

PERBANDINGAN METODE MCD-BOOTSTRAP DAN LAD- BOOTSTRAP DALAM MENGATASI PENGARUH PENCILAN PADA ANALISIS REGRESI LINEAR BERGANDA

Ni Luh Putu Ratna Kumalasari^{1§}, Ni Luh Putu Suciptawati^{2§},
Made Susilawati³

¹Jurusan Matematika, FMIPA – Universitas Udayana [Email: ratnakumalasari22@gmail.com]

²Jurusan Matematika, FMIPA – Universitas Udayana [Email: putusuciptawati@yahoo.co.id]

³Jurusan Matematika, FMIPA – Universitas Udayana [Email: susilawati.made@gmail.com]

[§]Corresponding Author

ABSTRACT

Outliers are observations that are far away from other observations. Outlier can be interfered with the process of data analysis which influence the regression parameters estimation. Methods that are able to deal with outliers are Minimum Covariance Determinant and Least Absolute Deviation methods. However, if both methods are applied with small sample the validity of both methods is being questioned. This research applies bootstrap to MCD and LAD methods to small sample. Resampling using 500, 750, and 1000 with confidence interval of 95% and 99% shows that both methods produce an unbiased estimators at 10%, 15%, and 20% outliers. The confidence interval of MCD-Bootstrap method is shorter than LAD-Bootstrap method. Both are, MCD-Bootstrap method is a better than LAD-Bootstrap method.

Keywords: Bootstrap, Least Absolute Deviation (LAD), Minimum Covariance Determinant (MCD), Outliers.

1. PENDAHULUAN

Analisis regresi merupakan metode statistika yang bertujuan untuk menganalisis hubungan atau pengaruh antara satu atau lebih peubah prediktor terhadap peubah respons. Analisis regresi linear dibagi menjadi dua jenis yaitu analisis regresi linear sederhana dan analisis regresi linear berganda. Analisis regresi linear berganda adalah analisis regresi yang melibatkan dua atau lebih peubah prediktor dengan satu peubah respons. Model regresi linear berganda dapat dituliskan sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \dots + \beta_k X_{ik} + \varepsilon_i \quad (1)$$

Pada analisis regresi, metode estimasi yang digunakan untuk menduga parameter serta memiliki sifat tidak bias adalah Metode Kuadrat Terkecil (MKT) atau *Ordinary Least of Square* (OLS). Penggunaan MKT memerlukan asumsi-asumsi tertentu yang harus dipenuhi antara lain adalah galat (sisaan) harus memenuhi asumsi kenormalan. Namun dalam berbagai kasus

penelitian, tidak jarang ditemukan pelanggaran terhadap asumsi kenormalan. Salah satu penyebabnya adalah adanya pencilan (*outlier*) pada data amatan.

Pencilan merupakan data amatan yang menyimpang jauh lebih besar atau jauh lebih kecil dari data amatan lainnya. Oleh karena itu, diperlukan alternatif dalam pemilihan metode estimasi yang tepat tanpa menghilangkan data pencilan. Metode yang mampu menghasilkan penduga parameter yang *robust* atau kekar terhadap adanya pencilan pada analisis regresi linear berganda adalah *Minimum Covariance Determinant* (MCD) dan *Least Absolute Deviation* (LAD). Akan tetapi, jika kedua metode ini dihadapkan pada data yang berjumlah sedikit maka diragukan. Oleh sebab itu, perlu dilakukannya penelitian terhadap pengaplikasian *bootstrap* atau *resampling* pada metode MCD dan metode LAD dengan data yang berjumlah sedikit.

Prinsip *bootstrap* adalah menduga parameter untuk masing-masing jumlah sampel yang

diperoleh dengan mengambil sampel berukuran n dari nilai-nilai data asli. Sampel tersebut merupakan sampel acak dengan pengembalian, artinya bahwa beberapa nilai sampel asli akan berulang dan beberapa diantaranya tidak akan terambil sama sekali (Sprenst, 1989). Metode *bootstrap* yang digunakan pada regresi linear adalah *bootstrap residual*. Metode *bootstrap residual* adalah metode yang meresampling sisaannya.

Berdasarkan pemaparan di atas, penulis tertarik melakukan perbandingan terhadap metode MCD-*Bootstrap* dan LAD-*Bootstrap* dalam mengatasi pengaruh pencilan pada analisis regresi linear berganda. Hal ini dapat ditunjukkan dengan membandingkan bias penduga parameter dan lebar selang yang dihasilkan oleh metode MCD-*Bootstrap* dan LAD-*Bootstrap*.

Dengan demikian, tujuan dari penelitian ini adalah untuk mengetahui bias yang dihasilkan oleh metode MCD-*Bootstrap* dan LAD-*Bootstrap*.

2. TINJAUAN PUSTAKA

Metode Kuadrat Terkecil (MKT)

Metode kuadrat terkecil adalah salah satu metode yang sering digunakan untuk menduga parameter dalam analisis regresi linear berganda. MKT meminimumkan jumlah kuadrat sisaan (galat), sehingga nilai parameter mendekati nilai sesungguhnya. Menurut Teorema Gauss-Markov, setiap pendugaan MKT yang asumsinya terpenuhi akan bersifat BLUE (*Best Linear Unbiased Estimator*).

