

Perbandingan Pembobot *Welsch* dan *Tukey Bisquare* pada Regresi *Robust S-estimator*

Fifi Nurhafifah^{1*}, Khoirin Nisa², Nusyirwan³, Rizki Agung Wibowo⁴

^{1,2,3}Jurusan Matematika, Fakultas MIPA, Universitas Lampung, Bandar Lampung, Indonesia

⁴Jurusan Manajemen, Fakultas Ekonomi, Universitas Malahayati, Bandar Lampung, Indonesia

*corresponding author: Fifi Nurhafifah99@gmail.com

Received June 25, 2024; Received in revised July 02, 2024; Accepted July 03, 2024

Abstrak. 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*; *tukey bisquare*; *welsch*; *tukey bisquare*

Abstract. 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

Keyword: robust regression; *S-estimator*; *Tukey Bisquare*; *Welsch*

PENDAHULUAN

Analisis regresi adalah sebuah analisis statistika yang bertujuan untuk mengestimasi model hubungan antara variabel terikat (y) dengan satu atau lebih variabel bebas (x). Salah satu metode yang sering digunakan untuk mengestimasi parameter model regresi adalah Metode Kuadrat Terkecil (MKT) (James et al., 2023; Flatt & Jacobs, 2019). 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 (Herawati, 2010). 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 (Rahayu et al., 2023).

MKT sangat peka terhadap adanya penyimpangan asumsi. Jika terdapat asumsi yang dilanggar, maka MKT menjadi tidak efisien (Karunasingha, 2022; Barron, 2020). 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 (Mahapatra, 2020). 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 (Hodson, 2022).

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 (Filzmoser & Nordhausen, 2021). 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* (Lakshmi & Sajesh, 2023). *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 (Amato et al., 2021; Saha et al., 2020).

Perbandingan antara pembobot *Welsch* dan *Tukey Bisquare* pada *S-estimator* regresi yang kuat dieksplorasi dalam berbagai penelitian. Nurcahyani et al. (2022) menunjukkan bahwa dalam studi tentang data indeks perkembangan manusia di Papua, fungsi pembobotan *Welsch* mengungguli *Tukey Bisquare* dalam hal fungsi pembobotan yang disesuaikan dan MKT. Demikian pula, Azizah & Wachidah (2022) berfokus pada data tingkat pengangguran di Indonesia dan menemukan bahwa pembobotan *Tukey Bisquare* menghasilkan model regresi yang lebih kuat dengan R-kuadrat yang disesuaikan lebih tinggi dan nilai Kesalahan Standar Residual yang lebih rendah dibandingkan dengan bobot Huber. Selain itu, Nurbaroqah et al. (2022) menyoroiti efektivitas *Tukey Bisquare* dalam regresi yang kuat untuk data indeks pembangunan manusia di Jawa Tengah. Temuan ini secara kolektif menunjukkan bahwa timbangan *Welsch* dan *Tukey Bisquare* memainkan peran penting dalam meningkatkan ketahanan dan akurasi penaksir S dalam model regresi yang kuat. Namun belum ada yang meneliti tentang perbandingan pembobot *Welsch* dan *Tukey Bisquare* dengan berbagai ukuran sampel. Oleh karena itu, tujuan dari penelitian ini adalah membandingkan kinerja dua fungsi pembobot, yaitu *Welsch* dan *Tukey Bisquare*, pada metode *S-estimator* dalam regresi *robust*. Selanjutnya dengan berbagai ukuran sampel, akan dilakukan perbandingan keefektifan ditinjau dari nilai MSE (*Mean Square Error*) dan bias agar diperoleh pembobot terbaik.

METODOLOGI PENELITIAN

Data Penelitian

Data yang digunakan dalam penelitian ini merupakan data simulasi berukuran 30, 60, 100 dan 200 dengan replikasi atau 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%.

