



ANALISIS REGRESI ROBUST ESTIMASI-M PEMBOBOT *HUBER* DAN *TUKEY BISQUARE* PADA TINGKAT KEMISKINAN INDONESIA

ROBUST REGRESSION ANALYSIS OF HUBER AND TUKEY BISQUARE WEIGHTED M-ESTIMATES ON THE INDONESIAN POVERTY RATE

Adelia Damayanti, Prodi Matematika FMIPA UNY
Mathilda Susanti*, Prodi Matematika FMIPA UNY
*e-mail: mathilda@uny.ac.id

Abstrak

Regresi robust merupakan metode yang penting untuk menganalisis data yang mengandung *outlier*. Untuk mengetahui faktor-faktor yang mempengaruhi tingkat kemiskinan di Indonesia, dapat diterapkan analisis regresi linear dengan melihat pencilan atau *outlier*. Metodenya, yaitu dengan metode estimasi regresi robust estimasi M. Penelitian ini bertujuan untuk membandingkan keefektifan antara model regresi robust estimasi-M pembobot *huber* dan *tukey bisquare* untuk mengatasi *outlier* serta mengetahui faktor yang mempengaruhi tingkat kemiskinan di Indonesia tahun 2021. Berdasarkan uji signifikansi yang dilakukan, variabel independen yang paling berpengaruh terhadap tingkat kemiskinan di Indonesia adalah penduduk dengan sumber penerangan listrik PLN dan tingkat pengangguran terbuka. Dari dua metode estimasi-M yang dipilih, metode estimasi-M pembobot *huber* menghasilkan nilai RSE 3,706 dan *adj R-square* 53,15% sedangkan pembobot *tukey bisquare* menghasilkan nilai RSE 3,294 dan *adj R-square* 51,59%. Berdasarkan data tersebut, dapat disimpulkan bahwa metode regresi robust estimasi-M pembobot *huber* lebih efektif digunakan untuk mengatasi *outlier* pada data tingkat kemiskinan di Indonesia.

Kata kunci: kemiskinan, pencilan (*outlier*), regresi robust, estimasi-M.

Abstract

Robust regression is an important method for analyzing data containing outliers. To find out the factors that affect the poverty rate in Indonesia, linear regression analysis can be applied by looking at outliers. The method, namely the M-estimation robust regression estimation method. This study aims to compare the effectiveness between the huber and tukey bisquare weighted M-estimation robust regression models to overcome outliers and determine the factors that affect the poverty rate in Indonesia in 2021. Based on the significance test conducted, the independent variables that have the most influence on the poverty rate in Indonesia are the population with a PLN electricity source and the open unemployment rate. Of the two M-estimation methods chosen, the huber weighting M-estimation method produces an RSE value of 3.706 and an adj R-square of 53.15% while the tukey bisquare weighting produces an RSE value of 3.294 and an adj R-square of 51.59%. Based on these data, it can be concluded that the huber-weighted M-estimation robust regression method is more effective in overcoming outliers in poverty rate data in Indonesia.

Keywords: poverty, outlier, robust regression, M-estimation.

PENDAHULUAN

Analisis regresi linear merupakan suatu metode statistik yang dapat digunakan untuk menyelidiki dan memodelkan hubungan antar variabel. Analisis regresi linear memiliki keunggulan, diantaranya mampu melakukan generalisasi dan ekstraksi dari pola data tertentu dan melakukan perhitungan secara paralel (Amrin, 2016). Dalam menentukan koefisien regresi, metode yang paling umum digunakan adalah *ordinary least squares* (OLS) atau metode kuadrat terkecil. Metode OLS harus memenuhi beberapa asumsi klasik yang akan menghasilkan model regresi yang bersifat BLUE (*Best Linear Unbiased Estimator*). Selain itu, untuk memperoleh model regresi yang baik sebagai syarat penggunaan metode OLS dalam memenuhi BLUE adalah tidak ditemukan adanya pencilan (*outlier*) dalam data pengamatan.

Pencilan atau *outlier* merupakan pengamatan yang berbeda dari data lainnya (Montgomery & Peck, 2012). Keberadaan *outlier* dapat mengakibatkan estimasi parameter regresi yang diperoleh menjadi kurang optimal karena memberikan nilai *error* yang besar. Hal ini dapat mempengaruhi hasil analisis data yang diperoleh. Meskipun demikian, data *outlier* tidak dapat dibuang percuma karena dapat mempengaruhi model regresi. Oleh karena itu, diperlukan suatu metode untuk mengestimasi koefisien regresi yang dapat mengatasi *outlier* pada data tanpa harus membuang data *outlier*. Metode regresi yang dapat menangani data yang mengandung *outlier* adalah regresi robust.