Pencilan (*Outlier*)

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

Minimum Covariance Determinant (MCD)

MCD merupakan metode penduga parameter dengan meminimumkan determinan matriks kovarians. Prinsip MCD adalah dengan menggunakan vektor rata-rata dan matriks kovarians yang didapat dari penduga MCD untuk menentukan bobot dari setiap data, sehingga akan didapat penduga parameter model MCD. Metode ini bertujuan untuk mencari subsampel H yang berukuran h dari keseluruhan n amatan dengan $h \leq n$ yang matriks kovariansnya memiliki determinan terkecil di antara semua kombinasi kemungkinan data (Hubert & Debruyne, 2009).

Adapun nilai h dapat ditentukan dengan rumus sebagai berikut:

$$h = \frac{n+p+1}{2} \quad (2)$$

Untuk nilai vektor rata-rata \mathbf{V}_{MCD} dan matriks kovarians S_{MCD} diberikan:

$$\mathbf{V}_{MCD} = \frac{1}{h} \sum_{i \in H} x_i \quad (3)$$

$$S_{MCD} = \frac{1}{h} \sum_{i \in H} [x_i - \mathbf{V}_{MCD}] [x_i - \mathbf{V}_{MCD}]^T \quad (4)$$

Dari persamaan (3) dapat diketahui nilai determinan matriks kovarians. Jika $\det(S_{MCD}) \neq 0$, maka kemudian hitung jarak Mahalanobis. Menghitung jarak Mahalanobis yang kekar diperoleh dengan rumus:

$$RD(x_i) = \sqrt{(x_i - \mathbf{V}_{MCD})^T S_{MCD}^{-1} (x_i - \mathbf{V}_{MCD})} \quad (5)$$

Adapun langkah-langkah penduga MCD dengan Fast-MCD adalah mengambil himpunan bagian dari matriks X secara acak, misalkan himpunan bagian tersebut H_1 dengan jumlah elemen sebanyak h . Hitung vektor rata-rata \mathbf{V}_{MCD} dan matriks kovarians S_{MCD} pada H_1 dengan persamaan (3) dan (4).

Misalkan $\mathbf{V}_{MCD} = \mathbf{V}_1$ dan $S_{MCD} = S_1$, serta hitung $\det(S_1)$. Apabila $\det(S_1) = 0$ maka berhenti, tetapi jika $\det(S_1) \neq 0$, maka hitung *Robust Distance (RD)* menggunakan persamaan (5). Nilai RD yang diperoleh akan diurutkan dari nilai terkecil. Demikian seterusnya hingga mencapai $\det(S_{i+1}) = \det(S_1)$. Ulangi langkah diatas dengan mengambil himpunan H selanjutnya. Pilih himpunan H yang memiliki nilai determinan matriks kovarians terkecil, selanjutnya mencari nilai \mathbf{V}_{MCD} dan S_{MCD} .

Berdasarkan anggota h tersebut, selanjutnya data diboboti:

$$W_i = \begin{cases} 1, & \text{jika } (x_i - V_{MCD})^T S_{MCD}^{-1} (x_i - V_{MCD}) \leq \chi_{p;1-\alpha}^2 \\ 0, & \text{lainnya} \end{cases}$$

Pembobot W_i dapat membentuk matriks W_{MCD} berukuran $n \times n$, sebagai berikut:

$$W_{MCD} = \begin{bmatrix} w_{11} & w_{12} & w_{13} & \dots & w_{1n} \\ w_{21} & w_{22} & w_{23} & \dots & w_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & w_{n3} & \dots & w_{nn} \end{bmatrix}$$

Sehingga diperoleh penduga MCD dengan persamaan:

$$\hat{\beta}_{MCD} = (X^T W_{MCD} X)^{-1} (X^T W_{MCD} Y) \quad (6)$$

Least Absolute Deviation (LAD)

Least Absolute Deviation (LAD) merupakan metode penduga parameter untuk menangani masalah galat yang tidak berdistribusi normal. Metode LAD dikenal dengan berbagai istilah yaitu Minimum Absolute Deviation dan Least Absolute Value (LAV). Penaksiran LAD untuk mendapatkan β adalah meminimumkan jumlah nilai mutlak dari galat yaitu:

$$\begin{aligned} \hat{\beta} &= \min \sum_{i=1}^n |\varepsilon_i| \\ &= \min \sum_{i=1}^n |Y_i - X_i^t \beta_{LAD}| \end{aligned} \quad (7)$$

Metode penaksiran LAD dapat dimodelkan sebagai berikut:

$$\hat{\beta} = (X^T W_{LAD} X)^{-1} (X^T W_{LAD} Y) \quad (8)$$

dengan W adalah matriks diagonal dengan elemen diagonalnya w_{ii} .

$$w_{ii} = \begin{cases} \frac{1}{|\varepsilon_i|}, & \text{jika } |\varepsilon_i| \neq 0 \\ 1, & \text{jika } |\varepsilon_i| = 0 \end{cases}$$

dengan ε_i adalah galat dari nilai awal yang telah diperoleh dari metode kuadrat terkecil (El-Salam, 2013).