Regresi Robust metode S Estimator

Regresi robust harus mampu menggambarkan struktur yang sesuai dengan sebagian besar data dan mengidentifikasi titik data yang menyimpang dari asumsi. Untuk menentukan kinerja metode *robust* dalam arti teoritis digunakan sifat efisiensi, *breakdown point*, dan *high leverage point* (Fitrianto & Xin, 2022). 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 (De Menezes et al., 2021).

Metode *S-estimator* merupakan penduga parameter yang berdasarkan perkiraan "skala". Metode ini memiliki cara yang sama seperti MKT yang bertujuan untuk meminimalkan ragam galat (Leski, 2023). Metode *S-estimator* didefinisikan sebagai (1):

$$\beta = (X^T_{((k+1) \times n)} W_{(n \times n)} X_{(n \times (k+1))})^{-1} (X^T_{((k+1) \times n)} W_{(n \times n)} Y_{(n \times 1)}). \quad (1)$$

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

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 β
 $\beta^{m+1} = (X^T W^m X)^{-1} (X^T W^m Y)$.
- g. Mengulangi langkah b hingga f sampai diperoleh nilai β 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 akan bergantung dengan fungsi penduga awal pada metode kuadrat terkecil, sedangkan iterasi kedua dan seterusnya bergantung dengan fungsi pada iterasi sebelumnya. Iterasi akan terus berlanjut hingga konvergen atau selisih β^{m+1} dengan β^m mendekati atau sama dengan nol.

Tahapan Penelitian

Topik yang dibahas terkait regresi *robust S-estimator* dengan pembobotan *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 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: $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$.
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*.

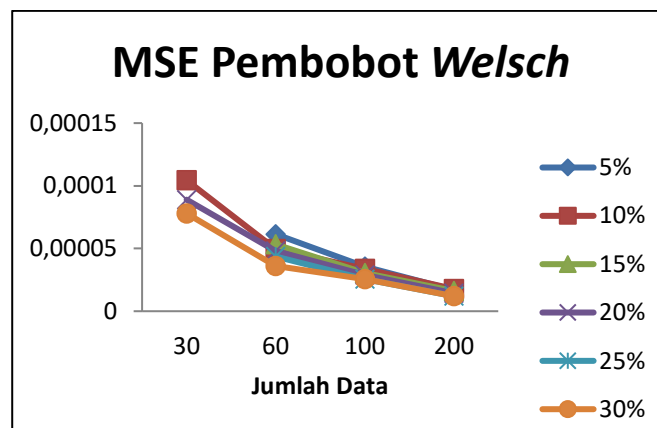
HASIL DAN PEMBAHASAN

Data yang digunakan merupakan data simulasi. Data simulasi dibangkitkan menggunakan perangkat lunak R-4.0.3 dengan jumlah data 30, 60, 100 dan 200 dan dilakukan pengulangan sebanyak seribu kali. Metode *S-estimator* merupakan salah satu metode *robust* yang resisten terhadap adanya pengaruh pencilan. Pada data disimulasikan mengandung pencilan sebesar 5%, 10%, 15%, 20%, 25% dan 30% sedangkan untuk ukuran data $n=30$ pencilan disimulasikan 10%, 20% dan 30%.

Berdasarkan nilai dugaan koefisien regresi dari 1000 ulangan, diperoleh nilai MSE dan bias menggunakan metode MKT dan *S-estimator* dengan masing-masing pembobot dan berbagai ukuran sampel. Nilai MSE dan bias digunakan sebagai tolok ukur keefektifan penggunaan metode *S-estimator* ketika diterapkan pada data yang mengandung pencilan.