Regresi *robust* memiliki keuntungan, yaitu metode ini parameternya tidak mudah dipengaruhi oleh keberadaan *outlier*. Untuk mengestimasi parameter regresi *robust* dapat digunakan beberapa metode estimasi, salah satunya adalah metode estimasi M (*Maximum Likelihood type*) (Chen, 2002). Metode estimasi-M merupakan metode *robust* yang paling sederhana serta memiliki tingkat efisiensi tertinggi yang ditentukan melalui nilai ragam/varians minimum suatu penduga. (Aflakhah *et al*, 2019). Dalam penerapannya, metode estimasi-M menggunakan dua fungsi pembobot yaitu pembobot *huber* dan *tukey bisquare*. Telah banyak penelitian yang meneliti efektivitas penggunaan metode *huber* dan *tukey bisquare*. Penelitian mengenai efektivitas penggunaan metode *huber* dan *tukey bisquare* salah satunya dilakukan oleh Pradewi & Sudarno (2012). Penelitian ini melakukan kajian estimasi-M IRLS menggunakan fungsi pembobot *huber* dan *tukey bisquare* pada data ketahanan pangan di Jawa Tengah. Perbandingan ini dilakukan karena kedua pembobot memiliki efisiensi yang tinggi dibandingkan dengan LAD (*Least Absolut Deviation*) yang sebesar 0,64 ketika error terdistribusi normal. Pembobot *huber* dan *tukey bisquare* menghasilkan efisiensi relatif sekitar 95% ketika densitas kesalahan normal dengan penggunaan $c = 1,345$ dan $c = 4,685$ pada persamaan *robust loss function* (Yu *et. al*, 2017). Dengan demikian, efisiensi relatif yang sama antara kedua pembobot ini mendorong peneliti untuk mengetahui fungsi pembobot terbaik dari dua pembobot ini dalam berbagai kasus.

Regresi linier dapat digunakan untuk menyelesaikan masalah sosial. Azizah *et. al* (2022) dan Cahyandari *et. al.* (2013) melakukan penelitian dengan menggunakan regresi linier untuk memprediksi Indeks Pembangunan Manusia (IPM), inflasi, produksi pangan, pengangguran, dan kemiskinan (Azizah & Wachidah, 2022; Cahyandari & Hisani, 2013). Kemiskinan dapat dilihat sebagai ketidakmampuan dari sisi ekonomi untuk memenuhi kebutuhan dasar, baik kebutuhan makanan maupun bukan makanan (BPS, 2021). Berdasarkan data yang diperoleh dari Badan Pusat Statistik, pada bulan September 2020 terjadi kenaikan jumlah penduduk miskin sebesar 1,13 juta jiwa. Selain itu, pada bulan Maret 2021, jumlah penduduk miskin di Indonesia mengalami kenaikan sebesar 1,12 juta jiwa atau sebesar 0,36%. Berdasarkan data tersebut, tingkat kemiskinan di Indonesia tidak juga mengalami penurunan yang signifikan. Untuk mengatasi permasalahan ini, perlu diketahui faktor-faktor yang lebih akurat dalam mempengaruhi tingkat kemiskinan di Indonesia untuk menghasilkan suatu kebijakan yang efektif untuk menurunkan tingkat kemiskinan di Indonesia.

Beberapa penelitian telah dilakukan untuk mengetahui faktor-faktor yang mempengaruhi tingkat kemiskinan di Indonesia. Penelitian yang dilakukan oleh Tisniwati (2012) menunjukkan bahwa pemodelan prediksi tingkat kemiskinan di Indonesia menggunakan metode regresi linear berganda menghasilkan satu dari tiga variabel bebas berpengaruh signifikan terhadap tingkat kemiskinan di Indonesia yaitu variabel angka harapan hidup. Berdasarkan penelitian di atas dapat diketahui bahwa metode regresi linear dapat diterapkan untuk mengetahui faktor-faktor yang mempengaruhi tingkat kemiskinan di Indonesia.

Berdasarkan keunggulan metode regresi robust estimasi-M pembobot *huber dan tukey bisquare*, penelitian ini akan membahas keefektifan dari regresi robust tersebut untuk mengetahui indikator apa saja yang berpengaruh terhadap tingkat kemiskinan di Indonesia.

METODE

Penelitian ini menggunakan data kuantitatif berupa data tingkat kemiskinan di Indonesia tahun 2021. Metode yang digunakan dalam pengumpulan data yaitu pengumpulan data sekunder tentang kemiskinan provinsi di Indonesia dan indikator-indikator pengaruh kemiskinan pada tahun 2021 yang diperoleh dari Badan Pusat Statistik Indonesia. Variabel yang digunakan dalam penelitian ini ditunjukkan pada Tabel 1.

Tabel 1. Variabel penelitian

Variabel	Keterangan
Y	Persentase penduduk miskin di Indonesia
X_1	Persentase tingkat penyelesaian pendidikan SMA
X_2	Persentase penduduk dengan sumber air minum layak
X_3	Persentase penduduk dengan sumber penerangan listrik PLN
X_4	Persentase penduduk dengan sanitasi layak
X_5	PDRB per kapita
X_6	Persentase tingkat pengangguran terbuka
X_7	Persentase <i>unmet need</i> pelayanan kesehatan

Langkah pertama yang dilakukan yaitu melakukan pengujian uji asumsi klasik regresi. Uji asumsi klasik regresi pada penelitian ini terdiri dari empat uji yaitu pengujian normalitas, pengujian multikolinearitas, pengujian homoskedastisitas, dan pengujian nonautokorelasi.

