat Terkecil (MKT). Ketiga jenis pencilan tersebut antara lain adalah *vertical outlier*, *bad leverage point* dan *good leverage point*. *Vertical outlier* adalah pencilan yang berada di luar sumbu (y) akan tetapi tidak berada di luar variabel bebas (x). *Vertical outlier* mempengaruhi metode MKT khususnya pada intersep yang akan diduga. *Bad leverage point* adalah pengamatan yang berada di luar ruang variabel bebas dan terletak jauh dari garis regresi sebenarnya, kehadiran pencilan ini mempengaruhi estimasi intersep dan kemiringan. Sedangkan *good leverage point* adalah pencilan yang berada di luar variabel bebas akan tetapi dekat dengan garis regresi^[15].

2.2.1 Uji DFFITS (*Difference in Fit Standardized*)

Salah satu metode yang dapat digunakan untuk mendeteksi pencilan yang berpengaruh terhadap variabel terikat y adalah DFFITS (*Difference in Fit Standardized*)^[3]. Adapun rumus DFFITS^[12] didefinisikan:

$$(DFFITS_i) = t_i \left(\frac{h_{ii}}{1 - h_{ii}} \right)^{\frac{1}{2}} \quad (2.8)$$

dimana t_i adalah *studentized deleted* residual, dan h_{ii} adalah elemen diagonal dari matriks $\mathbf{H}$.

Dengan:

$$t_i = \frac{\varepsilon_i}{\sqrt{S_{(i)}^2(1-h_{ii})}}, i = 1, 2, \dots, n; \quad S_{(i)}^2 = \frac{(n-p)MS_E - \left(\frac{\varepsilon_i^2}{1-h_{ii}}\right)}{n-p-1};$$

$$MS_E = \frac{\mathbf{Y}^T \mathbf{Y} - \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{Y}}{n-k-1} \quad ; \quad \mathbf{H} = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T.$$

i. Uji Hipotesis

H_0 : Pencilan ke-i tidak berpengaruh

H_1 : Pencilan ke-i berpengaruh

ii. Statistik Uji

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

iii. Keputusan uji

Tolak H_0 ketika nilai $|DFFITS| > 1$ untuk gugus data kecil ($n < 30$),

sedangkan untuk data berukuran besar menggunakan $|DFFITS| > 2 \sqrt{\frac{p}{n}}$, k

adalah banyaknya variabel bebas, dan $p = k + 1$ ^[7].

2.3 Regresi *Robust*

Regresi *robust* merupakan metode yang dirancang agar memiliki kinerja yang baik ketika bentuk model yang dihasilkan menyimpang dari asumsi yang mendasarinya misalnya terdapat sebuah pencilan. Regresi *robust* harus mampu menggambarkan struktur yang sesuai dengan sebagian besar data dan mengidentifikasi titik data yang menyimpang dari asumsi^[4]. Untuk menentukan kinerja metode *robust* dalam arti teoritis digunakan sifat efisiensi, *breakdown point*, dan *high leverage point*^[16].

Metode *robust* merupakan metode analisis yang memiliki sifat sebagai berikut:

1. Sama baiknya dengan MKT ketika semua asumsi terpenuhi dan tidak terdapat titik data yang berpengaruh.
2. Dapat menghasilkan model regresi yang lebih baik daripada MKT ketika asumsi tidak dipenuhi dan terdapat titik data yang berpengaruh.
3. Perhitungannya cukup sederhana dan mudah dimengerti, tetapi dilakukan secara iterasi hingga diperoleh estimasi terbaik yang memiliki *standard error* parameter yang paling kecil^[2].

Regresi *robust* terdiri dari lima metode estimasi yakni *M-estimator*, *LMS-estimator*, *LTS-estimator*, *S-estimator* dan *MM-estimator*. Dengan *breakdown point* yang sama, metode *S-estimator* memiliki efisiensi statistik yang lebih tinggi jika dibandingkan dengan metode *LTS-estimator*^[5]. *Breakdown point* adalah jumlah maksimum data terkontaminasi pencilan yang dapat ditoleransi oleh suatu metode.

2.4 Pembobotan Welsch

Metode regresi *robust S-estimator* memiliki beberapa fungsi pembobot antara lain adalah *Welsch* dan *Tukey Bisquare*. Penggunaan fungsi pembobot *Welsch* dan *Tukey Bisquare* pada *S-estimator* adalah untuk menghasilkan nilai skala pembobot dengan cara melakukan iterasi hingga *estimator* yang diperoleh konvergen^[6].

