

Regresi Robust Dengan Metode Estimasi-S*Robust Regression Method To Estimate - S***Febryanto Perdana Hidayatulloh¹, Desi Yuniarti², dan Sri Wahyuningsih³**

Laboratorium Statistika Terapan FMIPA Universitas Mulawarman

Program Studi Statistika FMIPA Universitas Mulawarman

¹Mahasiswa Program Studi Statistika FMIPA Universitas Mulawarman^{2,3}Dosen Program Studi Statistika FMIPA Universitas MulawarmanE-mail: febryantoperdana@yahoo.com¹, Desy_yunt@yahoo.com², swahyuningsih@gmail³**Abstract**

Electrical energy is a daily necessity for human beings may not be able to do the activity without electricity. Electrical energy needs will continue to increase in line with the regional economy. Sales of electrical energy can be predicted by using regression analysis. Sales of electric energy and the amount of electrical energy customer categories of households in East Kalimantan from 2010 until 2013 there is a data outlier, so it requires the appropriate method to perform data analysis. Robust Regression is a regression to overcome the irregularities caused by outliers. If there are outliers in the dependent variable (Y) and the independent variable (X), then the robust regression-S exact estimation used to estimate parameters. The parameter estimation obtained based on the analysis result are [-9822,16 ; 2,683]. Electrical energy sales data (Y) household categories East Kalimantan province from 2010 until 2013 contains outliers in observation to 16 , 21 and 30. The number of customers of electrical energy (X) household categories East Kalimantan province from 2010 until 2013 contains outliers in observation to 8 , 17 , and 26. By using a robust regression analysis estimates - S is known that there are significant customers of electrical energy to electrical energy sales .

Keywords : Electrical energy, S-estimation, robust regression.

Pendahuluan

Analisis regresi adalah suatu metode sederhana untuk melakukan investigasi tentang hubungan fungsional di antara beberapa variabel. Hubungan antara beberapa variabel tersebut diwujudkan dalam suatu model matematis. Pada model regresi, variabel dibedakan menjadi dua bagian, yaitu variabel respon (*response*) atau biasa disebut variabel bergantung (*dependent variable*) serta variabel *explanatory* atau biasa juga disebut variabel penduga (*predictor variable*) atau disebut juga variabel bebas (*independent variable*) (Nawari, 2010).

Dalam kasus model regresi linier, dimungkinkan terdapat data *outlier* (pencilan) yaitu pengamatan dengan nilai mutlak residual jauh lebih besar daripada residual-residual lain sehingga akan mempengaruhi model regresi yang terbentuk. Data pencilan tersebut tidak boleh dibuang begitu saja karena akan mempengaruhi model prediksi serta menghasilkan estimasi parameter yang kurang tepat. Untuk menyelesaikan masalah tersebut diperlukan adanya metode yang bersifat *robust* dimana nilai estimasinya tidak banyak dipengaruhi oleh perubahan kecil dalam data. Regresi *robust* merupakan metode regresi yang digunakan ketika distribusi dari residual tidak normal atau adanya beberapa pencilan yang berpengaruh pada model. Metode ini merupakan alat penting untuk menganalisa data yang

dipengaruhi oleh pencilan sehingga dihasilkan model yang *robust* atau *resistance* terhadap pencilan. Dalam regresi *robust* terdapat beberapa metode estimasi seperti estimasi-M (*M-estimation*), estimasi *Least Median Square* (LMS), estimasi *Least Trimmed Square* (LTS), estimasi-S (*S-estimation*), estimasi *Method of Moment* (MM) (Chen, 2002).