Langkah kedua yaitu menentukan nilai estimasi parameter untuk membuat model regresi berganda dengan metode *Ordinary Least Square* (OLS) yang dinotasikan dengan persamaan yaitu (Montgomery & Peck, 2012)

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

Berikutnya adalah pendeteksian *outlier* dengan metode grafis (*scatter plot*) dan perhitungan nilai *DfFITS*. Adapun persamaan *DfFITS* dapat dinotasikan sebagai berikut (Montgomery & Peck, 2012)

$$DfFITS_i = e_i \left(\frac{n-p-1}{SSE(1-h_{ii}) - e_i^2} \right)^{\frac{1}{2}} \left(\frac{h_{ii}}{1-h_{ii}} \right)^{\frac{1}{2}} ; i = 1, 2, \dots, n$$

dengan e_i adalah nilai residual ke- i , SSE adalah jumlah kuadrat residual, n adalah banyak data penelitian, p adalah banyak variabel independen, dan h_{ii} adalah nilai *leverage* berupa elemen dari diagonal matriks hat H yang dirumuskan sebagai $h_{ii} = X_i'(X'X)^{-1}X_i$ dan kriteria

keputusan dari pengujian *DfFITS*

$$DfFITS_i = \begin{cases} \leq 2\sqrt{\frac{p}{n}} & , \text{ pencilan ke } - i \text{ tidak berpengaruh} \\ > 2\sqrt{\frac{p}{n}} & , \text{ pencilan ke } - i \text{ berpengaruh} \end{cases}$$

Langkah selanjutnya, yaitu menentukan nilai estimasi parameter dengan metode regresi *robust* estimasi-M dengan langkah-langkah sebagai berikut.

1. Menghitung nilai residual e_i .
2. Menghitung nilai $\hat{\sigma}$ dan nilai skala residual u_i

$$\hat{\sigma} = \frac{MAD}{0,6745} = \frac{\text{median}|e_i - \text{median } e_i|}{0,6745}$$

$$u_i = \frac{(Y_i - \hat{Y}_i)}{\hat{\sigma}} = \frac{e_i}{\hat{\sigma}}$$

dengan MAD adalah *Median Absolute Deviation*.

3. Menghitung nilai pembobot awal $w_i = w(u_i)$ dengan konstanta untuk pembobot *huber* yaitu $c = 1,345$ dan *tukey bisquare* yaitu $c = 4,685$. Adapun persamaan fungsi obyektif dan fungsi pembobot *huber* dan *tukey bisquare* ditunjukkan pada Tabel 2.

Tabel 2. Fungsi obyektif dan fungsi pembobot *huber* dan *tukey bisquare*

Pembobot	Fungsi Obyektif	Fungsi Pembobot
<i>Huber</i>	$\rho(e_i) = \begin{cases} \frac{1}{2}u_i^2 & , u_i \leq c \\ c u_i - \frac{1}{2}c^2 & , u_i > c \end{cases}$	$w_i = w(u_i) = \frac{\psi(u_i)}{u_i} = \begin{cases} 1 & , u_i \leq c \\ \frac{c}{ u_i } & , u_i > c \end{cases}$
<i>Tukey Bisquare</i>	$\rho(e_i) = \begin{cases} \frac{c^2}{6} \left\{ 1 - \left[1 - \left(\frac{u_i}{c} \right)^2 \right]^2 \right\} & , u_i \leq c \\ \frac{c^2}{6} & , u_i > c \end{cases}$	$w_i = w(u_i) = \frac{\psi(u_i)}{u_i} = \begin{cases} \left[1 - \left(\frac{u_i}{c} \right)^2 \right]^2 & , u_i \leq c \\ 0 & , u_i > c \end{cases}$

4. Menghitung estimasi $\hat{\beta}$ dengan pembobot w_i sehingga diperoleh $\hat{\beta}$ yang baru pada iterasi ke-1 yang dinamakan nilai estimasi parameter *robust* $\hat{\beta}_m$ dengan persamaan

$$\hat{\beta}_m = (X'W_0X)^{-1}X'W_0Y$$
 dengan W_0 adalah diagonal dari matriks pembobot w_i berukuran $n \times n$.
5. Melakukan iterasi sehingga nilai w_i akan berubah di setiap iterasi sehingga diperoleh nilai $\hat{\beta}_m$ yang konvergen.

Langkah selanjutnya yaitu melakukan uji signifikansi parameter menggunakan uji F (simultan) dan uji t (parsial) untuk mengetahui adanya pengaruh antara variabel independen faktor-faktor pengaruh kemiskinan dengan variabel dependen tingkat kemiskinan di Indonesia pada model regresi. Persamaan uji F dan uji t dapat dilihat pada Tabel 3.

Tabel 3. Uji signifikansi parameter

Uji signifikansi	Persamaan
<i>Uji F</i>	$\frac{\sum(\hat{Y}_i - \bar{Y})^2 / (K - 1)}{\sum(Y_i - \hat{Y}_i)^2 / (n - K)}$
<i>Uji t</i>	$\frac{b_i}{Sb_i}$

Langkah berikutnya adalah menghitung nilai koefisien determinasi *adjusted R – squared* R_a^2 dari model regresi robust pembobot *huber* dan *tukey bisquare*. Selanjutnya, akan dilakukan analisis perbandingan nilai R_a^2 dari kedua model regresi robust tersebut untuk menentukan model terbaik dengan perolehan nilai R_a^2 terkecil.

