

Comparison of Estimate Method of Moment and Least Trimmed Squares in Models Robust Regression

Tri Wahyuni Nurmulyati, Dony Permana*, Nonong Amalita, Zamahsary Martha

Departemen Statistika, Universitas Negeri Padang, Padang, Indonesia

*Corresponding author: donypermana@fmipa.unp.ac.id

Submitted : 27 Mei 2024

Revised : 30 Mei 2024

Accepted : 31 Mei 2024

ABSTRACT

The poverty line represents the minimum income required for an individual to achieve a standard of living considered adequate in a given area. In 2022, the poverty line in West Sumatra Province was higher than the poverty line in Indonesia. A multiple linear regression analysis was conducted to identify the factors influencing the poverty line in West Sumatra Province. However, the observational data on the poverty line and its influencing factors contained outliers. Using the DFFITS test to check for outliers, it was found that the 8th and 18th observations contained outliers. Consequently, robust regression analysis was performed to address the data with outliers by comparing two estimators: MM estimation and LTS estimation. Based on the R^2 value, the best estimation was found to be the MM estimation, with significant factors being average net wages/salaries, unemployment rate (TPT), net enrollment rate (APM), and average years of schooling (AMH). If the average net wages/salaries, TPT, APM, and AMH increase, the poverty line in West Sumatra will rise. With an R^2 of 0.9582, the model explains 95.82% of the variation in the poverty line, while the remaining 4.18% is explained by other factors not included in the model.

Keywords: Ordinary Least Squares, Outlier, Method of Moment, Least Trimmed Squares.



This is an open access article under the Creative Commons 4.0 Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ©2022 by author and Universitas Negeri Padang.

I. PENDAHULUAN

Regresi adalah metode dalam ilmu statistika yang dipakai untuk menganalisis hubungan antara variabel respon dengan variabel penjelas. Regresi terdiri dari regresi liner sederhana dan regresi linear berganda. Regresi linier berganda merupakan metode yang digunakan untuk menganalisis hubungan antara satu variabel respon dengan dua atau lebih variabel penjelas. Menurut (Montgomery et al., 1992) metode estimasi yang baik digunakan untuk menentukan estimasi parameter regresi adalah *Ordinary Least Square* (OLS). OLS mempunyai asumsi-asumsi tertentu, asumsi-asumsi tersebut bisa saja tidak terpenuhi karena adanya masalah pada data. Masalah yang sering timbul pada data penelitian ialah karena adanya pencilan (*outlier*). Jika pada data penelitian mengandung *outlier*, maka akan menyebabkan asumsi-asumsi pada OLS tidak terpenuhi. Menurut (Draper & Smith, 1981) *outlier* merupakan suatu nilai data yang memiliki nilai yang tidak sama dengan nilai data lainnya. Atau dengan kata lain *outlier* adalah objek pengamatan observasi yang memiliki nilai yang eskrim atau jauh berbeda dari nilai yang lainnya. Dalam mengatasi *outlier* dapat menggunakan regresi *robust*. Menurut Irfagutami dkk. (2014), analisis regresi *robust* meruapaka salah satu metode dalam statistika yang bisa digunakan dalam memprediksi dengan melakukan estimasi. Regresi *robust* juga menghasilkan estimasi parameter yang kekar (kuat) terhadap *outlier*.

Pada berbagai kasus, tidak jarang ditemukannya data pengamatan yang mengandung data *outlier*. Pada studi kasus penelitian ini diambil studi kasus pada data garis kemiskinan di Provinsi Sumatera Barat Tahun 2022 yang mengandung *outlier*. Menurut Pratama (2014), kemiskinan adalah dimana keadaan seseorang yang tidak memiliki kemampuan, asset, dan akses untuk kebutuhan mereka di masa yang akan datang. Pada tahun 2022, penduduk miskin di Indonesia mencapai 26.363.270 jiwa, sedangkan di Provinsi Sumatera Barat jumlah penduduk miskin mencapai 343.820 jiwa. Menurut (Aprilia & Sembiring, 2021), garis kemiskinan menjadi tolak ukur kemiskinan. Garis kemiskinan adalah batas minimum pendapatan yang harus diperoleh oleh seseorang agar bisa dikatakan mendapatkan kehidupan yang layak sesuai standar di suatu daerah. Setiap tahun garis kemiskinan relatif akan naik sesuai dengan kondisi perekonomian daerah tersebut yang digambarkan dengan inflasi yang selalu positif. Bersumber dari data di Badan Pusat Statistika (BPS), dari tahun 2021-2022 jumlah penduduk miskin dan garis kemiskinan di Provinsi Sumatera Barat meningkat. Pada penelitian kali

ini, ditujukan untuk menganalisis faktor-faktor yang berpengaruh kuat, sedang dan tidak berpengaruh sama sekali terhadap garis kemiskinan di Sumatera Barat pada tahun 2022, yang mengandung *outlier*.