Penelitian ini penulis menggunakan metode lain dalam regresi *robust* yaitu metode estimasi-S (*S-estimation*). Metode estimasi-S pertama kali dikembangkan oleh Rousseeuw dan Yohai (1984) dimana metode ini merupakan keluarga *high breakdown point* yaitu ukuran umum proporsi dari data pencilan yang dapat ditangani sebelum pengamatan tersebut mempengaruhi model prediksi. Disebut estimasi-S karena mengestimasi berdasarkan skala. Skala yang digunakan adalah standar deviasi residual. Metode estimasi-S mempunyai kelebihan yaitu bisa digunakan untuk pencilan dengan proporsi hingga 50% serta digunakan ketika variabel dependen dan variabel independen terdapat pencilan. Metode ini menggunakan nilai pembobot dengan fungsi *Tukey's biweight* dengan nilai *breakdown value* = 0,5 dimana nilai konstanta $c = 1,547$. Menggunakan fungsi pembobotan *Tukey's biweight* karena iterasi yang digunakan lebih sedikit dibandingkan fungsi pembobot yang lain (Salibian dan Yohai, 2006).

Salah satu contoh data yang dapat dianalisis menggunakan analisis regresi adalah data tentang penjualan energi listrik. Energi listrik merupakan kebutuhan sehari-hari karena manusia tidak mungkin bisa melakukan aktivitas tanpa adanya listrik. Kebutuhan energi listrik akan terus meningkat sejalan dengan roda perekonomian daerah. Penjualan energi listrik dapat diprediksi berdasarkan data yang diperoleh menggunakan analisis regresi untuk mengetahui hubungan antara variabel dependen dalam hal ini adalah penjualan energi listrik (Y) dengan variabel independen dalam hal ini jumlah pelanggan (X). Penjualan energi listrik tersebut memiliki keragaman nilai sehingga dapat menyebabkan terjadinya beberapa pencilan.

Adapun tujuan yang ingin dicapai dari penelitian ini adalah mengetahui pengaruh jumlah pelanggan pada data penjualan energi listrik dan jumlah pelanggan energi listrik kategori rumah tangga di Kalimantan Timur tahun 2010 sampai 2013 dengan menggunakan regresi *robust* estimasi-S.

Analisis Regresi

Dalam ilmu statistika salah satu teknik yang umum digunakan dalam menganalisis hubungan diantara dua atau lebih variabel adalah analisis regresi, karena pada dasarnya perubahan nilai suatu variabel tidak selalu terjadi dengan sendirinya, hal tersebut dapat disebabkan oleh berubahnya variabel lain yang berhubungan dengan variabel. Analisis regresi juga merupakan salah satu alat statistika yang digunakan dalam pengambilan keputusan yang banyak digunakan dalam pembangunan model matematis. Perkembangan ilmu pengetahuan yang begitu pesat memerlukan model matematis yang andal yang akan mempermudah proses identifikasi variabel kunci, peramalan, maupun dalam merancang suatu proyek simulasi (Sembiring, 1995).

Analisis Regresi Linier Sederhana

Analisis regresi merupakan analisis ketergantungan dari satu atau lebih variabel independen terhadap satu variabel dependen, dengan tujuan untuk menduga atau memprediksi nilai rata-rata populasi berdasarkan nilai-nilai variabel independennya. Analisis regresi yang digunakan untuk memprediksi satu variabel dependen berdasarkan pada satu variabel independen disebut dengan analisis regresi sederhana (Suliyanto, 2011).

Estimasi Parameter Model Regresi Linier Sederhana Menggunakan Metode Kuadrat Terkecil

Prinsip dasar metode kuadrat terkecil ini adalah meminimumkan jumlah kuadrat residual. Dalam matematika, untuk mendapatkan nilai

minimum dari suatu fungsi maka syaratnya adalah differensiasi atau turunan pertama dari fungsi tersebut harus sama dengan nol. Dengan demikian untuk meminimumkan kuadrat residual tersebut maka kita harus melakukan differensiasi parsial jumlah residual kuadrat terhadap $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$ (Widarjono, 2007).

Pengujian Signifikan Parameter Model Regresi Sederhana

Sebelum menarik kesimpulan dari suatu model regresi, terlebih dahulu dilakukan pengujian terhadap parameter regresi baik secara simultan (bersama-sama) maupun secara parsial (individu). Hal ini bertujuan untuk mengetahui variabel-variabel independen yang mempunyai pengaruh maupun tidak terhadap variabel dependen (Sudjana, 2002).