HASIL DAN PEMBAHASAN

Hasil

1. Uji Asumsi Klasik

Dalam membentuk model persamaan regresi, perlu dilakukan uji asumsi klasik untuk mengetahui apakah model regresi yang akan dibentuk sudah memenuhi asumsi atau tidak. Uji asumsi klasik yang dilakukan pada penelitian ini antara lain uji normalitas, uji multikolinearitas, uji heterokedastisitas, dan uji non autokorelasi.

a. Uji Normalitas

Uji normalitas digunakan untuk mengetahui apakah residual pada model berdistribusi normal atau tidak. Pengujian kenormalan dapat dilakukan menggunakan uji *Kolmogorov-Smirnov* yang hasilnya dapat dilihat pada Tabel 4.

Tabel 4. Nilai *Kolmogorov Smirnov*

<i>df</i>	<i>p – value</i>
33	0,0324

Hipotesis nol dari uji ini yaitu residual berdistribusi normal. Berdasarkan Tabel 4 di atas diperoleh *p-value* untuk uji *Kolmogorov-Smirnov* yaitu sebesar 0,0324. Berdasarkan kriteria keputusan, *p – value* yang diperoleh lebih kecil dari $\alpha = 0,05$ yaitu $0,0324 < \alpha = 0,05$ sehingga H_0 ditolak (Siregar, 2017). Dengan demikian, dapat disimpulkan bahwa residual tidak berdistribusi normal pada taraf signifikansi 5%.

b. Uji Multikolinearitas

Uji multikolinearitas digunakan untuk menguji apakah dalam model regresi ditemukan adanya korelasi antar variabel independen. Multikolinearitas dapat dilihat dari nilai *Tolerance* (TOL) dan *Variabel Inflation Factor* (VIF). Nilai VIF lebih besar dari 10 menunjukkan adanya multikolinearitas pada data. Hasil pengujian multikolinearitas dapat dilihat pada Tabel 5.

Tabel 5. Nilai TOL dan VIF

<i>variabel independen</i>	<i>TOL</i>	<i>VIF</i>
X_1	0,2710	3,689
X_2	0,4953	2,018
X_3	0,2498	4,003
X_4	0,2511	3,980
X_5	0,5467	1,829
X_6	0,3666	2,727
X_7	0,4753	2,103

Hipotesis nol dari uji ini yaitu tidak terjadi multikolinearitas pada model regresi. Berdasarkan Tabel 5, nilai VIF yang diperoleh dari masing-masing variabel independen lebih kecil dari 10 sehingga H_0 diterima. Dengan demikian, dapat disimpulkan bahwa tidak terjadi multikolinearitas antar variabel independen dalam model regresi.

c. Uji Heteroskedastisitas

Uji heteroskedastisitas digunakan untuk menguji apakah varians dari error tidak konstan (heteroskedastik) untuk semua variabel independen. Hasil pengujian homoskedastisitas dilakukan dengan menggunakan uji Glejser yang hasilnya dapat dilihat pada Tabel 6.

Tabel 6. Hasil p -value masing-masing variabel independent pada uji Glejser

Variabel independen	p - value	Keterangan
X_1	0,637	homoskedastisitas
X_2	0,176	homoskedastisitas
X_3	0,834	homoskedastisitas
X_4	0,905	homoskedastisitas
X_5	0,893	homoskedastisitas
X_6	0,721	homoskedastisitas
X_7	0,779	homoskedastisitas

Berdasarkan Tabel 6, diperoleh p - value dari ketujuh variabel independen memiliki p - value $> \alpha = 0,05$ sehingga dapat disimpulkan bahwa H_0 diterima. Dengan demikian, data sudah bersifat homokedastik atau tidak terjadi heteroskedastisitas.

d. Uji Non Autokorelasi

Pengujian nonautokorelasi pada penelitian ini dilakukan dengan uji Durbin-Watson. Berikut hasil pengujian autokorelasi pada Tabel 7.

Tabel 7. Output hasil uji autokorelasi (*durbin-watson test*)

DW value	p - value
2,0827	0,4474

Berdasarkan Tabel 7, diperoleh p -value untuk uji *durbin watson* yaitu 0,4474. Berdasarkan kriteria keputusan, p - value yang diperoleh lebih besar dari $\alpha = 0,05$ yaitu $0,4474 > \alpha = 0,05$ sehingga H_0 diterima. Dengan demikian, dapat disimpulkan bahwa model yang dihasilkan bebas autokorelasi.

Berdasarkan keempat uji asumsi klasik yang telah dilakukan, dapat diketahui bahwa terdapat salah satu asumsi klasik yang tidak terpenuhi yaitu uji normalitas. Menurut Montgomery & Peck (2012), tidak terpenuhinya salah satu asumsi tersebut menunjukkan adanya pencilan pada model. Dengan demikian, akan dilakukan estimasi parameter β dengan metode OLS yang bertujuan untuk membentuk model persamaan regresi awal yang mengandung *outlier*.

2. Estimasi Parameter β dengan Metode *Ordinary Least Square* (OLS)

Setelah dilakukan uji asumsi regresi, langkah selanjutnya adalah melakukan estimasi parameter β dengan metode OLS. Berdasarkan hasil *output* dengan program R, diperoleh nilai koefisien variabel antara variabel dependen dan variabel independen pada data tingkat kemiskinan provinsi di Indonesia tahun 2021 sebagai berikut.