Perbandingan nilai MSE dan Bias untuk Seluruh Jumlah Data

Berdasarkan nilai dugaan koefisien regresi dengan 1000 ulangan yang diterapkan pada metode *S-estimator* pembobot *Welsch* dan *Tukey Bisquare* nilai MSE dan bias yang dihasilkan dari setiap n data dan pencilan menunjukkan hasil yang baik untuk presentase pencilan 30%. Meski memiliki fungsi objektif yang berbeda untuk setiap pembobot, akan tetapi berdasarkan nilai MSE dan bias menunjukkan bahwa tidak terdapat perbedaan yang cukup signifikan serta $\hat{\beta}$ yang diperoleh tidak memiliki perbedaan yang signifikan. Nilai MSE secara keseluruhan untuk masing-masing pembobot disajikan pada Gambar 1. dan Gambar 2.



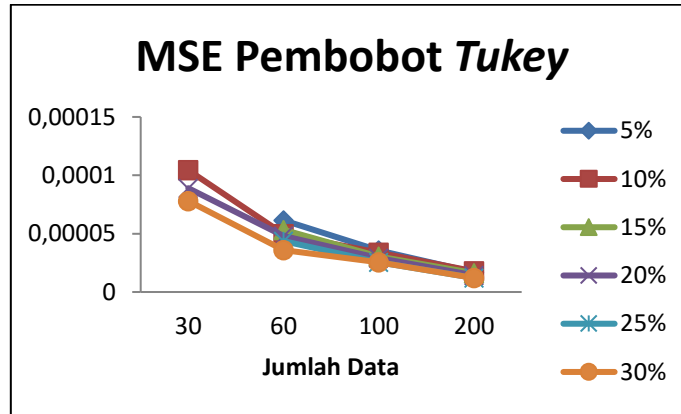
Gambar 1. Grafik MSE untuk Pembobot *Welsch*

Gambar 1. menampilkan grafik *Mean Square Error* (MSE) dari pembobot *Welsch* dalam regresi *robust* untuk berbagai ukuran data dan tingkat kontaminasi pencilan. Sumbu horizontal menunjukkan jumlah data yang digunakan dalam simulasi, yaitu 30, 60, 100, dan 200, sedangkan sumbu vertikal menunjukkan nilai MSE dengan skala mulai dari 0 hingga 0,00015. Garis dan simbol yang berbeda pada grafik mewakili berbagai tingkat kontaminasi pencilan dalam data, yakni 5%, 10%, 15%, 20%, 25%, dan 30%.

Secara umum, nilai MSE cenderung menurun seiring dengan peningkatan jumlah data, menunjukkan bahwa dengan lebih banyak data, model regresi *robust* dengan pembobot *Welsch* menjadi lebih akurat. Pada jumlah data yang lebih sedikit (30 dan 60), terdapat perbedaan yang lebih besar antara nilai MSE untuk berbagai tingkat kontaminasi pencilan. Namun, seiring bertambahnya jumlah data (100 dan 200), nilai MSE untuk berbagai tingkat kontaminasi pencilan menjadi lebih serupa dan lebih rendah secara keseluruhan.

Pada jumlah data yang lebih kecil (30), kontaminasi pencilan yang lebih tinggi (10% dan 30%) cenderung menghasilkan MSE yang sedikit lebih tinggi dibandingkan dengan kontaminasi pencilan yang lebih rendah (5%). Namun, perbedaan ini menjadi kurang signifikan dengan bertambahnya jumlah data, menunjukkan bahwa pembobot *Welsch* cukup efektif dalam menangani pencilan, terutama dengan jumlah data yang lebih besar. Kesimpulannya, grafik tersebut menunjukkan bahwa pembobot *Welsch* dalam regresi *robust* bekerja dengan baik untuk berbagai tingkat kontaminasi pencilan, terutama ketika jumlah data cukup besar (100 atau lebih). Penurunan nilai MSE seiring dengan peningkatan

jumlah data mengindikasikan bahwa metode ini menjadi lebih stabil dan akurat dengan dataset yang lebih besar, dan efektif dalam mengurangi pengaruh pencilan pada estimasi parameter.