Fungsi *Welsch* didefinisikan dalam persamaan berikut^[17]:

$$\rho(u_i) = \frac{c^2}{2} \left[1 - \exp \left(- \left(\frac{u_i}{c} \right)^2 \right) \right].$$

Kemudian untuk fungsi pengaruh Ψ yang merupakan turunan pertama dari $\rho(u_i)$ adalah sebagai berikut:

$$\Psi(u_i) = u_i \left[\exp \left(- \left(\frac{u_i}{c} \right)^2 \right) \right]$$

dan untuk fungsi pembobot *Welsch*:

$$w(u_i) = \exp \left(- \left(\frac{u_i}{c} \right)^2 \right).$$

Dengan nilai $c = 2,985$, dimana c adalah nilai *tunning constant* yang telah ditetapkan agar memperoleh tingkat *robust* yang paling efisien dari suatu pembobot. Tingkat *robust* (resisten) dari suatu pembobot dapat dilihat dari skala *robust* yang dihasilkan. Semakin kecil nilai skala yang dihasilkan dari suatu pembobot, maka akan semakin resisten terhadap adanya pencilan^[6].

2.5 Pembobotan *Tukey Bisquare*

Fungsi objektif untuk pembobotan *Tukey Bisquare* adalah sebagai berikut^[18]:

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

fungsi *Tukey Bisquare* merupakan turunan pertama dari $\rho(u_i)$, sehingga diperoleh persamaan:

$$\psi(u_i) = \begin{cases} u_i \left(1 - \left(\frac{u_i}{c} \right)^2 \right)^2, & |u_i| \leq c \\ 0, & |u_i| > c \end{cases}$$

fungsi pembobot *Tukey Bisquare* adalah:

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

dimana nilai u_i adalah skala sisaan pada pengamatan ke- i dan nilai $c = 1,547$ adalah nilai *tuning constan* yang telah ditetapkan untuk menentukan tingkat keresistenan (*robust*). Tabel 1 adalah nilai *tunning constant* untuk setiap *breakdown point*^[19].

Tabel 1. Nilai *Tuning Constant* untuk Setiap *Breakdown Point*

<i>Breakdown Point</i>	<i>c</i>	<i>K</i>
50%	1,547	0,1995
45%	1,756	0,2312
40%	1,988	0,2634
35%	2,251	0,2957
30%	2,560	0,3278
25%	2,937	0,3593
20%	3,420	0,3899
25%	4,096	0,4194
10%	5,182	0,4475

2.6 Metode S-Estimator

Metode ini dikenal sebagai *S-estimator* karena penduga yang dihasilkan berdasarkan perkiraan skala. Cara kerja dari metode *S-estimator* sama seperti MKT (Metode Kuadrat Terkecil) yakni bertujuan untuk meminimalkan ragam galat^[19]. Metode *S-estimator* didefinisikan sebagai berikut:

$$\beta_s = \arg \beta \min \sigma_s(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) \quad (2.9)$$

ε_i merupakan residual ke- i dari β dan $\sigma_s(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$ didefinisikan sebagai solusi dari:

$$\frac{1}{n} \sum_{i=1}^n \rho\left(\frac{\varepsilon_i}{\sigma_s}\right) = \frac{1}{n} \sum_{i=1}^n \rho\left(\frac{Y_i - \sum_{j=0}^k X_{ij}\beta}{\sigma_s}\right) = K. \quad (2.10)$$

Dengan K merupakan suatu konstanta yang didefinisikan sebagai $K = 0,1995$ agar mencapai *breakdown point* 50%. Sedangkan σ_s merupakan skala *robust*.