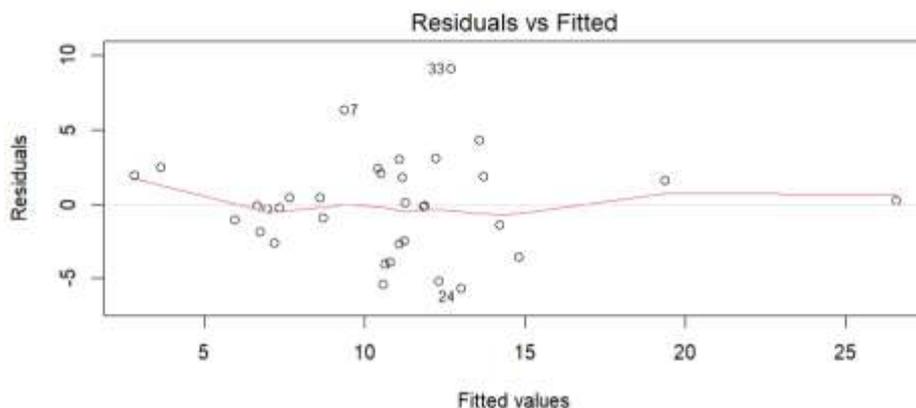
$$\hat{Y} = 44,0473 + 0,0360X_1 + 0,0481X_2 - 0,3804X_3 - 0,1488X_4 - 0,0740X_5 + 0,2548X_6 + 0,1485X_7$$

3. Pendeteksian *Outlier*

Pendeteksian *outlier* pada penelitian ini dilakukan dengan dua metode, yaitu *scatter plot* dan metode *DfFITS*.

a. *Scatter Plot*

Berdasarkan *scatter plot* pada Gambar 1, dapat dilihat bahwa data ke-7, 24, dan 33 mempunyai residual yang besar dan ketiga data tersebut terletak jauh dari pola kumpulan data. Dengan demikian dapat disimpulkan bahwa data ke-7, 24, dan 33 merupakan *outlier* dari data uji (Soemartini, 2007).



Gambar 1. *Scatter plot* data tingkat kemiskinan provinsi di Indonesia pada tahun 2021 Langkah selanjutnya yaitu akan dilakukan pendeteksian *outlier* dengan perhitungan statistis menggunakan metode *DfFITS*.

b. Metode *DfFITS*

Suatu data dapat dikatakan sebagai *outlier* apabila nilai mutlak *DfFITS*-nya lebih besar dari $2\sqrt{\frac{p}{n}}$ dengan p adalah banyak variabel independent dan n adalah banyak data pengamatan. Pada penelitian ini diketahui $p = 7$ dan $n = 34$. Suatu data dapat dikatakan *outlier* jika nilai $DfFITS > 2\sqrt{\frac{7}{34}} = 0,9074$. Hasil identifikasi *outlier* dapat dilihat pada Tabel 8.

Tabel 8. Identifikasi *outlier*

Data Pengamatan ke – i	$ DfFITS_i $
7	1,7157
14	0,9307
24	1,2672
33	1,2251

Berdasarkan Tabel 8, nilai *DfFITS* pada data ke-7, 14, 24, dan 33 memiliki nilai *DfFITS* yang lebih besar dari 0,9074. Dengan demikian dapat disimpulkan bahwa data ke-7, 14, 24, dan 33 merupakan *outlier* dari data pengamatan (Montgomery & Peck, 2012).

Berdasarkan pendeteksian *outlier* menggunakan metode grafis *scatter plot* dan metode statistis *DfFITS*, dapat disimpulkan bahwa data pengamatan mengandung *outlier* sehingga akan dilakukan pengestimasi parameter regresi menggunakan metode regresi *robust* estimasi-M pembobot *huber* dan *tukey bisquare*.

4. Estimasi Parameter β dengan Estimasi-M Pembobot *Huber*

Setelah terbukti ditemukan adanya pencilan pada data, langkah selanjutnya, yaitu mengestimasi parameter β dengan estimasi-M pembobot *huber*. Perolehan output nilai skala *robust* estimasi-M pembobot *huber* yaitu sebesar 2,912 yang diperoleh dari iterasi ke-7 hingga konvergen. Berikut fungsi untuk pembobot *huber*.

$$w(u_i) = \begin{cases} 1 & , |u_i| \leq 1,345 \\ \frac{1,345}{|u_i|} & , |u_i| > 1,345 \end{cases}$$

Berdasarkan fungsi pembobot *huber* di atas, diperoleh hasil estimasi parameter regresi estimasi-M *huber* yang *robust* (kekar) terhadap *outlier* sebagai berikut.

$$\hat{Y}_{m-huber} = 41,6115 + 0,0317X_1 + 0,0829X_2 - 0,3519X_3 - 0,1906X_4 - 0,0078X_5 + 0,2652X_6 + 0,0617X_7$$

Untuk selanjutnya, akan dilakukan perhitungan model regresi estimasi-M pembobot *tukey bisquare*.