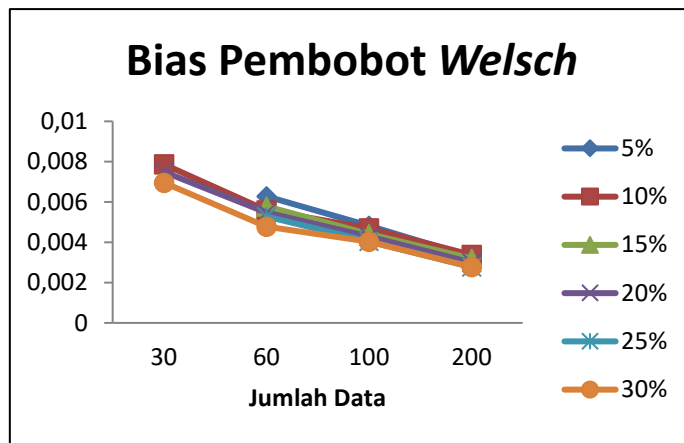


Gambar 2. Grafik MSE untuk Pembobot *Tukey Bisquare*

Berdasarkan Gambar 2. Secara keseluruhan, nilai MSE cenderung menurun seiring dengan peningkatan jumlah data, menunjukkan bahwa dengan lebih banyak data, model regresi robust dengan pembobot Tukey menjadi lebih akurat. Pada jumlah data yang lebih sedikit (30 dan 60), terdapat perbedaan yang lebih besar antara nilai MSE untuk berbagai tingkat kontaminasi pencilan. Namun, seiring bertambahnya jumlah data (100 dan 200), nilai MSE untuk berbagai tingkat kontaminasi pencilan menjadi lebih serupa dan lebih rendah secara keseluruhan.

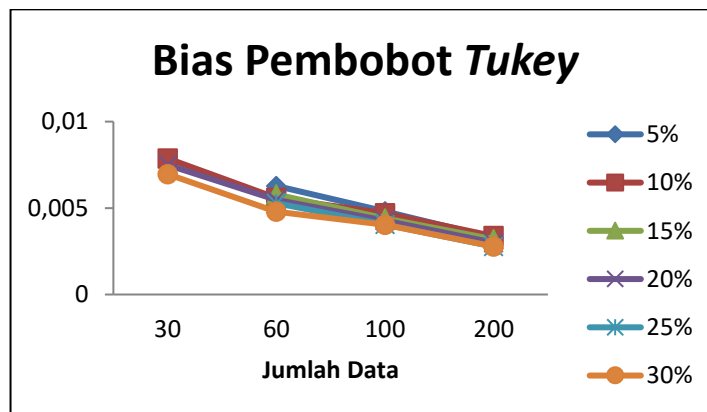
Pada jumlah data yang lebih kecil (30), kontaminasi pencilan yang lebih tinggi (10% dan 30%) cenderung menghasilkan MSE yang sedikit lebih tinggi dibandingkan dengan kontaminasi pencilan yang lebih rendah (5%). Namun, perbedaan ini menjadi kurang signifikan dengan bertambahnya jumlah data, menunjukkan bahwa pembobot Tukey cukup efektif dalam menangani pencilan, terutama dengan jumlah data yang lebih besar. Kesimpulannya, grafik tersebut menunjukkan bahwa pembobot Tukey dalam regresi robust bekerja dengan baik untuk berbagai tingkat kontaminasi pencilan, terutama ketika jumlah data cukup besar (100 atau lebih). Penurunan nilai MSE seiring dengan peningkatan jumlah data mengindikasikan bahwa metode ini menjadi lebih stabil dan akurat dengan dataset yang lebih besar, dan efektif dalam mengurangi pengaruh pencilan pada estimasi parameter.

Berdasarkan Gambar 1. dan Gambar 2. Dapat disimpulkan bahwa secara keseluruhan berdasarkan jumlah data dan masing-masing pencilan pada setiap pembobot, semakin besar pencilan dan jumlah data yang disimulasikan maka akan semakin kecil nilai MSE yang diperoleh. Selanjutnya nilai bias secara keseluruhan untuk masing-masing pembobot disajikan pada Gambar 3. dan Gambar 4.