Sejak tahun 2014-2022 rata-rata garis kemiskinan di Provinsi Sumatera Barat meningkat sebesar 7,58% dan terjadi peningkatan tertinggi pada tahun 2021-2022 yaitu sebesar 12,88%. Sedangkan rata-rata garis kemiskinan di Indonesia meningkat sebesar 6,99% dan terjadi peningkatan tertinggi pada tahun 2014-2015 yaitu sebesar 10,40%. Garis kemiskinan di Provinsi Sumatera Barat pada tahun 2022 berbeda setiap kabupaten/kota. Garis kemiskinan tertinggi terletak di Kota Padang yaitu sebesar Rp. 634.581 rupiah/kapita/bulan, dan garis kemiskinan terendah terletak di Kepulauan Mentawai yaitu sebesar Rp. 386.152 rupiah/kapita/bulan. Pada masalah garis kemiskinan terdapat banyak faktor yang bisa berpengaruh, namun pada penelitian ini digunakan faktor-faktor seperti Produk Domestik Regional Bruto (PDRB), Rata-Rata Upah/Gaji Bersih, Tingkat Pengangguran Terbuka (TPT), Indeks Pembangunan Manusia (IPM), Angka Partisipasi Murni (APM), Jumlah Penduduk, dan Angka Melek Huruf (AMH).

Penelitian ini diharapkan dapat memberikan informasi dan menjadi referensi bagi pembaca dalam mengatasi kasus data penelitian yang mengandung *outlier*. Pada penelitian Akolo & Nadjamuddin (2022), dilakukan perbandingan estimasi LTS dan M pada data IPM di Pulau Sulawesi. Dari penelitian tersebut disimpulkan bahwa estimasi LTS lebih baik dalam mengatasi data outlier pada data IPM di Pulau Sulawesi. Penelitian Rositawati & Fitri (2022), dengan membandingkan estimasi S dan MM pada data inflasi pada bank. Hasil penelitian tersebut menyimpulkan bahwa estimasi MM lebih baik dari estimasi S dalam mengatasi data outlier pada data inflasi bank.

II. METODE PENELITIAN

A. Sumber Data dan Variabel Penelitian

Penelitian ini mengguankan data sekunder, yang bersumber dari *website* resmi BPS Provinsi Sumatera Barat tahun 2022. Variabel yang digunakan yaitu PDRB (X_1), rata-rata upag/gaji bersih (X_2), TPT (X_3), IPM (X_4), APM (X_5), jumlah penduduk (X_6), AMH (X_7), dan garis kemiskinan (Y).

B. Teknik Analisis Data

Dengan menggunakan bantuan *software Rstudio*, maka langkah-langkah analisis data dengan menggunakan regresi *robust* ialah sebagai berikut:

1. Melakukan eksplorasi dan statistika deskriptif pada data penelitian.
2. Pemodelan dan pengestimasi parameter menggunakan *Ordinary Least Square* (OLS).

Regresi linear berganda memiliki persamaan sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_k X_k + \varepsilon \quad (1)$$

Dalam model regresi terdapat parameter yang nilainya belum diketahui, untuk mengestimasi parameter regresi maka digunakan metode OLS, bentuk umum estimasi OLS yaitu:

$$\hat{\beta}_{OLS} = \arg \min \sum_{i=1}^n \varepsilon_i^2$$

Pada persamaan regresi linear berganda dilakukan estimasi menggunakan OLS, maka didapatkan persamaan dalam bentuk matriks penduga $\hat{\beta}_{OLS}$ ialah,

$$\bar{\beta} = (\underline{X}'\underline{X})^{-1}\underline{X}'\bar{Y} \quad (2)$$

Dengan $\bar{\beta}$ adalah vektor parameter regresi, \underline{X}' adalah matriks variabel penjelas yang di *transpose*, \underline{X} adalah matriks variabel penjelas, $(\underline{X}'\underline{X})^{-1}$ adalah perkalian matriks yang di *invers*, dan \bar{Y} adalah vektor variabel respon.