5. Estimasi Parameter β dengan Estimasi-M Pembobot *Tukey Bisquare*

Perolehan output nilai skala *robust* estimasi-M pembobot *tukey bisquare* yaitu sebesar 2,496 yang diperoleh dari iterasi ke-88 hingga konvergen. Berikut fungsi untuk pembobot *tukey bisquare*.

$$w(u_i) = \begin{cases} \left[1 - \left(\frac{u_i}{4,685}\right)^2\right]^2 & , |u_i| \leq 4,685 \\ 0 & , |u_i| > 4,685 \end{cases}$$

Berdasarkan fungsi pembobot *tukey bisquare* di atas, diperoleh hasil estimasi parameter regresi estimasi-M *tukey bisquare* yang *robust* (kekar) terhadap *outlier* sebagai berikut.

$$\hat{Y}_{m-tb} = 41,1161 + 0,0401X_1 + 0,0816X_2 - 0,3073X_3 - 0,2459X_4 - 0,0445X_5 + 0,2789x_6 - 0,0269x_7$$

Setelah dilakukan pemodelan regresi *robust* dengan metode estimasi-M pembobot *huber* dan *tukey bisquare*, langkah selanjutnya yaitu melakukan uji signifikansi parameter dengan uji F (simultan) dan uji t (parsial) untuk melihat adanya pengaruh dari variabel-variabel independent faktor-faktor pengaruh kemiskinan terhadap variabel dependen tingkat kemiskinan di Indonesia pada penelitian ini.

6. Uji Signifikansi Parameter

Uji signifikansi parameter bertujuan untuk melihat faktor-faktor yang berpengaruh terhadap tingkat kemiskinan. Uji signifikansi ini terdiri dari dua bagian, yaitu uji F (uji simultan) dan uji t (uji parsial).

a. Uji F (simultan)

Hipotesis nol pada uji ini yaitu tidak terdapat variabel independen pengaruh tingkat kemiskinan X_i yang secara simultan berpengaruh terhadap variabel dependen tingkat kemiskinan di Indonesia Y model regresi pada taraf signifikansi 5%. Berikut merupakan ringkasan hasil uji F pada model regresi estimasi-M pembobot *huber* dan *tukey bisquare* pada Tabel 9.

Tabel 9. Ringkasan hasil uji F model regresi estimasi-M pembobot *huber* dan *tukey bisquare*

Metode Estimasi-M	F value	p – value	Keputusan	Kesimpulan
<i>Huber</i>	6,314	0,00022	H_0 ditolak	Signifikan
<i>Tukey Bisquare</i>	6,411	0,00019	H_0 ditolak	Signifikan

Berdasarkan Tabel 9, diperoleh hasil yaitu *p-value* pada masing-masing model regresi estimasi-M pembobot *huber* dan *tukey bisquare* lebih kecil daripada taraf signifikansi 5% sehingga H_0 ditolak. Dengan demikian, dapat disimpulkan bahwa terdapat paling sedikit satu variabel independen pengaruh tingkat kemiskinan X_i yang secara simultan mempengaruhi variabel dependen tingkat kemiskinan di Indonesia Y pada model regresi *robust* estimasi-M pembobot *huber* dan *tukey bisquare*.

Setelah dilakukan uji F (simultan) pada model regresi *robust* estimasi-M pembobot *huber* dan *tukey bisquare*, diperoleh kesimpulan yang sama yaitu terdapat paling sedikit satu variabel independen pengaruh tingkat kemiskinan X_i yang secara simultan mempengaruhi variabel dependen tingkat kemiskinan di Indonesia Y pada masing-masing model regresi *robust* yang telah diuji. Oleh karena itu, perlu dilakukan uji lanjut untuk mengetahui variabel-variabel independen apa saja yang berpengaruh secara individu (parsial) terhadap tingkat kemiskinan di Indonesia.

b. Uji t (parsial)

Hipotesis nol pada uji ini yaitu tidak terdapat variabel independent pengaruh tingkat kemiskinan X_i yang secara individu mampu menjelaskan perubahan nilai variabel dependen tingkat kemiskinan di Indonesia Y model regresi pada taraf signifikansi 5%. Berikut merupakan hasil uji t pada model regresi estimasi-M pembobot *huber* pada Tabel 10.

Tabel 10. Hasil perhitungan uji t pada model regresi *robust* estimasi-M pembobot *huber*.

X_i	t – value	p – value	Keputusan	Kesimpulan
X_1	0,266	0,791	H_0 diterima	tidak signifikan
X_2	0,759	0,454	H_0 diterima	tidak signifikan
X_3	0,266	0,009	H_0 ditolak	signifikan
X_4	0,759	0,156	H_0 diterima	tidak signifikan
X_5	0,266	0,963	H_0 diterima	tidak signifikan
X_6	0,759	0,007	H_0 ditolak	signifikan
X_7	0,266	0,906	H_0 diterima	tidak signifikan

Dan berikut merupakan hasil uji t pada model regresi estimasi-M pembobot *tukey bisquare* pada Tabel 11.

Tabel 11. Hasil perhitungan uji t pada model regresi *robust* estimasi-M pembobot *tukey bisquare*.