Dengan skala *robust* (σ_s)

$$\sigma_s = \sqrt{\frac{1}{nK} \sum_{i=1}^n w_i \varepsilon_i^2}. \quad (2.11)$$

Pembobot $w_i = w_\sigma(u_i) = \frac{\rho(u_i)}{u_i^2}$ dan dipilih estimasi awal

$$\sigma_s = \frac{\text{median}|\varepsilon_i - \text{median}(\varepsilon_i)|}{0,6745}. \quad (2.12)$$

Untuk menyelesaikan persamaan (2.10) dengan mencari turunan parsial pertama terhadap β sehingga diperoleh

$$\begin{aligned} \sum_{i=1}^n \rho' \left(\frac{Y_i - \sum_{j=0}^k X_{ij}\beta}{\sigma_s} \right) &= 0; & j &= 0, 1, 2, \dots, k \\ \sum_{i=1}^n X_{ij} \psi \left(\frac{Y_i - \sum_{j=0}^k X_{ij}\beta}{\sigma_s} \right) &= 0; & j &= 0, 1, 2, \dots, k \end{aligned} \quad (2.13)$$

Ψ disebut fungsi pengaruh yang merupakan turunan dari ρ ($\rho' = \Psi$).

Persamaan (2.13) dapat diselesaikan dengan IRLS (*Iteratively Reweighted Least Square*) sehingga dapat ditulis menjadi

$$\sum_{i=1}^n X_{ij} w_i \left(Y_i - \sum_{j=0}^k X_{ij} \beta \right) = 0; \quad j = 0, 1, 2, \dots, k. \quad (2.14)$$

Dalam menggunakan IRLS, diasumsikan bahwa suatu estimasi awal β^0 dan $\hat{\sigma}_i$ suatu estimasi skala ada. j adalah jumlah parameter yang akan diestimasi, maka persamaan (2.14) menjadi

$$\sum_{i=1}^n X_{ij} w_i^0 \left(Y_i - \sum_{j=0}^k X_{ij} \beta^0 \right) = 0; \quad j = 0, 1, 2, \dots, k. \quad (2.15)$$

Dengan β^0 estimasi parameter pada iterasi pertama dan w_i^0 adalah nilai bobot pada iterasi awal. Dalam notasi matriks, persamaan (2.15) dapat ditulis menjadi

$$\mathbf{X}^T \mathbf{W} \mathbf{X} \boldsymbol{\beta} = \mathbf{X}^T \mathbf{W} \mathbf{Y} \quad (2.16)$$

dengan $\mathbf{W}_{(n \times n)}$ dengan elemen-elemen diagonal yang berisi pembobot.

Penyelesaian persamaan (2.16) tersebut akan memberikan *estimator* untuk $\hat{\boldsymbol{\beta}}$ yaitu:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T_{((k+1) \times n)} \mathbf{W}_{(n \times n)} \mathbf{X}_{(n \times (k+1))})^{-1} (\mathbf{X}^T_{((k+1) \times n)} \mathbf{W}_{(n \times n)} \mathbf{Y}_{(n \times 1)}).$$

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & X_{11} & X_{12} & \dots & X_{1k} \\ 1 & X_{21} & X_{22} & \dots & X_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \dots & X_{nk} \end{bmatrix}, \quad \hat{\boldsymbol{\beta}} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix},$$

$$\mathbf{W} = \begin{bmatrix} w_1 & 0 & 0 & \dots & 0 \\ 0 & w_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & w_n \end{bmatrix}.$$

Adapun tahapan iterasi metode *S-estimator* adalah sebagai berikut:

1. Melakukan estimasi parameter regresi pada data dengan menggunakan MKT.
2. Menguji asumsi klasik dari model regresi.
3. Mendeteksi adanya pencilan dalam data.
4. Mengestimasi koefisien regresi *robust* dengan menggunakan metode *S-estimator*.
 - a. Menghitung parameter β_0 dengan MKT.
 - b. Menghitung nilai sisaan ε_i .

c. Menghitung nilai

$$\hat{\sigma}_i = \begin{cases} \frac{\text{median}|\varepsilon_i - \text{median}(\varepsilon_i)|}{0,6745} & ; \text{iterasi} = 1. \\ \sqrt{\frac{1}{nk} \sum_{i=1}^n w_i \varepsilon_i^2} & ; \text{iterasi} > 1. \end{cases}$$

d. Menghitung nilai $u_i = \frac{\varepsilon_i}{\hat{\sigma}_i}$.

e. Menghitung pembobot $w(u)$

a. Pembobot *Welsch*

$$w(u) = \exp\left(-\left(\frac{u}{c}\right)^2\right).$$

b. Pembobot *Tukey Bisquare*

$$w(u) = \begin{cases} \left(1 - \left(\frac{u}{1,547}\right)^2\right)^2 & ; |u| \leq 1,547. \\ 0 & ; |u| > 1,547. \end{cases}$$

f. Menghitung parameter $\hat{\beta}$

$$\beta^{m+1} = (X^T W^m X)^{-1} (X^T W^m Y).$$

g. Mengulangi langkah b hingga f sampai diperoleh nilai $\hat{\beta}$ yang konvergen.