3. Melakukan uji asumsi klasik.
 - a. Uji Normalitas

Uji normalitas, untuk melihat berdistribusi normal atau tidaknya residual pada data observasi. Pengujian normalitas residual menggunakan uji *shapiro-wilk* dengan ($n < 30$).

Hipotesis (Rini & Faisal, 2015) :

H_0 : Berdistribusi normal

H_1 : Tidak berdistribusi normal

Statistik Uji:

$$W = \frac{b^2}{(n-1)s^2} \quad \text{dengan } b^2 = \sum_{i=1}^n a_{n-i+1} (X_{(n-i+1)} - X_{(i)}), \quad s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})}{n-1} \quad (3)$$

Dimana, $X_{(n-i+1)}$ adalah objek ke- $n-i+1$, $X_{(i)}$ adalah Objek ke- i , a_{n-i+1} adalah koefisien tes *shapiro wilk*, \bar{x} adalah rata-rata data, s^2 adalah varian data, dan n adalah jumlah objek data.

Kriteria uji : Tolak H_0 , jika $W_{hit} < W_{\alpha(n)}$ atau $p\text{-value} < \alpha$.

b. Uji Homoskedastisitas

Uji homoskedastisitas, untuk melihat apakah sama atau tidaknya varians dan residual nilai satu data observasi dengan data lainnya. Pada uji homoskedastisitas menggunakan uji *Breusch Pagan* (BP) dengan ($n < 50$).

Hipotesis :

H_0 : $var(S^2) = \sigma^2$ (homoskedastisitas)

H_1 : $var(S^2) \neq \sigma^2$ (heterokedastisitas)

Statistik Uji:

$$BP = n \times R^2 \quad (4)$$

Dimana, n adalah objek ke- i , R^2 adalah koefisien determinasi.

Kriteria uji : Tolak H_0 , jika $BP_{hitung} \geq X_{p-1}^2$ atau $p\text{-value} < \alpha$

c. Uji Non Autokorelasi

Uji non autokorelasi, untuk melihat nilai satu data dengan data lainnya memiliki korelasi pada residualnya. Uji non autokorelasi menggunakan uji *Durbin Watson* (DW) ($n < 100$).

Hipotesis:

H_0 : Tidak ada autokorelasi

H_1 : Ada autokorelasi

Statistik uji:

$$DW = \frac{\sum_{t=2}^{t=n} (e_t - e_{t-1})^2}{\sum_{t=2}^{t=n} e_t^2} \quad (5)$$

Kriteria uji : Terima H_0 , jika $dU < DW$, $4 - DW > dU$ atau $p\text{-value} < \alpha$. dU adalah batas atas pada tabel DW dan dL adalah batas bawah pada tabel DW

d. Uji Non Multikolinearitas

Uji non multikolinearitas digunakan untuk melihat antara variabel penjelas terdapat korelasi atau tidak. Uji non multikolinearitas memakai nilai *Variance Inflation Factors* (VIF)

Statistik Uji:

$$VIF_j = \frac{1}{1 - R_j^2} \quad (6)$$

$j : 1, 2, 3, \dots, k$, k adalah variabel penjelas, R_j^2 adalah koefisien determinasi variabel penjelas ke- j

Kriteria uji : Jika $VIF > 10$, maka ada multikolinearitas antara variabel penjelas

4. Mendeteksi *outlier* pada data

Pendeteksian *outlier* dilakukan menggunakan *Difference Fitted of FITS* (DFFITS). DFFITS adalah salah satu cara pendeteksian *outlier* yang melakukan pengukuran dengan tidak menghapus nilai data observasi karena akan mempengaruhi nilai-nilai data yang akan di prediksi (Montgomery et al., 2012). Persamaan DFFITS ialah:

$$DFFITS = t_i \sqrt{\frac{h_{ii}}{1-h_{ii}}} \quad \text{dengan } t_i = e_i \sqrt{\frac{n-k-1}{JKR(1-h_{ii})-e_i^2}}, \quad h_{ii} = x_i'(X'X)^{-1}x_i \quad (7)$$

e_i adalah residual ke- i , JKR adalah jumlah kuadrat dari residual, h_{ii} adalah nilai pengaruh dari elemen diagonal ke- i .