X_i	$t - value$	$p - value$	Keputusan	Kesimpulan
X_1	0,338	0,738	H_0 diterima	tidak signifikan
X_2	0,748	0,460	H_0 diterima	tidak signifikan
X_3	2,452	0,021	H_0 ditolak	signifikan
X_4	1,884	0,070	H_0 diterima	tidak signifikan
X_5	0,261	0,795	H_0 diterima	tidak signifikan
X_6	3,038	0,005	H_0 ditolak	signifikan
X_7	0,051	0,958	H_0 diterima	tidak signifikan

Setelah dilakukan uji t (parsial) pada model regresi *robust* estimasi-M pembobot *huber* dan *tukey bisquare*, diperoleh kesimpulan yang sama yaitu variabel independen X_3 yaitu persentase penduduk dengan sumber listrik PLN dan X_6 yaitu persentase tingkat pengangguran terbuka berpengaruh secara signifikan terhadap variabel dependen tingkat kemiskinan di Indonesia Y pada masing-masing model regresi *robust* yang telah diuji. Dengan demikian, akan dilakukan perbandingan model regresi *robust* estimasi-M pembobot *huber* dan *tukey bisquare* untuk menentukan metode mana yang memberikan hasil terbaik ditinjau dari nilai *Adjusted R-square* dan *Residual Standard Error* (RSE).

7. Perbandingan Model Regresi *Robust* Estimasi-M *Huber* dan *Tukey Bisquare* Data Tingkat Kemiskinan di Indonesia Tahun 2021

Selanjutnya, akan dilakukan perbandingan antara hasil estimasi parameter menggunakan estimasi-M pembobot *huber* dan *tukey bisquare* untuk menentukan metode mana yang memberikan hasil terbaik terhadap model. Kriteria yang digunakan dalam perbandingan ini adalah dengan membandingkan nilai *Adjusted R-square* dan *Residual Square Error* (RSE). Hasil perhitungan RSE dan *Adjusted R-square* pada masing-masing model regresi estimasi-M dapat dilihat pada Tabel 12.

Tabel 12. Perbandingan model regresi *robust* estimasi-M pembobot *huber* dan *tukey bisquare*.

Metode Estimasi-M	RSE	<i>Adjusted R-Square</i>
<i>Huber</i>	3,706	0,5315
<i>Tukey Bisquare</i>	3,767	0,5199

Pembahasan

Berdasarkan penelitian yang telah dilakukan, diperoleh hasil yaitu terdapat salah satu asumsi klasik yang tidak terpenuhi yaitu uji normalitas. Menurut Montgomery & Peck (2012), tidak terpenuhinya salah satu asumsi tersebut menunjukkan adanya penciliran pada model. Oleh karena itu, dilakukan pendeteksian *outlier* menggunakan *scatter plot* dan metode *DfFITS*. Berdasarkan pendeteksian *outlier* dengan metode grafis *scatter plot* pada gambar 5, dapat dilihat bahwa data ke-7, 24, dan 33 mempunyai residual yang besar dan ketiga data tersebut terletak jauh dari pola kumpulan data. Dengan demikian dapat disimpulkan bahwa data ke-7, 24, dan 33 merupakan *outlier* dari data uji (Soemartini, 2007). Langkah selanjutnya yaitu dilakukan pendeteksian *outlier* dengan perhitungan statistis menggunakan metode *DfFITS*. Berdasarkan hasil perhitungan pada tabel 8, diperoleh hasil yaitu nilai *DfFITS* pada data ke-7, 14, 24, dan 33 memiliki nilai *DfFITS* yang lebih besar dari 0,9074. Dengan demikian dapat disimpulkan bahwa data ke-7, 14, 24, dan 33 merupakan *outlier* dari data pengamatan (Montgomery & Peck, 2012). Berdasarkan pendeteksian *outlier* menggunakan metode grafis *scatter plot* dan metode statistis *DfFITS*, dapat disimpulkan bahwa data pengamatan

mengandung *outlier* sehingga dilakukan pengestimasi parameter regresi menggunakan metode regresi *robust* estimasi-M pembobot *huber* dan *tukey bisquare*.

Berdasarkan penelitian yang telah dilakukan, diperoleh hasil yaitu model regresi *robust* estimasi-M pembobot *huber* menghasilkan nilai RSE sebesar 3,706 sedangkan model regresi *robust* estimasi-M pembobot *tukey bisquare* menghasilkan nilai RSE 3,767. Selain itu, model regresi *robust* estimasi-M pembobot *huber* memperoleh nilai *adjusted R-square* yang lebih tinggi yaitu sebesar 0,5315 dibandingkan dengan estimasi-M pembobot *tukey bisquare* sebesar 0,5199. Menurut Kurniawan & Yuniarto (2016), semakin dekat nilai *adjusted R-square* dengan 1 maka semakin baik tingkat kecocokan model regresi dengan data pengamatan. Berdasarkan hasil tersebut, maka dapat diketahui bahwa metode estimasi-M pembobot *huber* merupakan model regresi yang lebih baik karena memiliki tingkat kecocokan model regresi sebesar 0,5315 atau 53,15% variabel dependen Y dapat dijelaskan oleh variabel-variabel independen X_i pada model regresi estimasi-M pembobot *huber*.

Selain itu, pada penelitian ini juga dilakukan uji signifikansi parameter menggunakan uji F dan uji t. Berdasarkan pengujian signifikansi dengan uji F (uji simultan) pada kedua model regresi estimasi-M, diperoleh kesimpulan yang sama yaitu terdapat variabel independen pengaruh tingkat kemiskinan X_i yang secara simultan berpengaruh terhadap variabel dependen tingkat kemiskinan di Indonesia Y pada taraf signifikansi 5%. Untuk pengujian signifikansi parameter menggunakan uji t (uji parsial) pada kedua model regresi estimasi-M, diperoleh juga kesimpulan yang sama yaitu variabel independen X_3 yaitu persentase penduduk dengan sumber listrik PLN dan X_6 yaitu persentase tingkat pengangguran terbuka secara parsial mempengaruhi variabel dependen tingkat kemiskinan di Indonesia Y yaitu tingkat kemiskinan di Indonesia pada tahun 2021.