Gambar 3. Grafik Bias untuk Pembobot *Welsch*

Gambar 3. Menunjukkan bahwa nilai Bias cenderung menurun seiring dengan peningkatan jumlah data, menunjukkan bahwa dengan lebih banyak data, model regresi robust dengan pembobot *Welsch* menjadi lebih akurat dalam estimasi parameter. Pada jumlah data yang lebih sedikit (30 dan 60), terdapat perbedaan yang lebih besar antara nilai Bias untuk berbagai tingkat kontaminasi pencilan. Namun, seiring bertambahnya jumlah data (100 dan 200), nilai Bias untuk berbagai tingkat kontaminasi pencilan menjadi lebih serupa dan lebih rendah secara keseluruhan.



Gambar 4. Grafik Bias untuk Pembobot *Tukey Bisquare*

Sedangkan Gambar 4. menunjukkan bahwa nilai bias juga cenderung menurun seiring dengan peningkatan jumlah data, menunjukkan bahwa dengan lebih banyak data, model regresi robust dengan pembobot *Tukey* menjadi lebih akurat dalam estimasi parameter. Pada jumlah data yang lebih sedikit (30 dan 60), terdapat perbedaan yang lebih besar antara nilai Bias untuk berbagai tingkat kontaminasi pencilan. Namun, seiring bertambahnya jumlah data (100 dan 200), nilai Bias untuk berbagai tingkat kontaminasi pencilan menjadi lebih serupa dan lebih rendah secara keseluruhan. Berdasarkan Gambar 3. dan Gambar 4. Jika dipandang secara keseluruhan berdasarkan jumlah data dan masing-masing pencilan pada setiap pembobot, semakin besar pencilan dan jumlah data yang disimulasikan maka akan semakin kecil nilai bias yang diperoleh.

Pengujian Hipotesis Terhadap Rata-rata Penduga

Uji-t sampel independen merupakan salah satu analisis statistik parametrik. Uji ini digunakan untuk melihat ada atau tidaknya perbedaan rata-rata antar dua kelompok data yang saling bebas atau tidak berpasangan (Dankel & Loenneke, 2021).

Adapun syarat yang harus dipenuhi sebelum dilakukan uji-t independen adalah sebagai berikut:

1. Normalitas.
Masing-masing kelompok data harus berdistribusi normal.
2. Independen.
Masing-masing kelompok data harus independen dari semua pengamatan, atau tidak ada hubungan antar kelompok.
3. Homogenitas.
Setiap kelompok harus memiliki ragam yang homogen.
Rumus uji-t [20] didefinisikan sebagai berikut:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-1} \tag{9}$$

dengan:

- \bar{x}_1 : nilai rata-rata kelompok 1
- \bar{x}_2 : nilai rata-rata kelompok 2
- n : jumlah data
- S_p : $\sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}}$.

Berikut adalah uji hipotesis dari rata-rata penduga yang disajikan pada Tabel 1.

Tabel 1. Uji Hipotesis Rata-rata Penduga untuk n=30 dan Pencilan 10%

Koefisien	Uji Levene untuk Persamaan Varian		Uji-t untuk Persamaan Rata-rata		
	F	Sig	t	dB	Sig.(2 - tailed)
Sama dengan Varian yang di Asumsikan	.000	1.000	.000	1998	1.000
Tidak Sama dengan Varian yang di Asumsikan			.000	1998.000	1.000

i. Uji Hipotesis

$$H_0: \text{Nilai rata-rata } \hat{\beta}_{welsch} = \hat{\beta}_{tukey}$$

$$H_1: \text{Nilai rata-rata } \hat{\beta}_{welsch} \neq \hat{\beta}_{tukey}$$

ii. Taraf Signifikansi $\alpha = 0,05$.