Dimana selisih nilai β^{m+1} dengan β^m mendekati atau sama dengan nol.

h. Uji hipotesis untuk mengetahui apakah variabel bebas mempunyai pengaruh yang signifikan terhadap variabel tak bebas.

Iterasi pertama pada metode regresi *robust S-estimator* akan bergantung dengan fungsi penduga awal Metode Kuadrat Terkecil (MKT), sedangkan iterasi kedua dan seterusnya bergantung dengan fungsi pada iterasi sebelumnya. Iterasi akan terus berlanjut hingga konvergen atau selisih nilai β^{m+1} dengan β^m mendekati atau sama dengan nol^[20].

III. METODOLOGI PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian ini dilakukan pada semester ganjil tahun akademik 2020/2021, bertempat di Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung.

3.2 Data Penelitian

Data yang digunakan dalam penelitian ini merupakan data simulasi berukuran 30, 60, 100 dan 200 dengan pengulangan sebanyak 1000. Pencilan disimulasikan untuk kondisi jika terdapat pencilan 5%, 10%, 15%, 20%, 25% dan 30%. Untuk n data 30 pencilan yang disimulasikan 10%, 20% dan 30%.

3.3 Metode Penelitian

Penelitian ini dilakukan dengan teknik pengumpulan studi literatur secara sistematis, yang bersumber dari buku-buku, jurnal serta media penunjang lainnya. Topik yang dibahas terkait regresi *robust S-estimator* dengan pembobot *Welsch* dan *Tukey Bisquare*. Perangkat lunak yang digunakan adalah R-4.0.3.

Adapun tahapan yang akan dilakukan pada penelitian ini adalah sebagai berikut:

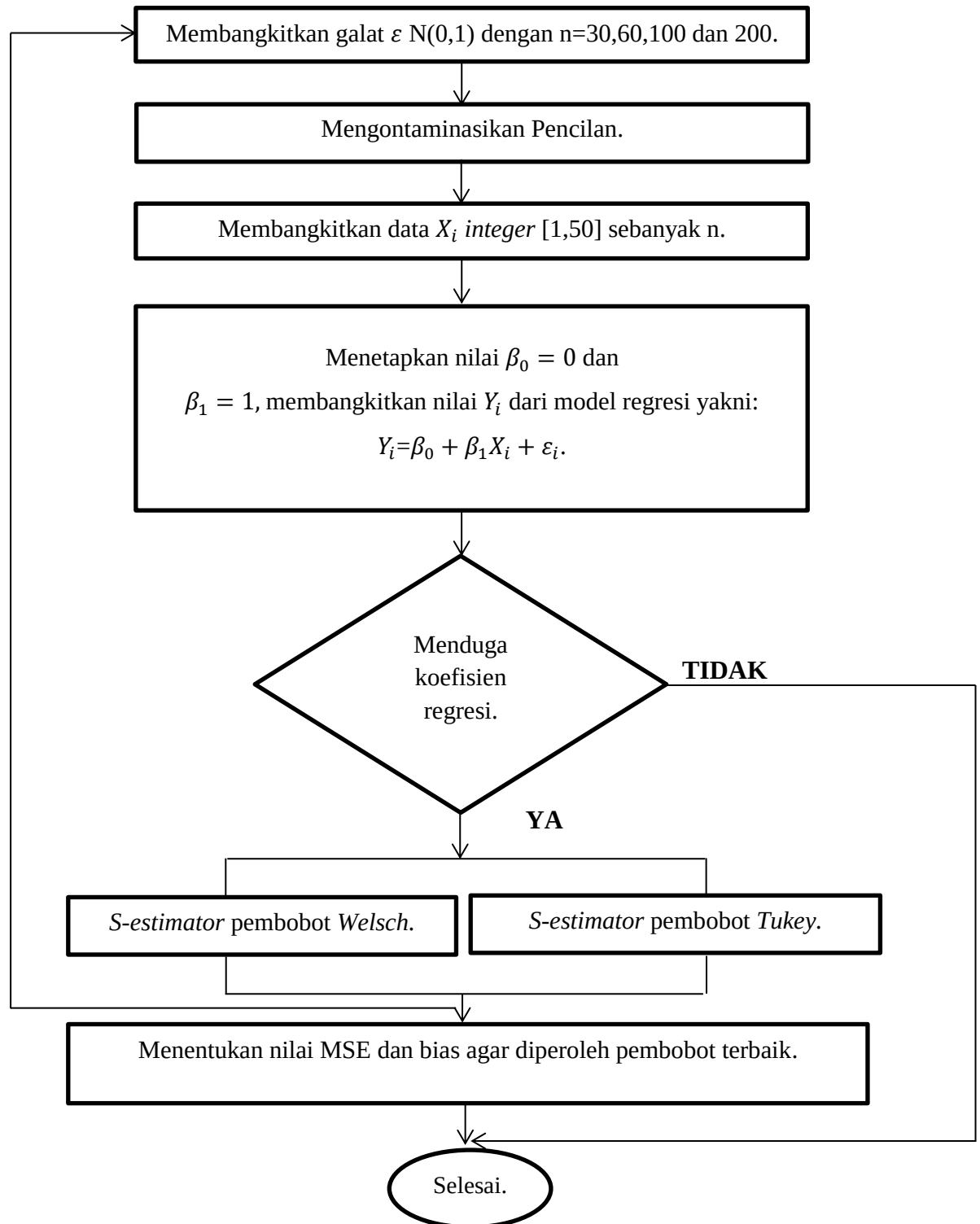
1. Membangkitkan satu kelompok galat ε dari sebaran normal baku $N(0,1)$ berukuran 30, 60, 100 dan 200.
2. Mengontaminasikan kelompok galat dari sebaran $N(0,1)$ dengan sebaran $N(50,1)$. Pencilan yang diberikan yaitu 5%, 10%, 15%, 20%, 25% dan 30% dengan j persentase pencilan dan n jumlah data, banyaknya pencilan didefinisikan sebagai $\text{pencilan} = (j \times n)/100$. Untuk n data 30 pencilan yang disimulasikan 10%, 20% dan 30%.
3. Membangkitkan data X_i sebagai variabel bebas dengan *random integer* pada interval $[1,50]$. Banyaknya tergantung jumlah data yang diamati.
4. Menetapkan nilai $\beta_0 = 0$ dan $\beta_1 = 1$, membangkitkan nilai Y_i dari model regresi yakni: $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$.
5. Menduga koefisien regresi dengan *S-estimator* menggunakan pembobot *Welsch* dan *Tukey Bisquare*.
6. Ulangi langkah 1 sampai 5 sebanyak 1000 kali untuk seluruh jumlah data.
7. Menghitung nilai MSE (*Mean Square Error*) dan bias untuk metode *S-estimator* dengan masing-masing pembobot menggunakan rumus sebagai berikut:

$$MSE = \frac{1}{m} \sum_{i=1}^m (1 - \hat{\beta}_i)^2, m = 1000.$$

$$Bias = \frac{1}{m} \sum_{i=1}^m |1 - \hat{\beta}_i|, m = 1000.$$

8. Membandingkan MSE dan bias dari metode *S-estimator* pembobot *Welsch* dan *Tukey Bisquare*.

Ulangi Sebanyak m



Gambar. 1 Diagram Alir Metodologi Penelitian.

V. KESIMPULAN

Kesimpulan dari penelitian ini adalah:

1. Pembobot *Welsch* dan *Tukey Bisquare* memiliki fungsi objektif yang berbeda. Meskipun demikian, nilai parameter $\hat{\beta}$ yang diperoleh menunjukkan nilai yang tidak jauh berbeda.
2. Nilai MSE dan bias untuk pembobot *Welsch* dan *Tukey Bisquare* pada regresi *robust* metode *S-estimator* tidak memiliki perbedaan yang cukup signifikan. Sehingga dapat dikatakan bahwa kedua pembobot tersebut sama baiknya jika digunakan pada metode *S-estimator*.