Suatu data observasi mengandung *outlier* jika $DFFITS > 2 \sqrt{\frac{p}{n}}$ dengan p adalah banyaknya parameter dan n adalah banyaknya observasi (Atamia et al., 2021).

5. Pemodelan dan pengestimasi parameter regresi *robust*

Regresi *robust* bertujuan untuk menganalisis data yang mengandung *outlier* dan asumsi tidak terpenuhi dengan hasil model estimasi yang kuat (*robust*) dari pengaruh *outlier*.

a. Estimasi MM

Estimasi MM merupakan salah satu estimasi regresi *robust*, menggabungkan estimasi *Maximum Likelihood* (M) dan estimasi *Scale* (S). Menurut Elveny dkk. (2021), bentuk umum estimasi MM yaitu:

$$\hat{\beta}_{MM} = \arg \min \sum_{i=1}^n \rho \left(\frac{\varepsilon_i}{\hat{\sigma}_s} \right)$$

Dengan menggunakan fungsi pembobot Tukey Bisquare yaitu

$$w_i = \frac{\psi(u_i)}{\varepsilon_i} \begin{cases} \left[1 - \left(\frac{u_i}{c}\right)^2\right]^2, & |u_i| \leq c \\ 0, & |u_i| > c \end{cases} \quad (8)$$

dimana $c = 4,685$

Pada bentuk umum estimasi MM, maka didapatkan model persamaan dalam bentuk matriks penduga $\hat{\beta}_{MM}$ ialah

$$\hat{\beta}_{MM} = (\underline{X}'\underline{W}\underline{X})^{-1}\underline{X}'\underline{W}\underline{Y}$$

Pendugaan parameter dengan estimasi MM:

1) Menentukan nilai residual $\varepsilon_i = y_i - \hat{y}_i$

2) Menentukan standar deviasi sisaan $\hat{\sigma}_s = \sqrt{\frac{n \sum_{i=1}^n (\varepsilon_i^2) - (\sum_{i=1}^n \varepsilon_i)^2}{n(n-1)}}$

3) Kemudian menghitung nilai $u_i = \frac{\varepsilon_i}{\hat{\sigma}_i}$

4) Menentukan nilai pembobot (w_i)

5) Menentukan OLS berbobot untuk menghasilkan penaksir kuadrat terkecil berbobot

$$\hat{\beta}_{MM} = (\underline{X}'\underline{W}\underline{X})^{-1}\underline{X}'\underline{W}\underline{Y} \quad (9)$$

6) Mengulangi tahap 2 – 5 hingga nilai $\hat{\beta}_{mm}$ konvergen, yaitu selisih β_j^{m+1} dengan β_j^m mendekati 0 atau kurang dari 10^{-4} , dimana m adalah banyaknya iterasi.

b. Estimasi LTS

Menurut Shodiqin dkk. (2018), estimasi LTS merupakan salah satu estimasi regresi *robust* yang meminimumkan banyaknya kuadrat residual h . Dimana h adalah banyaknya subset data yang memiliki kuadrat kecil dari fungsi obyektif. Menurut Rahman & Widodo (2018), persamaan estimasi LTS sebagai berikut:

$$\hat{\beta}_{LTS} = \arg \min \sum_{i=1}^h \varepsilon_i^2 \quad \text{dengan } h = \left\lfloor \frac{n}{2} \right\rfloor + \left\lfloor \frac{k+2}{2} \right\rfloor, \varepsilon_i = (Y_i - \sum_{j=0}^k \beta_j X_j)$$

Setelah dicari kuadrat *error* dan menurunkan persamaan terhadap $\hat{\beta}$ dan disama dengan kan nol, maka estimasi dari parameter $\hat{\beta}_{LTS}$ diperoleh hasil sebagai berikut:

$$\hat{\beta}_{LTS} = (\underline{X}\underline{X}')^{-1}\underline{X}\underline{Y} \quad (10)$$

Pendugaan parameter dengan estimasi LTS:

1) Menghitung estimasi koefisien dengan OLS

2) Menghitung n residual $\varepsilon_i^2 = (Y_i - \sum_{j=0}^k \beta_j X_j)^2$ yang sesuai dengan $(\hat{\beta}_0)$, setelah itu menentukan banyaknya $h_0 = \left\lfloor \frac{n}{2} \right\rfloor + \left\lfloor \frac{k+2}{2} \right\rfloor$ observasi dengan nilai ε_i^2 terkecil.