SIMPULAN

Berdasarkan uji F (simultan) yang telah dilakukan terhadap masing-masing model regresi *robust* estimasi-M pembobot *huber* dan *tukey bisquare*, dapat disimpulkan bahwa terdapat paling sedikit satu variabel independen X_i atau variabel yang mempengaruhi tingkat kemiskinan di Indonesia yang secara simultan memiliki pengaruh yang signifikan terhadap variabel dependen Y yaitu tingkat kemiskinan di Indonesia pada taraf signifikansi 5%.

Berdasarkan uji t (parsial) yang telah dilakukan terhadap masing-masing variabel independen pada model regresi *robust* estimasi-M pembobot *huber* dan *tukey bisquare*, dapat disimpulkan bahwa variabel independen persentase jumlah penduduk dengan sumber penerangan listrik PLN dan persentase tingkat pengangguran terbuka memiliki pengaruh yang signifikan terhadap tingkat kemiskinan di Indonesia pada tahun 2021.

Metode regresi *robust* estimasi-M pembobot *huber* merupakan metode yang lebih efektif digunakan dalam mengatasi *outlier* pada data tingkat kemiskinan di Indonesia pada tahun 2021 daripada pembobot *tukey bisquare* dengan perolehan nilai *adjusted R-square* yaitu 0,5315 dan nilai *Residual Standard Error* (RSE) 3,706 dengan persamaan yaitu

$$\hat{Y}_{m-huber} = 41,6115 + 0,0317X_1 + 0,0829X_2 - 0,3519X_3 - 0,1906X_4 - 0,0078X_5 + 0,2652X_6 + 0,0617X_7$$

UCAPAN TERIMA KASIH

Terima kasih kepada semua pihak yang telah membantu khususnya ibu Dr. Dra. Mathilda Susanti, M. Si selaku dosen pembimbing yang telah memberikan bimbingan dan pengarahan dalam penyusunan artikel ilmiah ini.

DAFTAR PUSTAKA

- Aflakhah, Z., Jajang, J., & Br. Sb., A. T. (2019). Kajian Metode Ordinary Least Square Dan Robust Estimasi M Pada Model Regresi Linier Sederhana Yang Memuat Outlier. *Jurnal Ilmiah Matematika Dan Pendidikan Matematika*, 11(1), 21.
- Amrin, A. (2016). Data Mining Dengan Regresi Linier Berganda Untuk Peramalan Tingkat Inflasi. *Jurnal Techno Nusa Mandiri*, XIII(1), 74–79.
- Azizah, R. J., & Wachidah, L. (2022). Regresi Robust Estimasi-M dengan Pembobot Huber dan Tukey Bisquare pada Data Tingkat Pengangguran di Indonesia Menurut Provinsi Tahun 2020. *Bandung Conference Series: Statistics*, 2(2), 18–26.
- BPS Indonesia. (2021). *Statistik Indonesia 2021*. Jakarta: Badan Pusat Statistik.
- Cahyandari, R., & Hisani, N. (2012). *Model Regresi Linier Berganda Menggunakan Penaksir Parameter Regresi Robust M-Estimator (Studi Kasus: Produksi Padi Di Provinsi Jawa Barat Tahun 2009)*. 6(1), 1–2.
- Candraningtyas, S., Safitri, D., & Ispriyanti, D. (2013). Regresi Robust MM-Estimator Untuk Penanganan Pencilan Pada Regresi Linier Berganda. *Jurnal Gaussian*, 2(4), 395–404.
- Chen, C. (2002). Robust Regression and Outlier Detection with the ROBUSTREG Procedure. *SAS Institute Inc., September*, 265–27.
- Kurniawan, R., & Yuniarto, B. (2016). *ANALISIS REGRESI: Dasar dan Penerapannya (Edisi Pertama)*. Jakarta: Kencana.
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to linear regression analysis 5th ed.* John Wiley & Sons.
- Pradewi, E. D., & Sudarno, S. (2012). Kajian Estimasi-M IRLS Menggunakan Fungsi Pembobot Huber dan Bisquare Tukey Pada Data Ketahanan Pangan di Jawa Tengah. *Media Statistika*, 5(1), 1-10.
- Santosa, P. B. & Ashari (2005). *Analisis Statistik Dengan Microsoft Excel Dan SPSS*. Yogyakarta: Andi.
- Siregar, Syofian. (2017). *Statistika Terapan Untuk Perguruan Tinggi*. Jakarta: Kencana.
- Soemartini. (2007). *Pencilan (Outlier)*. Jatinangor: Universitas Padjajaran.
- Tisniwati, B. (2012). Analisis Faktor-Faktor Yang Mempengaruhi Tingkat Kemiskinan Di Indonesia. *Jurnal Ekonomi Pembangunan*, 10(1), 33.
- Yu, C., & Yao, W. (2017). Robust linear regression: A review and comparison. *Communications in Statistics: Simulation and Computation*, 46(8), 6261–6282.