DAFTAR PUSTAKA

- [1] Herawati, N., Nisa, K., dan Setiawan, E. 2011. Analisis Ketegaran Regresi Robust Terhadap Letak Pencilan: Studi Perbandingan. *Bulletin of Mathematics*. (1): 49-60.
- [2] Cahyawati, D., Tanuji, H., dan Abdiati, R. 2009. Efektivitas Metode Regresi Robust Penduga Welsch dalam Mengatasi Pencilan pada Pemodelan Regresi Linear Berganda. *Jurnal Penelitian Sains*. 12 (1): 1-7.
- [3] Romdi, Wahyuningsih, S., dan Yuniarti, D. 2015. Regresi Robust Linear Sederhana dengan Menggunakan Estimasi MM (Method of Moment). *Jurnal Eksponensial*. 6 (2): 179-186.
- [4] Olive, D.J. 2008. *Applied Robust Statistics*. Southern Illinois University, Carbondale.
- [5] Chen, C. 2002. *Robust Regression and Outlier Detection with the ROBUSTREG Procedure*. Paper 265-27. SAS Institute Inc., Cary, NC.
- [6] Setiarini, Z. dan Listyani, E. 2017. Analisis Regresi Robust Estimasi-S Menggunakan Pembobot Welsch dan Tukey Bisquare. *Jurnal Matematika*. 6 (1): 48-55.
- [7] Setiawan, W., Nessyana, N., dan Sulistianingsih, E. 2019. Metode Estimasi-S pada Analisis Regresi Robust dengan Pembobotan Tukey Bisquare. *Bimaster*. 8 (2): 289-296.
- [8] Sembiring, R.K. 1995. *Analisis regresi*. Penerbit ITB, Bandung.
- [9] Ghozali, I. 2011. *Aplikasi Analisis Multivariate dengan Progam SPSS*. Badan Penerbit Universitas Diponegoro, Semarang.

- [10] Widyaningsih, A., Susilawati, M., dan Sumarjaya, I.W. 2014. Estimasi Model Seemingly Unrelated Regression (SUR) dengan Metode Generalized Least Square (GLS). *Jurnal Matematika*. 4 (2): 102-110.
- [11] Basuki, A.T. dan Prawoto, N. 2015. *Analisis Regresi dalam Penelitian Ekonomi dan Bisnis*. Raja Grafindo Persada, Depok.
- [12] Montgomery, D.C., Peck, E.A., and Vining, G.G. 2012. *Introduction to Linear Regression Analysis*. Fifth Edition. John Wiley and Sons, Inc., New York.
- [13] Maharani, I.F., Satyahadewi, N., dan Kusnandar, D. 2014. Metode Ordinary Least Square dan Least Trimmed Squares dalam Mengestimasi Parameter Regresi Ketika Terdapat Outlier. *Bimaster*. 3 (3): 163-168.
- [14] Desgagné, A. 2020. Efficient and Robust Estimation of Regression and Scale Parameters with Outlier Detection. *Computational Statistics and Data Analysis*. 1-19.
- [15] Rousseeuw, P. and Leroy, A.M. 2005. *Robust Regression and Outlier Detection*. Vol.589. John Wiley & Sons, New York.
- [16] Alma, G.Ö. 2011. Comparison of Robust Regression Methods. *Int. J. Contemp. Math. Sciences*. 6 (9): 409-421.
- [17] Holland, P.W. and Welsch, R.E. 1977. Robust Regression Using Iteratively Reweighted Least-square. *Communications in Statistics-Theory and Method*. 6 (9): 813-827.
- [18] Fox, J. 2002. *Robust Regression*. Appendix to An R and S-PLUS Companion to Applied Regression.
- [19] Rousseeuw, P. and Yohai, V. 1984. *Robust Regression by Means of S Estimators*. Paper 256-272. Springer Verlag, New York.
- [20] Pitselis, G. 2013. A Review on Robust Estimator Applied to Regression Credibility. *Journal of Computational and Applied Mathematics*. 231-249.

- [21] Njodi, I. 2019. *Outlook on Human Capacity Building and Development: A Handbook of Research in Honour of Professor Ibrahim Njodi edited by Jacinta A. Opara*. University of Maiduguri Press, Nigeria. 532-538.
- [22] Usman, M. 2016. On Consistency and Limitation of Independent t-test Kolmogorov Smirnov Test and Mann Whitney U test. *IOSR Journal of Mathematics (IOSR-JM)*. 12 (4): 22-27.