iii. Statistik Uji

$$t_{hitung} = \frac{\bar{x}_1 - \bar{x}_2}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

iv. Keputusan Uji

Tolak H_0 ketika $|t_{hitung}| > t_{(\frac{\alpha}{2}, n_1+n_2-2)}$ atau $Sig. (2 - tailed) < \alpha$.

v. Kesimpulan

Pada Tabel 1 nilai $sig.(2-tailed) > \alpha$ maka tidak cukup bukti untuk menolak H_0 sehingga dapat ditarik kesimpulan bahwa nilai rata-rata $\hat{\beta}_{welsch}$ sama dengan nilai rata-rata $\hat{\beta}_{tukey}$.

Berdasarkan Tabel 1. uji rata-rata penduga dapat ditarik kesimpulan bahwa untuk $n=30$ dan pencilan sebesar 10%, 20% dan 30% dengan seribu ulangan menunjukkan bahwa nilai rata-rata $\hat{\beta}_{welsch}$ sama dengan nilai rata-rata $\hat{\beta}_{tukey}$. Sehingga dapat ditarik suatu kesimpulan bahwa pembobot *Welsch* dan *Tukey Bisquare* memiliki nilai yang sama baiknya jika digunakan pada regresi *robust S-estimator*.

Berdasarkan hasil simulasi dan analisis yang dilakukan, terdapat beberapa temuan penting mengenai efektivitas metode *S-estimator* dengan pembobot *Welsch* dan *Tukey Bisquare* dalam menangani data yang mengandung pencilan. Pertama, kedua pembobot menunjukkan kemampuan yang baik dalam mengurangi pengaruh pencilan pada estimasi parameter regresi. Meskipun terdapat perbedaan nilai MSE dan bias pada jumlah data dan tingkat kontaminasi pencilan yang lebih kecil, perbedaan tersebut menjadi tidak signifikan dengan bertambahnya jumlah data. Kedua, nilai MSE dan bias cenderung menurun seiring dengan peningkatan jumlah data, menunjukkan bahwa metode regresi robust dengan *S-estimator* menjadi lebih akurat dan stabil pada dataset yang lebih besar. Ketiga, hasil uji hipotesis menunjukkan bahwa tidak ada perbedaan yang signifikan antara rata-rata nilai β' yang diperoleh menggunakan pembobot *Welsch* dan *Tukey Bisquare*. Kedua pembobot memiliki kinerja yang sama baiknya dalam estimasi parameter regresi pada data yang mengandung pencilan.

Faktor-faktor yang menyebabkan hasil tersebut antara lain adalah jumlah data yang cukup, yang membuat estimasi parameter menjadi lebih stabil karena distribusi data mendekati distribusi yang sebenarnya, sehingga metode robust dapat bekerja lebih efektif. Selain itu, keefektifan metode *S-estimator* dalam mengurangi pengaruh pencilan juga berperan penting. Meskipun pembobot *Welsch* dan *Tukey Bisquare* memiliki fungsi objektif yang berbeda, keduanya mampu mengurangi pengaruh pencilan secara efektif, terutama pada jumlah data yang besar.

Penelitian ini memiliki beberapa kelebihan, seperti pengulangan simulasi yang banyak, memberikan hasil yang lebih akurat dan representatif, serta variasi ukuran data dan tingkat pencilan, memungkinkan analisis yang komprehensif mengenai kinerja metode yang diuji. Namun, ada juga beberapa kekurangan, seperti fokus yang hanya pada metode *S-estimator*, sehingga tidak membandingkan dengan metode robust lainnya, serta penggunaan data simulasi yang mungkin tidak sepenuhnya mencerminkan karakteristik data nyata dalam berbagai aplikasi.