Bootstrap Residual

Metode *bootstrap* akan menghasilkan dugaan koefisien regresi ($\hat{\beta}^*$) sebanyak jumlah ulangan (B). Langkah-langkah *bootstrap residual* adalah menentukan nilai \hat{Y} dari model analisis regresi yang telah ditetapkan. Selanjutnya menghitung nilai *residual* yaitu

$e = Y - \hat{Y}$, diperoleh $e = \{e_1, e_2, \dots, e_n\}$. Mengambil sampel *bootstrap* berukuran n dari e_1, e_2, \dots, e_n secara acak dengan pengembalian, diperoleh sampel *bootstrap* $e^* = \{e_1^*, e_2^*, \dots, e_n^*\}$. Menghitung nilai *bootstrap* untuk Y^* dengan menambahkan e^* , sehingga menghasilkan $\hat{\beta}^*$. Mengulangi proses diatas sesuai dengan ulangan (B) yang diinginkan (Sungkono, 2013).

2. METODE PENELITIAN

Data yang digunakan dalam penelitian ini adalah data simulasi yang diperoleh dengan membangkitkan data berdistribusi normal. Data dibangkitkan dengan bantuan program R i386 3.2.0. Data yang dibangkitkan dimodelkan dalam regresi linear berganda. Banyaknya sampel yang digunakan adalah 40 amatan. Persentase pencilan yang akan digunakan yaitu 10%, 15%, dan 20%.

Langkah pertama dari penelitian ini adalah membangkitkan nilai sisaan (e) yang berdistribusi normal sebanyak 40 dengan rata-rata 0 dan varians 1. Selanjutnya, membangkitkan peubah prediktor sebanyak 40 amatan $X_1 \sim N(60,5)$ dan $X_2 \sim N(80,3)$. Menentukan nilai peubah respons dengan nilai parameter yang telah ditentukan yakni $\beta_0 = 3$, dan $\beta_1, \beta_2 = 1$. Pencilan dibangkitkan pada peubah respons sebesar 10%, 15%, dan 20%. Kemudian, menghitung nilai peubah respons yang telah ditambahkan pencilan. Langkah berikutnya, dilakukan uji Anderson-Darling untuk melihat kenormalan data, pendeteksian pencilan dengan plot *robust distance* dan dianalisis dengan menggunakan Metode Kuadrat Terkecil (MKT).

Langkah selanjutnya, data dianalisis dengan metode MCD-*Bootstrap* yaitu menduga nilai parameter dengan mencari nilai vektor rata-rata dan matriks kovarians MCD. Selanjutnya untuk menemukan β_0, β_1 , dan β_2 , sisaan yang diperoleh dari metode MCD kemudian diresampling dengan metode *bootstrap residual* sebanyak 500, 750, dan 1000 kali ulangan dengan selang kepercayaan 95% dan 99%. Selanjutnya, dilakukan analisis dengan metode

LAD-*Bootstrap*. Dilakukan *resampling* pada sisaan yang diperoleh dari metode LAD-*Bootstrap*. *Resampling* dilakukan sebanyak 500, 750, dan 1000 kali dengan selang kepercayaan 95% dan 99%. Langkah terakhir, yakni membandingkan hasil yang diperoleh dari kedua metode tersebut.

4. HASIL DAN PEMBAHASAN

4.1 Pemeriksaan Kenormalan Data

Pemeriksaan kenormalan data menggunakan uji kenormalan Anderson-Darling dilakukan dengan menggunakan program Minitab 17. Hipotesis sebagai berikut:

H_0 : sisaan menyebar secara normal

H_1 : sisaan tidak menyebar secara normal

Sisaan dikatakan menyebar secara normal (terima H_0) apabila nilai *p-value* lebih besar dari alfa (α). Demikian sebaliknya, sisaan dikatakan tidak menyebar secara normal (tolak H_0) apabila nilai *p-value* lebih kecil dari alfa (α). Pada penelitian uji kenormalan data ini, alfa (α) yang digunakan sebesar 0,05. Nilai *p-value* pada uji Anderson-darling disajikan pada table berikut:

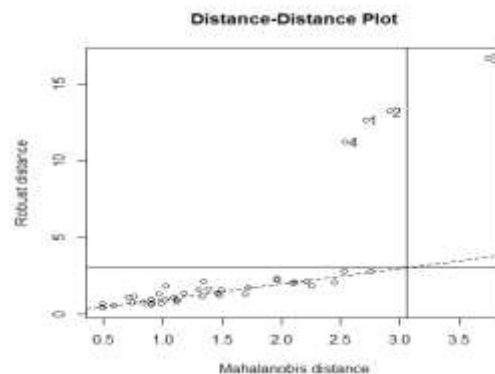
Tabel 1. Uji Kenormalan Data

Persentase Pencilan	P-value	Keterangan
Data awal (tanpa pencilan)	0,551	Terima H_0 (normal)
10%	< 0,005	Tolak H_0 (tidak normal)
15%	< 0,005	Tolak H_0 (tidak normal)
20%	< 0,005	Tolak H_0 (tidak normal)

Berdasarkan hasil uji kenormalan data dengan uji Anderson-Darling, dapat dilihat bahwa data awal (tanpa pencilan) memiliki nilai *p-value* yang lebih besar dari alfa (α) sebesar 0,05. Hal ini berarti data awal (tanpa pencilan) merupakan data yang menyebar secara normal. Namun nilai *p-value* pada data yang memiliki persentase pencilan sebesar 10%, 15%, dan 20% menunjukkan bahwa data yang mengandung pencilan merupakan data tidak menyebar secara normal.