3) Menentukan $\hat{\beta}_{LTS}$.

4) Mengestimasi parameter $\hat{\beta}_{LTS}$ dari h observasi.

5) Menghitung n kuadrat residual $\varepsilon_i^2 = (Y_i - \sum_{j=0}^k \beta_j X_j)^2$ yang baru, menentukan h_{baru} pengamatan,

6) Mengestimasi $\hat{\beta}_{baru}$ yang sesuai dengan h_{baru} dengan melihat nilai ε_i^2 terkecil.

7) Mengulang tahap 4 – 7 untuk menghasilkan fungsi obyektif yang terkecil dan konvergen.

6. Membandingkan estimasi MM dan LTS dengan melihat koefisien determinasi (R^2).

Nilai R^2 dipakai untuk melihat seberapa baik variabel penjelas bisa menjelaskan variabel respon dalam suatu model, ($0 \leq R^2 \leq 1$). Menurut Gujarati dan Porter (2012:493) R^2 digunakan jika variabel penjelas hanya satu saja yaitu pada regresi linear sederhana. Sedangkan *adjusted* R^2 digunakan jika variabel penjelas lebih dari satu. Berikut ini merupakan rumus uji dari koefisien determinasi (Gujarati, 2009)

$$R^2 = \frac{[\sum(\hat{Y}_i - \bar{Y})(\hat{Y}_i - \bar{Y})]}{\sum(Y_i - \bar{Y})^2 \sum(\hat{Y}_i - \bar{Y})^2} \quad (11)$$

Semakin tinggi R^2 maka menunjukkan semakin baik dan bagus variabel penjelas dalam menjelaskan variabel respon dalam suatu model.

III. HASIL DAN PEMBAHASAN

A. Statistika Deskriptif

Sebelum dilakukannya pemodelan pengestimasi parameter regresi, dilakukan terlebih dahulu statistika deskriptif terhadap data penelitian secara umum. Statistika deskriptif pada Tabel 1.

Tabel 1. Statistika Deskriptif

Variabel	Satuan	Jumlah Observasi	Rata-rata	Maksimum	Minimum	St.Dev
<i>Y</i>	Rupiah/ Kapita/ Bulan	19	504.980	634.581	386.152	56.563,98
<i>X</i> ₁	Juta Rupiah	19	14.998.598	72.961.651	4.086.546	15.292.021
<i>X</i> ₂	Rupiah	19	2.560.424	3.339.847	1.987.340	462.700,5
<i>X</i> ₃	Persen	19	5,276	11,69	1,39	1,954022
<i>X</i> ₄	Persen	19	73,31	83,29	62,19	5,399309
<i>X</i> ₅	Persen	19	81,4	86,59	67,49	4,425431
<i>X</i> ₆	Jiwa	19	296.875	919.145	57.850	220.557
<i>X</i> ₇	Persen	19	99,65	100	98,46	0,3781178

Pada Tabel 1, dapat dilihat karakteristik setiap variabel yang berbeda. Nilai standar deviasi variabel *Y* sebesar 56563,98 lebih kecil dari nilai rata-rata variabel *Y* yaitu sebesar 504980, yang berarti bahwa garis kemiskinan setiap kabupaten/kota di Provinsi Sumatera Barat kurang bervariasi. Garis kemiskinan tertinggi pada tahun 2022 terletak pada Kota Padang yaitu sebesar Rp. 634.581 rupiah/kapita/bulan sedangkan garis kemiskinan terendah terletak di Kabupaten Kepulauan Mentawai yaitu sebesar Rp. 386.152 rupiah/kapita/bulan.

B. Pemodelan *Ordinary Least Squares* (OLS)

Pemodelan regresi linear berganda dilakukan dengan metode estimasi OLS, hasil awal estimasi OLS pada Tabel 2.