Penelitian ini sejalan dengan penelitian sebelumnya yang menunjukkan efektivitas metode robust dalam menangani pencilan. Sebagai contoh, penelitian oleh Khan et al. (2021) dan She (2022) juga menunjukkan bahwa metode *S-estimator* memiliki ketahanan yang baik terhadap pencilan. Penelitian ini menambah bukti bahwa pembobot *Welsch* dan *Tukey Bisquare* dalam metode *S-estimator* dapat digunakan secara efektif pada dataset yang mengandung pencilan.

Implikasi penelitian ini mencakup aplikasi praktis dalam analisis data, memberikan panduan bagi praktisi dalam memilih metode robust yang efektif untuk analisis data yang mengandung pencilan. Kedua pembobot yang diuji dapat digunakan secara efektif dalam berbagai aplikasi. Selain itu, penelitian ini dapat menjadi dasar untuk pengembangan lebih lanjut dari metode robust lainnya, dengan fokus pada peningkatan ketahanan terhadap pencilan dan efisiensi dalam estimasi parameter. Penelitian ini juga menambah literatur mengenai kinerja metode *S-estimator* dengan pembobot *Welsch* dan *Tukey Bisquare*, terutama dalam konteks simulasi data dengan berbagai ukuran dan tingkat kontaminasi

pencilan. Dengan demikian, penelitian ini tidak hanya mengkonfirmasi temuan sebelumnya tetapi juga memperluas pemahaman mengenai kinerja metode robust dalam berbagai kondisi data yang mengandung pencilan.

KESIMPULAN

Berdasarkan hasil penelitian, maka kesimpulan dari penelitian ini adalah pembobot *Welsch* dan *Tukey Bisquare* memiliki fungsi objektif yang berbeda. Meskipun demikian, nilai parameter $\hat{\beta}$ yang diperoleh menunjukkan nilai yang tidak jauh berbeda. Kemudian, 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*. Untuk penelitian selanjutnya, disarankan agar metode *S-estimator* dibandingkan dengan metode robust lainnya dan diuji pada data nyata untuk memperkuat temuan ini. Penelitian lebih lanjut juga bisa mengeksplorasi pembobot tambahan atau kombinasi pembobot untuk meningkatkan ketahanan terhadap pencilan.

DAFTAR PUSTAKA

- Amato, U., Antoniadis, A., De Feis, I., & Gijbels, I. (2021). Penalised robust estimators for sparse and high-dimensional linear models. *Statistical Methods & Applications*, 30(1), 1-48. <https://doi.org/10.1007/s10260-020-00511-z>
- 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. In *Bandung Conference Series: Statistics*, 2(2), 18-26. <https://doi.org/10.29313/bcss.v2i2.2648>
- Barron, A. R. (2020). Predicted squared error: a criterion for automatic model selection. In *Self-organizing methods in modeling* (pp. 87-103). CRC Press.
- Dankel, S. J., & Loenneke, J. P. (2021). Effect sizes for paired data should use the change score variability rather than the pre-test variability. *The Journal of Strength & Conditioning Research*, 35(6), 1773-1778. <https://doi.org/10.1519/JSC.0000000000002946>
- De Menezes, D. Q. F., Prata, D. M., Secchi, A. R., & Pinto, J. C. (2021). A review on robust M-estimators for regression analysis. *Computers & Chemical Engineering*, 147, 107254. <https://doi.org/10.1016/j.compchemeng.2021.107254>
- Filzmoser, P., & Nordhausen, K. (2021). Robust linear regression for high-dimensional data: An overview. *Wiley Interdisciplinary Reviews: Computational Statistics*, 13(4), e1524. <https://doi.org/10.1002/wics.1524>
- Fitrianto, A., & Xin, S. H. (2022). Comparisons between robust regression approaches in the presence of outliers and high leverage points. *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, 16(1), 243-252. <https://doi.org/10.30598/barekengvol16iss1pp241-250>
- Flatt, C., & Jacobs, R. L. (2019). Principle assumptions of regression analysis: Testing, techniques, and statistical reporting of imperfect data sets. *Advances in Developing Human Resources*, 21(4), 484-502. <https://doi.org/10.1177/1523422319869915>
- Herawati, N. (2010). Analisis Ketegaran Regresi Robust Terhadap Letak Pencilan: Studi Perbandingan. *Bulletin of Mathematics*, 3(1).
- Hodson, T. O. (2022). Root mean square error (RMSE) or mean absolute error (MAE): When to use them or not. *Geoscientific Model Development Discussions*, 2022, 1-10. <https://doi.org/10.5194/gmd-15-5481-2022>
- James, G., Witten, D., Hastie, T., Tibshirani, R., & Taylor, J. (2023). Linear regression. In *An introduction to statistical learning: With applications in python* (pp. 69-134). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-38747-0_3