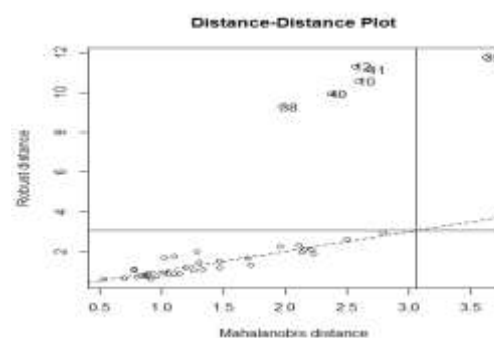
4.2 Pemeriksaan Pencilan

Pemeriksaan pencilan dilakukan menggunakan *Robust Distance* (RD) dengan bantuan program Ri386 3.2.0. Pencilan dapat diklasifikasikan dengan plot *outlier* yang berguna untuk membedakan data amatan. Plot yang dihasilkan akan membagi amatan atas 4 kuadran berdasarkan jenis pencilan yang terdeteksi. Pada pemeriksaan pencilan dengan *Robust Distance* diperoleh plot sebagai berikut:



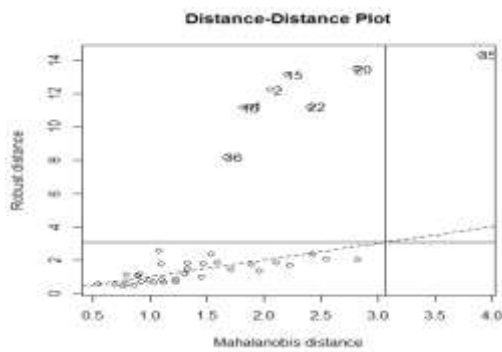
Gambar 1. Plot RD pada Pencilan 10%

Amatan yang terdeteksi pencilan antara lain adalah amatan ke-1, 2, 3 dan 4. Terlihat bahwa terdapat sebanyak 1 amatan yang berada pada kuadran I yang merupakan *bad leverage* dan terdapat sebanyak 3 amatan yang berada pada kuadran II yang merupakan *outlier orthogonal*. Berikutnya pencilan dengan persentase 15%.



Gambar 2. Plot RD pada Pencilan 15%

Amatan yg terdeteksi pencilan adalah amatan ke-10, 11, 12, 38, 39 dan 40. Pada kuadran I terdapat 1 amatan dan pada kuadran II terdapat 5 amatan. Hal ini berarti 1 amatan termasuk *bad leverage* yaitu amatan ke-39 dan 5 amatan termasuk *outlier orthogonal*. Selanjutnya plot RD dengan pencilan 20% :



Gambar 3. Plot RD pada Pencilan 20%

Pada Gambar 3 terlihat pada pencilan 20% terdapat 8 amatan yang terdeteksi pencilan. Amatan tersebut adalah amatan ke-1, 2, 15, 16, 20, 22, 35 dan 36. Amatan jenis pencilan *bad leverage* sebanyak 1 amatan yang berada pada kuadran I. Sedangkan pada kuadran II terdapat sebanyak 7 amatan yang termasuk *outlier orthogonal*.

4.3 Analisis Data dengan Metode Kuadrat Terkecil (MKT)

Metode awal yang digunakan untuk menduga nilai parameter pada penelitian ini adalah MKT. Data awal tanpa pencilan dan data dengan pencilan, selanjutnya akan dianalisis dengan menggunakan MKT. Selang kepercayaan yang digunakan adalah 95% dan 99%. Diperoleh nilai parameter dengan MKT, sebagai berikut:

Tabel 2. Analisis dengan MKT

Persentase Pencilan	Parameter	Estimasi	Selang Kepercayaan	Keterangan
Data awal tanpa pencilan	β_0	28,88798	10,98113-39,75710	Tidak Bias
	β_1	0,95186	0,82283-1,08088	Tidak Bias
	β_2	0,83413	0,57599-1,09226	Tidak Bias
10%	β_0	40,53211	12,45217-43,19877	Tidak Bias
	β_1	1,47372	0,62026-1,34140	Bias
	β_2	1,20976	0,71817-1,02539	Bias
15%	β_0	52,93148	22,96745-48,09781	Bias
	β_1	1,39347	0,58459-1,29198	Bias
	β_2	1,25263	0,44370-1,13315	Bias
20%	β_0	56,76652	20,86231-52,98622	Bias
	β_1	1,49318	0,73625-1,20425	Bias
	β_2	1,38419	0,55607-1,15338	Bias

Tabel 3. Analisis dengan MKT

Persentase Pencilan	Parameter	Estimasi	Selang Kepercayaan	Keterangan
Data awal tanpa pencilan	β_0	28,88798	16,94413-42,98210	Tidak Bias
	β_1	0,95186	0,77894-1,12477	Tidak Bias
	β_2	0,83413	0,48818-1,18007	Tidak Bias
10%	β_0	40,53211	14,67215-49,01827	Tidak Bias
	β_1	1,47372	0,60905-1,42241	Bias
	β_2	1,20976	0,51058-1,15708	Bias
15%	β_0	52,93149	29,12641-48,67122	Bias
	β_1	1,39347	0,71209-1,23925	Bias
	β_2	1,25263	0,87063-1,00648	Bias
20%	β_0	56,76652	20,89761-50,97762	Bias
	β_1	1,49318	0,78471-1,37269	Bias
	β_2	1,38419	0,64305-1,20583	Bias