Tabel 2. Hasil Estimasi OLS

Koefisien	Estimate	Pr (> t)
<i>Intercept</i>	6,347000	0,0271
<i>X</i> ₁	- 0,0023	0,2681
<i>X</i> ₂	0,02243	0,3413
<i>X</i> ₃	9.962	0,2153
<i>X</i> ₄	5.571	0,1817
<i>X</i> ₅	4.572	0,2163
<i>X</i> ₆	0,1702	0,1955
<i>X</i> ₇	- 67.730	0,0247
<i>R-square</i>	0,7383	

Setelah melakukan reduksi pada satu persatu variabel yang tidak signifikan, maka didapatkan model terbaik untuk OLS yang signifikan yaitu sebagai berikut:

$$Y_{ij} = -99.870 + 7.880X_4 + 0,09151X_6$$

Dari model dapat dilihat bahwa variabel yang mempengaruhi garis kemiskinan adalah variabel *X*₄ yaitu IPM dan *X*₆ yaitu jumlah penduduk, dimana koefisien variabel *X*₄ dan *X*₆ bernilai positif. Hal ini menunjukkan apabila IPM dan jumlah penduduk meningkat maka garis kemiskinan di Sumatera Barat juga akan naik. Model estimasi OLS memiliki nilai koefisien determinasi (*R*²) sebesar 0.7383, artinya model mampu menjelaskan variasi penyebab garis kemiskinan sebesar 73,83% dan sisanya dijelaskan oleh faktor lain yang tidak disebutkan dalam model.

C. Uji Asumsi Klasik

1. Uji Normalitas

Hasil uji normalitas disajikan pada Tabel 3.

Tabel 3. Hasil Uji Normalitas

W	p-value
0,97936	0,9347

Berdasarkan Tabel 3, didapatkan nilai $p\text{-value} > \alpha$ yaitu $0,9347 > 0,05$, artinya terima H_0 . Maka disimpulkan residual data berdistribusi normal.

2. Uji Homoskedastisitas

Hasil uji homoskedastisitas disajikan pada Tabel 4.

Tabel 4. Hasil Uji Homoskedastisitas

BP	p-value
6,6342	0,4679

Berdasarkan Tabel 4, didapatkan nilai $p\text{-value} > \alpha$ yaitu $0,4679 > 0,05$, artinya terima H_0 . Maka dapat disimpulkan bahwa model estimasi OLS bersifat homoskedastisitas.

3. Uji Non Autokorelasi

Hasil uji non autokorelasi disajikan pada Tabel 5.

Tabel 5. Hasil Uji Non Autokorelasi

Durbin Watson	p-value
2,0303	0,3724

Berdasarkan Tabel 5, didapatkan nilai $p\text{-value} > \alpha$ yaitu $0,3724 > 0,05$, artinya terima H_0 . Maka disimpulkan model estimasi OLS tidak ada autokorelasi pada residual.

4. Uji Non Multikolinearitas

Hasil uji non multikolinearitas disajikan pada Tabel 6.

Tabel 6. Hasil Uji Non Multikolinearitas

Variabel	X_1	X_2	X_3	X_4	X_5	X_6	X_7
VIF	19,540322	2,341613	4,711526	9,569226	5,113155	15,942395	2,085436

Berdasarkan Tabel 6 diketahui bahwa variabel X_1 dan X_6 memiliki nilai VIF lebih besar dari 10. Maka dapat disimpulkan adanya multikolinearitas antara variabel penjelas, sehingga uji asumsi klasik tidak terpenuhi.

Uji asumsi klasik tidak terpenuhi, biasanya dikarenakan adanya *outlier* pada data penelitian. Maka dari itu data tersebut diatasi dengan menggunakan regresi *robust*.

D. Pendeteksian Outlier

Pendeteksian *outlier* dilakukan menggunakan uji *DFFITs*, berikut hasil pendeteksian *outlier* pada Tabel 7.

Tabel 7. Hasil Deteksi *Outlier*

Data ke-	DFFITs
8	1,66408884
18	1,21258364

Berdasarkan persamaan 7, nilai observasi dikatakan *outlier* jika $DFFITs > 2\sqrt{\frac{p}{n}}$ yaitu $DFFITs > 1,298$. Dari Tabel 7 diketahui adanya *outlier* pada data penelitian yaitu pada observasi ke 8 dan 18.

E. Regresi Robust

1. Estimasi *Method of Moment* (MM)

Hasil awal estimasi MM disajikan pada Tabel 8.