- Karunasingha, D. S. K. (2022). Root mean square error or mean absolute error? Use their ratio as well. *Information Sciences*, 585, 609-629. <https://doi.org/10.1016/j.ins.2021.11.036>
- Khan, D. M., Ali, M., Ahmad, Z., Manzoor, S., & Hussain, S. (2021). A New Efficient Redescending M-Estimator for Robust Fitting of Linear Regression Models in the Presence of Outliers. *Mathematical Problems in Engineering*, 2021(1), 3090537. <https://doi.org/10.1155/2021/3090537>
- Lakshmi, R., & Sajesh, T. A. (2023). Empirical Study on Robust Regression Estimators and Their Performance. *Reliability: Theory & Applications*, 18(2 (73)), 466-478.
- Leski, J. M. (2023). Fuzzy double-ordered c-regression models based on fuzzy S-estimators. *Fuzzy Sets and Systems*, 465, 108531. <https://doi.org/10.1016/j.fss.2023.108531>
- Mahapatra, A. P. K., Nanda, A., Mohapatra, B. B., Padhy, A. K., & Padhy, I. (2020). Concept of outlier study: The management of outlier handling with significance in inclusive education setting. *Asian Research Journal of Mathematics*, 16(10), 7-25. <https://doi.org/10.9734/arjom/2020/v16i1030228>
- Nurbaroqah, A., Pratikno, B., & Supriyanto, S. (2022). Pendekatan Regresi Robust Dengan Fungsi Pembobot Bisquare Tukey Pada Estimasi-M Dan Estimasi-S. *Jurnal Ilmiah Matematika Dan Pendidikan Matematika (JMP)*, 14(1), 19-30. <https://doi.org/10.20884/1.jmp.2022.14.1.5669>
- Nurcahyani, N., Pratikno, B., & Supriyanto, S. (2022). PENERAPAN REGRESI ROBUST ESTIMASI-S DENGAN FUNGSI PEMBOBOT TUKEY BISQUARE DAN WELSCH PADA KASUS DATA IPM DI PROVINSI PAPUA TAHUN 2021. *Jurnal Ilmiah Matematika dan Pendidikan Matematika (JMP)*, 14(2), 133-144. <https://doi.org/10.20884/1.jmp.2022.14.2.6668>
- Rahayu, D. A., Nursholihah, U. F., Suryaputra, G., & Surono, S. (2023). Comparasion of the m, mm and s estimator in robust regression analysis on indonesian literacy index data 2018. *EKSAKTA: Journal of Sciences and Data Analysis*, 11-22. <https://doi.org/10.20885/EKSAKTA.vol4.iss1.art2>
- Saha, A., Singh, K. N., Gurung, B., Rathod, S., & Yeasin, M. (2020). Robust Regression in the Presence of Leverage: An Application to the Baseball Data.
- She, Y., Wang, Z., & Shen, J. (2022). Gaining outlier resistance with progressive quantiles: Fast algorithms and theoretical studies. *Journal of the American Statistical Association*, 117(539), 1282-1295. <https://doi.org/10.1080/01621459.2020.1850460>