Berdasarkan Tabel 2 dan Tabel 3 diperoleh bahwa data awal tanpa pencilan untuk selang kepercayaan 95% dan 99% adalah tidak bias. Hal ini dapat disimpulkan karena nilai penduga parameter β_0 , β_1 dan β_2 yang dihasilkan oleh analisis metode MKT berada antara lebar selang kepercayaan yang sudah diberikan. Hasil yang sama diperlihatkan saat persentase pencilan 10% untuk parameter β_0 yang bersifat tidak bias. Namun untuk parameter β_1 dan β_2 pada persentase 10% tidak berada pada selang kepercayaan 95% dan 99%.

Dilain pihak, nilai penduga parameter pada data dengan persentase pencilan sebesar 15% dan 20% untuk selang kepercayaan 95% dan 99% tidak berada pada lebar selang yang sudah diberikan. Hal ini berarti, penduga parameter β_0 , β_1 , dan β_2 yang dihasilkan mengalami bias. Karena data dengan pencilan mengalami bias pada analisis MKT, maka akan dilanjutkan menganalisis dengan metode *MCD-Bootstrap* dan *LAD-Bootstrap*.

4.4 Analisis Data dengan Metode Minimum Covariance Determinant (MCD-Bootstrap)

Langkah pertama dari analisis MCD adalah menduga parameter regresi dengan mencari nilai vektor rata-rata dan matriks kovarians MCD. Selanjutnya untuk menemukan β_0 , β_1 , dan β_2 pada data yang mengandung pencilan sebesar 10%, 15%, dan 20%, sisaan yang diperoleh dari metode MCD kemudian diresampling dengan metode *bootstrap residual*.

Resampling dilakukan sebanyak 500, 750, dan 1000 kali ulangan. Selang kepercayaan yang digunakan adalah selang kepercayaan sebesar 95% dan 99%.

Tabel 4. Analisis Metode MCD-Bootstrap dengan B = 500, SK= 95%

Persentase pencilan	Parameter	Estimasi	Selang Kepercayaan 95%		Keterangan
			Batas Bawah	Batas Atas	
10%	β_0	27,36254	14,58490	39,77970	Tidak bias
	β_1	1,01700	0,93227	1,19616	Tidak bias
	β_2	0,68452	0,53585	0,94587	Tidak bias
15%	β_0	27,40943	12,42169	41,94291	Tidak bias
	β_1	1,01269	0,72193	1,10650	Tidak bias
	β_2	0,68725	0,50882	0,97289	Tidak bias
20%	β_0	27,19416	12,66719	41,69741	Tidak bias
	β_1	1,01218	0,92321	1,10522	Tidak bias
	β_2	0,69107	0,51780	0,96391	Tidak bias

Tabel 5. Analisis Metode MCD-Bootstrap dengan B = 500, SK= 99%

Persentase pencilan	Parameter	Estimasi	Selang Kepercayaan 99%		Keterangan
			Batas Bawah	Batas Atas	
10%	β_0	27,21720	10,78117	43,58343	Tidak bias
	β_1	1,01755	0,90371	1,22472	Tidak bias
	β_2	0,68592	0,49269	0,88903	Tidak bias
15%	β_0	27,05674	9,53557	45,82903	Tidak bias
	β_1	1,01676	0,88328	1,14515	Tidak bias
	β_2	0,68878	0,43932	1,04239	Tidak bias
20%	β_0	26,85559	11,52642	46,83818	Tidak bias
	β_1	1,01410	0,89732	1,13110	Tidak bias
	β_2	0,89375	0,45006	0,93165	Tidak bias

Tabel 6. Analisis Metode MCD-Bootstrap dengan B = 750, SK= 95%

Persentase pencilan	Parameter	Estimasi	Selang Kepercayaan 95%		Keterangan
			Batas Bawah	Batas Atas	
10%	β_0	27,10400	14,56568	39,79892	Tidak bias
	β_1	1,11461	0,93351	1,39491	Tidak bias
	β_2	0,58961	0,53588	0,84584	Tidak bias
15%	β_0	26,92436	11,20819	43,15641	Tidak bias
	β_1	1,01307	0,91736	1,11106	Tidak bias
	β_2	0,89313	0,50129	0,98043	Tidak bias
20%	β_0	26,83828	12,26572	42,09888	Tidak bias
	β_1	1,11758	0,91530	1,21313	Tidak bias
	β_2	0,99128	0,50995	1,07176	Tidak bias