Tabel 8. Hasil Estimasi MM

Koefisien	Estimate	Pr (> t)
<i>Intercept</i>	6.521.000	0,0000349
X_1	- 0,0004719	0,63106
X_2	0,03134	0,0000229
X_3	14.760	0,00204
X_4	- 3.021	0,12095
X_5	11.880	0,0000138
X_6	- 0,06664	0,25659
X_7	- 69.300	0.0000230
<i>R-square</i>	0,9582	

Setelah melakukan reduksi pada satu persatu variabel yang tidak signifikan dan iterasi sudah mendekati 0, maka didapatkan model yang signifikan yaitu sebagai berikut:

$$Y_{ij} = 6.160.000 + 0,0255X_2 + 10.920X_3 + 9.508X_5 - 65.740X_7$$

Dari model dapat dilihat bahwa variabel yang mempengaruhi garis kemiskinan adalah variabel X_2 yaitu rata-rata upah/gaji bersih, variabel X_3 yaitu TPT, variabel X_5 yaitu APM dan variabel X_7 yaitu AMH, dimana koefisien variabel X_2, X_3, X_5, X_7 bernilai positif. Hal ini menunjukkan apabila rata-rata upah/gaji bersih, TPT, APM, dan AMH meningkat maka garis kemiskinan di Sumatera Barat akan naik. Model estimasi MM memiliki nilai koefisien determinasi (R^2) sebesar 0,9582, artinya model mampu menjelaskan variasi penyebab garis kemiskinan sebesar 95,52% dan sisanya dijelaskan oleh faktor lain yang tidak disebutkan dalam model.

2. Estimasi *Least Trimmed Squares* (LTS)

Hasil awal estimasi LTS disajikan pada Tabel 9.

Tabel 9. Hasil Estimasi LTS

Koefisien	Estimate	Pr (> t)
<i>Intercept</i>	6.254.049,1333	-
X_1	- 0,0001	0,9538
X_2	0,0302	0,02664
X_3	12.712,7505	0,0132
X_4	- 980,4374	0,6963
X_5	10.186,7759	0,0006638
X_6	- 0,0136	0,8619
X_7	- 66.684,6121	0,0006294
<i>R-square</i>	0,653561	

Setelah melakukan reduksi pada satu persatu variabel yang tidak signifikan dan iterasi sudah mendekati 0, maka didapatkan model yang signifikan yaitu sebagai berikut:

$$Y_{ij} = 6.254.049,1333 + 0,0302X_2 + 12.712,7505X_3 + 10.186,7759X_5 - 66.684,6121X_7$$

Dari model dapat dilihat variabel yang mempengaruhi garis kemiskinan adalah variabel X_2 yaitu rata-rata upah/gaji bersih, variabel X_3 yaitu TPT, variabel X_5 yaitu APM dan variabel X_7 yaitu AMH, dimana koefisien variabel X_2, X_3, X_5, X_7 bernilai positif. Hal ini menunjukkan apabila rata-rata upah/gaji bersih, TPT, APM, dan AMH meningkat maka garis kemiskinan di Sumatera Barat akan naik. Model estimasi LTS memiliki nilai R^2 sebesar 0,653561, artinya model mampu menjelaskan variasi penyebab garis kemiskinan sebesar 65,36% dan sisanya dijelaskan oleh faktor lain yang tidak disebutkan dalam model.

F. Koefisien Determinasi (R^2)

Nilai R^2 digunakan untuk menentukan estimasi terbaik dari 3 estimasi yaitu estimasi OLS, MM dan LTS. Hasil perbandingan R^2 disajikan pada Tabel 10.

Tabel 10. Perbandingan Estimasi OLS, MM dan LTS

Estimasi	R -square
OLS	73,83 %
MM	95,82 %
LTS	65,35 %

Berdasarkan Tabel 10 dapat dilihat perbandingan estimasi OLS, MM dan LTS untuk data garis kemiskinan di Sumatera Barat. Didapatkan estimasi MM lebih baik digunakan daripada estimasi LTS, dilihat dari R^2 estimasi MM yaitu 95,82% lebih besar dibandingkan dengan R^2 estimasi LTS yaitu 65,35%.