Tabel 7. Analisis Metode MCD-Bootstrap dengan B = 750, SK= 99%

Persentase pencilan	Parameter	Estimasi	Selang Kepercayaan 99%		Keterangan
			Batas Bawah	Batas Atas	
10%	β_0	27,08484	9,69528	44,66932	Tidak bias
	β_1	0,91526	0,91279	1,21563	Tidak bias
	β_2	0,57731	0,47826	0,90346	Tidak bias
15%	β_0	27,24372	10,47284	44,89176	Tidak bias
	β_1	1,01512	0,78724	1,14118	Tidak bias
	β_2	0,78758	0,43919	0,94252	Tidak bias
20%	β_0	27,17589	8,22410	46,14050	Tidak bias
	β_1	1,01446	0,89477	1,23365	Tidak bias
	β_2	0,68948	0,46398	0,91773	Tidak bias

Tabel 8. Analisis Metode MCD-Bootstrap dengan B = 1000, SK= 95%

Persentase pencilan	Parameter	Estimasi	Selang Kepercayaan 95%		Keterangan
			Batas Bawah	Batas Atas	
10%	β_0	26,73376	14,01490	40,34970	Tidak bias
	β_1	1,01624	0,83014	1,19829	Tidak bias
	β_2	0,79298	0,53081	0,85090	Tidak bias
15%	β_0	27,02557	11,58506	42,77954	Tidak bias
	β_1	1,01295	0,81412	1,11430	Tidak bias
	β_2	0,59194	0,49840	1,08332	Tidak bias
20%	β_0	27,49510	12,28192	42,08268	Tidak bias
	β_1	1,01303	0,92174	1,12668	Tidak bias
	β_2	0,68654	0,50903	0,87268	Tidak bias

Tabel 9. Analisis Metode MCD-Bootstrap dengan B = 1000, SK= 99%

Persentase pencilan	Parameter	Estimasi	Selang Kepercayaan 99%		Keterangan
			Batas Bawah	Batas Atas	
10%	β_0	26,97037	9,76541	44,59919	Tidak bias
	β_1	0,91582	0,80968	1,11874	Tidak bias
	β_2	0,69038	0,47906	0,90266	Tidak bias
15%	β_0	27,02121	6,35750	48,00710	Tidak bias
	β_1	1,01370	0,88290	1,14552	Tidak bias
	β_2	0,69139	0,43357	0,94814	Tidak bias
20%	β_0	27,10866	7,88706	46,47753	Tidak bias
	β_1	1,01241	0,89734	1,33108	Tidak bias
	β_2	0,89192	0,45456	0,92715	Tidak bias

Sifat tidak bias pada selang kepercayaan 95% dan 99% terjadi apabila nilai parameter regresi yaitu β_0 , β_1 , dan β_2 berada pada selang kepercayaan yang sudah diberikan. Pada Tabel 4, 5, 6, 7, 8, dan 9 dapat dilihat bahwa nilai parameter dari data dengan pencilan

sebesar 10%, 15%, dan 20% berada pada selang kepercayaan. Hal ini berarti data bersifat tidak bias pada selang kepercayaan 95%, 99% dengan *resampling* 500, 750, dan 1000 kali ulangan.

4.5 Analisis Data dengan Metode Least Absolute Deviation (LAD)-Bootstrap

Langkah berikutnya yang dilakukan setelah data pencilan dianalisis dengan menggunakan metode *MCD-Bootstrap* adalah menganalisis data pencilan dengan metode *LAD-Bootstrap*. Untuk menduga nilai parameter regresi, data yang mengandung pencilan sebesar 10%, 15%, dan 20% akan dianalisis dengan metode LAD.

Selanjutnya dilakukan *resampling* sebanyak 500, 750, dan 1000 kali ulangan *bootstrap residual*. Berikut hasil analisis dengan metode *LAD-Bootstrap*.

Tabel 10. Analisis Metode *LAD-Bootstrap* dengan B=500 , SK= 95%

Persentase Pencilan	Parameter	Estimasi	Selang Kepercayaan 95%		Keterangan
			Batas Bawah	Batas Atas	
10%	β_0	40,07568	35,36376	63,40877	Tidak bias
	β_1	1,12953	0,73557	1,34715	Tidak bias
	β_2	0,82608	0,64581	1,19472	Tidak bias
15%	β_0	42,77278	27,95160	60,99660	Tidak bias
	β_1	1,25421	0,53560	1,44712	Tidak bias
	β_2	0,62913	0,40082	1,04973	Tidak bias
20%	β_0	40,09636	19,78189	59,82690	Tidak bias
	β_1	1,25642	0,61194	1,57078	Tidak bias
	β_2	0,83076	0,46264	1,01156	Tidak bias

Tabel 11. Analisis Metode *LAD-Bootstrap* dengan B=500 , SK= 99%

Persentase pencilan	Parameter	Estimasi	Selang Kepercayaan 99%		Keterangan
			Batas Bawah	Batas Atas	
10%	β_0	40,57057	23,18437	60,22938	Tidak bias
	β_1	1,23771	0,66961	1,31311	Tidak bias
	β_2	1,23771	0,37132	1,02023	Tidak bias
15%	β_0	44,32462	34,03279	71,07779	Tidak bias
	β_1	1,25721	0,43557	1,44715	Tidak bias
	β_2	0,50838	0,23100	0,97991	Tidak bias
20%	β_0	41,26789	26,81105	63,85606	Tidak bias
	β_1	1,05865	0,59634	1,48638	Tidak bias
	β_2	0,85175	0,46003	1,00895	Tidak bias