IV. KESIMPULAN

Pemodelan dengan estimasi terbaik ialah estimasi MM dan iterasi sudah mendekati 0 dengan koefisien determinasi sebesar 95,82%. Kemudian melakukan reduksi pada satu persatu variabel yang tidak maka model estimasi MM sebagai berikut.

$$Y_{ij} = 6.160.000 + 0,0255X_2 + 10.920X_3 + 9.508X_5 - 65.740X_7$$

Dari model dapat dilihat bahwa variabel yang mempengaruhi garis kemiskinan adalah variabel X_2 yaitu rata-rata upah/gaji bersih, variabel X_3 yaitu TPT, variabel X_5 yaitu APM dan variabel X_7 yaitu AMH, dimana koefisien variabel X_2, X_3, X_5, X_7 bernilai positif. Hal ini menunjukkan apabila rata-rata upah/gaji bersih, TPT, APM, dan AMH meningkat maka garis kemiskinan di Sumatera Barat akan naik. Model estimasi MM memiliki nilai koefisien determinasi (R^2) sebesar 0,9582, artinya model mampu menjelaskan variasi penyebab garis kemiskinan sebesar 95,52% dan sisanya dijelaskan oleh faktor lain yang tidak disebutkan dalam model. Penelitian selanjutnya disarankan untuk mencoba mengatasi data outlier dengan menggunakan estimasi lain.

DAFTAR PUSTAKA

- Akolo, I. R., & Nadjamuddin, A. (2022). Andriany, C. D., Susanti, Y., & Sugiyanto. (2021). Estimasi Parameter Regresi Robust Dengan Metode Estimasi Least Trimmed Squares (Lts) Pada Kematian Ibu Di Indonesia. Nasional Aplikasi Sains & Teknologi (SNAST), 9–14. *Jurnal Ilmiah Matematika, Sains Dan Teknologi*, 10(2), 211–221.
- Aprilia, K., & Sembiring, F. (2021). Analisis Garis Kemiskinan Makanan Menggunakan Metode Algoritma K-Means Clustering. *SISMATIK*, 1–10.
- Atamia, N. A., Susanti Yuliana, & Handajani, S. S. (2021). Perbandingan Analisis Regresi Robust Estimasi-S dan Estimasi-M dengan Pembobot Huber dalam Mengatasi Outlier. *PRISMA, Prosiding Seminar Nasional Matematika*, 4.
- Draper, N. R., & Smith, H. (1981). *Applied Regression Analysis* (R. A. Bradley, J. S. Hunter, D. G. Kendall, & G. S. Watson, Eds.; 2nd ed.). John Wiley & Sons, Inc .
- Elveny, M., Syah, R., & Al-khowarizmi. (2021). Modelling Under Uncertainty Business Competitive With Robust Estimation Measurement Model. *Turkish Journal of Computer and Mathematics Education*, 12(3), 4410–4417.
- Irfagutami, N. P. N., Srinadi, I. G. A. M., & Sumarjaya, I. W. (2014). Perbandingan Regresi Robust Penduga Mm Dengan Metode Random Sample Consensus Dalam Menangani Pencilan. *E-Jurnal Matematika*, 3(2), 45.
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (1992). *Introduction to Linear Regression Analysis* (V. Barnett, R. A. Bradley, N. I. Fisher, J. S. Hunter, J. B. Kadane, D. G. Kendall, A. F. M. Smith, S. M. Stigler, J. L. Teugels, & G. S. Watson, Eds.; 2nd ed.). John Wiley & Sons, Inc.

- Pratama, Y. C. (2014). Analisis Faktor-Faktor Yang Mempengaruhi Kemiskinan di Indonesia. *Jurnal Bisnis Dan Manajemen*, 4(2).
- Rahman, M. B., & Widodo, E. (2018). Perbandingan Metode Regresi Robust Estimasi Least Trimmed Square , Estimasi Scale , dan Estimasi Method Of Moment. *Prosiding Seminar Nasional Matematika*.
- Rini, D. S., & Faisal, F. (2015). *Perbandingan Power of Test dari Uji Normalitas Metode Bayesian, Uji Shapiro-Wilk, Uji Cramer-von Mises, dan Uji Anderson-Darling* (Vol. 11, Issue 2).
- Rositawati, A. F. D., & Fitri, H. Z. (2022). Pengendalian Inflasi Melalui Hasil Pemodelan Faktor yang Berpengaruh Terhadap Inflasi Menggunakan Metode Regresi Robust Time Series Estimasi-S dan Estimasi-MM. *Govstat*, 1(1), 13–28.
- Shodiqin, A., Aini, A. N., & Rubowo, M. R. (2018). Perbandingan Dua Metode Regresi Robsut Yakni Metode Least Trimmed Squares (LTS) dengan Metode Estimator MM (Estimasi-MM) (Studi Kasus DAta Ujian Tulis Masuk Terhadap Hasil IPK Mahasiswa UPGRIS). *Jurnal Ilmiah Teknosains*, 4(1).