Tabel 12. Analisis Metode *LAD-Bootstrap* dengan B=750 , SK= 95%

Persentase pencilan	Parameter	Estimasi	Selang Kepercayaan 95%		Keterangan
			Batas Bawah	Batas Atas	
10%	β_0	42,43778	28,97793	56,02294	Tidak bias
	β_1	1,25330	0,56247	1,22025	Tidak bias
	β_2	0,51643	0,32821	0,97713	Tidak bias
15%	β_0	43,02655	30,24352	67,28852	Tidak bias
	β_1	1,14947	0,64704	1,33568	Tidak bias
	β_2	0,82991	0,62262	1,17153	Tidak bias
20%	β_0	44,76784	17,86732	64,91232	Tidak bias
	β_1	1,12560	0,73958	1,44314	Tidak bias
	β_2	0,69551	0,55432	1,00323	Tidak bias

Tabel 13. Analisis Metode *LAD-Bootstrap* dengan B=750 , SK= 99%

Persentase Pencilan	Parameter	Estimasi	Selang Kepercayaan 99%		Keterangan
			Batas Bawah	Batas Atas	
10%	β_0	41,40304	26,33562	61,38063	Tidak bias
	β_1	1,23694	0,46899	1,41373	Tidak bias
	β_2	0,95377	0,48598	1,13489	Tidak bias
15%	β_0	40,08556	24,20022	61,24522	Tidak bias
	β_1	1,03437	0,42503	1,05769	Tidak bias
	β_2	0,77872	0,38757	0,93648	Tidak bias
20%	β_0	43,32162	27,10799	68,15299	Tidak bias
	β_1	1,04535	0,43056	1,15216	Tidak bias
	β_2	0,93639	0,66437	1,01328	Tidak bias

Tabel 14. Analisis Metode *LAD-Bootstrap* dengan B=1000 , SK= 95%

Persentase Pencilan	Parameter	Estimasi	Selang Kepercayaan 95%		Keterangan
			Batas Bawah	Batas Atas	
10%	β_0	42,90844	20,68882	54,73383	Tidak bias
	β_1	1,23405	0,67377	1,40895	Tidak bias
	β_2	1,01264	0,63523	1,18414	Tidak bias
15%	β_0	42,83407	18,41196	57,45697	Tidak bias
	β_1	1,23781	0,63526	1,44746	Tidak bias
	β_2	0,94167	0,73341	1,38232	Tidak bias
20%	β_0	41,93689	36,29397	68,33898	Tidak bias
	β_1	1,04316	0,61906	1,46366	Tidak bias
	β_2	0,63031	0,48550	1,03441	Tidak bias

Tabel 15. Analisis Metode LAD-Bootstrap dengan B=1000, SK= 99%

Persentase pencilan	Parameter	Estimasi	Selang Kepercayaan 99%		Keterangan
			Batas Bawah	Batas Atas	
10%	β_0	43,39485	25,57769	62,62269	Tidak bias
	β_1	1,24799	0,65705	1,32567	Tidak bias
	β_2	0,72084	0,34372	0,99263	Tidak bias
15%	β_0	43,61488	20,18021	68,22522	Tidak bias
	β_1	1,04325	0,46212	1,52060	Tidak bias
	β_2	0,52793	0,24731	0,89623	Tidak bias
20%	β_0	42,87202	22,86275	70,90776	Tidak bias
	β_1	1,34966	0,45913	1,52359	Tidak bias
	β_2	0,63904	0,35345	0,90237	Tidak bias

Pada Tabel 10, 11, 12, 13, 14, dan 15, menunjukkan bahwa data yang mengandung pencilan sebesar 10%, 15%, 20% yang telah dianalisis dengan metode LAD-Bootstrap bersifat tidak bias. Hal ini terlihat dari nilai parameter β_0, β_1 dan β_2 berada pada selang kepercayaan 95% dan 99%.

4.6 Perbandingan Hasil Analisis Metode MCD-Bootstrap dengan LAD-Bootstrap

Perbandingan antara hasil analisis metode MCD-Bootstrap dengan LAD-Bootstrap dapat dilihat dari bias parameter dan lebar selang kepercayaan yang diperoleh. Pada analisis MCD-Bootstrap dan LAD-Bootstrap sama-sama menghasilkan penduga parameter yang tidak bias pada data yang mengandung pencilan sebesar 10%, 15% dan 20% dengan *resampling* sebanyak 500, 750, dan 1000 kali ulangan.

Selang kepercayaan yang digunakan sebesar 95% dan 99%. Selang kepercayaan yang dianggap mengandung nilai parameter sebenarnya terdiri dari batas bawah dan batas atas. Lebar selang kepercayaan dapat menjadi perbandingan ketepatan dari metode MCD-Bootstrap dan LAD-Bootstrap dalam mengatasi pencilan. Lebar selang diperoleh dari selisih batas atas dengan batas bawah yang dihasilkan oleh penduga parameter pada selang kepercayaan. Lebar selang yang pendek menunjukkan pendugaan yang lebih akurat. Demikian pula sebaliknya, lebar selang yang

panjang menunjukkan pendugaan yang kurang akurat. Lebar selang pada metode MCD-Bootstrap dan LAD-Bootstrap diuraikan pada tabel berikut ini:

Tabel 16. Lebar selang saat B= 500

Persentase Pencilan	Parameter	Selang Kepercayaan 95%		Selang Kepercayaan 95%	
		MCD-Bootstrap	LAD-Bootstrap	MCD-Bootstrap	LAD-Bootstrap
10%	β_0	25,19479	28,04500	32,80226	37,04500
	β_1	0,26388	0,61157	0,32100	0,64350
	β_2	0,41002	0,54891	0,39634	0,64891
15%	β_0	29,52122	33,04500	36,29345	37,04500
	β_1	0,38456	0,91152	0,26186	1,01157
	β_2	0,46406	0,64891	0,60306	0,74891
20%	β_0	29,03021	40,04500	35,31176	36,14509
	β_1	0,18200	0,95884	0,23377	0,89003
	β_2	0,44611	0,54891	0,48159	0,54891

Tabel 17. Lebar selang saat B= 750

Persentase Pencilan	Parameter	Selang Kepercayaan 95%		Selang Kepercayaan 95%	
		MCD-Bootstrap	LAD-Bootstrap	MCD-Bootstrap	LAD-Bootstrap
10%	β_0	25,23323	27,04506	34,97404	35,04512
	β_1	0,46140	0,65777	0,30284	0,94473
	β_2	0,30996	0,64891	0,42519	0,64891
15%	β_0	31,94821	37,24701	34,41892	37,60590
	β_1	0,19369	0,68864	0,35394	0,63265
	β_2	0,47913	0,54891	0,50333	0,54891
20%	β_0	29,83315	47,02407	37,91640	41,04500
	β_1	0,29782	0,70356	0,33888	0,72159
	β_2	0,56180	0,64891	0,45374	0,64891

Tabel 18. Lebar selang saat B= 1000

Persentase Pencilan	Parameter	Selang Kepercayaan 95%		Selang Kepercayaan 95%	
		MCD-Bootstrap	LAD-Bootstrap	MCD-Bootstrap	LAD-Bootstrap
10%	β_0	26,33479	34,14533	34,83377	37,12497
	β_1	0,36814	0,73518	0,30906	0,66861
	β_2	0,32008	0,54891	0,42359	0,64891
15%	β_0	31,19448	39,12918	41,64960	48,20450
	β_1	0,30017	0,81219	0,26261	1,05848
	β_2	0,58492	0,64891	0,51456	0,64891
20%	β_0	29,80075	32,04522	38,59047	48,04576
	β_1	0,20493	0,84459	0,43373	1,06446
	β_2	0,36364	0,54891	0,47259	0,54891

Berdasarkan Tabel 16, 17, dan 18 menunjukkan bahwa lebar selang yang dihasilkan oleh metode MCD-Bootstrap lebih pendek dibandingkan lebar selang yang dihasilkan oleh metode LAD-Bootstrap saat

resampling sebanyak 500, 750, dan 1000 kali ulangan dan pada selang kepercayaan 95% dan 99% untuk penduga parameter β_0 , β_1 dan β_2 .

5. KESIMPULAN DAN SARAN

Berdasarkan hasil analisis dan pembahasan, didapatkan kesimpulan sebagai berikut:

1. Metode MCD–*Bootstrap* dapat mengatasi pencilan pada data yang mengandung pencilan sebesar 10%, 15% dan 20%. Hal ini dapat dilihat dari metode ini menghasilkan penduga parameter yang tidak bias saat ulangan 500, 750, dan 1000 kali ulangan dan pada selang kepercayaan 95% dan 99%.
2. Metode LAD–*Bootstrap* dapat mengatasi pencilan pada data yang mengandung pencilan sebesar 10%, 15% dan 20%. Hal ini dapat dilihat dari metode ini menghasilkan penduga parameter yang tidak bias saat ulangan 500, 750, dan 1000 kali ulangan dan pada selang kepercayaan 95% dan 99%.
3. Keakuratan metode MCD–*Bootstrap* lebih tinggi dibandingkan metode LAD–*Bootstrap*. Hal ini dapat ditunjukkan dari lebar selang yang dihasilkan oleh metode MCD–*Bootstrap* lebih pendek saat ulangan 500, 750, dan 1000 kali.

Adapun saran yang diberikan pada penelitian ini adalah perlu dikakukan penelitian lebih lanjut untuk membandingkan metode *robust* yang lainnya seperti *Least Trimmed Square* (LTS), estimasi-M, estimasi-MM dalam mengatasi pengaruh pencilan pada data.

DAFTAR PUSTAKA

- El-Salam, M. (2013). The Efeciency of Some Robust Ridge Regression for Handling Multicollinearity and Non-Normal Errors Problem. Vol.7, No.77, Hal. 3831-3846.
- Hubert, M., & Debruyne, M. (2009). Minimum Covariance Determinant. *WIREs Computational Statistics*, Vol.2, Hal.36-43.
- Neter, J., Wasserman, W., & Kutner, M. (1997). *Model Linear Terapan Buku I: Analisis Regresi Linear Sederhana*. (Terjemahan Bambang Sumantri) Bandung: FMIPA-IPB.
- Sprent, P. (1989). *Applied Nonparametric Statistical Methods*. New York: Chapman & Hall.
- Sungkono, J. (2013). Resampling Bootstrap Pada R. *Magistra* No. 84, Hal. 47-54.