Uji Simultan (Uji F)

Uji F pada dasarnya menunjukkan apakah semua variabel independen yang dimasukkan dalam model mempunyai pengaruh secara bersama-sama terhadap variabel dependen. Statistik uji yang digunakan adalah Uji F melalui tabel Analisis Variansi (ANAVA).

Tabel 1 Analisis Variansi (ANAVA)

Sumber Variansi	Derajat Bebas (db)	Jumlah Kuadrat(JK)	Rata-rata Jumlah Kuadrat (RJK)	Uji F
Regresi	k	JKR	RJKR	F_{hitung}
Residual	$n-k-1$	JKG	RJKG	
Total	$n-1$	JKT		

Sumber : Sudjana(2002)

Uji Parsial (Uji t)

Uji t pada dasarnya seberapa jauh pengaruh satu variabel independen secara individual dalam menerangkan variansi variabel dependen. Statistik uji yang akan dilakukan adalah Uji t dari distribusi t -student dengan rumus (widarjono, 2007) :

$$t_{hit} = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \tag{1}$$

Asumsi Regresi Linier Sederhana

Asumsi-asumsi yang harus dipenuhi adalah residual data berdistribusi normal, tidak terdapat autokorelasi, dan tidak terjadi masalah heteroskedastisitas (Suliyanto, 2011).

Residual Data Berdistribusi Normal

Uji kenormalan adalah untuk melihat apakah nilai residual berdistribusi normal atau tidak. Model regresi yang baik adalah memiliki nilai residual yang terdistribusi normal, jadi uji normalitas bukan dilakukan pada masing-masing variabel tetapi pada nilai residualnya.

Uji kenormalan dapat dilakukan dengan uji Jarque-Bera. Uji ini merupakan uji normalitas dengan berdasarkan pada koefisien keruncingan (*kurtosis*) dan koefisien kemiringan (*skewness*). Uji ini dilakukan dengan membandingkan statistik Jarque-Bera (JB) dengan nilai $\chi^2_{(\alpha,k)}$.

Jika nilai $JB \leq \chi^2_{(\alpha,k)}$ maka nilai residual terstandarisasi dinyatakan berdistribusi normal. Untuk menghitung nilai statistik Jarque-Bera (JB) digunakan rumus berikut (Suliyanto, 2011) :

$$JB = n \left[\frac{S^2}{6} + \frac{(K-3)^2}{24} \right] \tag{2}$$

Tidak Ada Autokorelasi

Autokorelasi merupakan asumsi residual yang memiliki komponen atau nilai yang berkorelasi berdasarkan waktu (urutan waktu) pada himpunan data itu sendiri. Autokorelasi menjelaskan adanya korelasi atau hubungan yang erat antar residualnya, dapat pula didefinisikan sebagai korelasi kesalahan pengganggu (e_i) yang satu dengan kesalahan pengganggu lainnya (e_j). Pada pengujian asumsi ini diharapkan asumsi autokorelasi tidak terpenuhi. Uji Durbin-Watson (uji D-W) merupakan uji yang sangat populer untuk menguji ada-tidaknya masalah autokorelasi dari model empiris yang diestimasi. Untuk menghitung nilai *Durbin Watson* digunakan rumus (Widarjono, 2007) :

$$d_W = \frac{\sum_{i=2}^n (\varepsilon_i - \varepsilon_{i-1})^2}{\sum_{i=1}^n \varepsilon_i^2} \tag{3}$$

Tidak Terjadi Masalah Heteroskedastisitas

Heteroskedastisitas berarti ada varian variabel pada model regresi yang tidak sama (konstan), yang diharapkan ada pada model regresi (tidak terjadi masalah heteroskedastisitas). Untuk mendeteksi adanya masalah heteroskedastisitas dapat digunakan metode *White*. Uji heteroskedastisitas menggunakan uji *White* dilakukan dengan meregresikan semua variabel independen, variabel independen kuadrat dan perkalian (interaksi) variabel independen terhadap nilai residual kuadratnya. Nilai χ^2 hitung dalam metode ini diperoleh dari (Suliyanto, 2011) :

$$\chi^2_{hitung} = n \times R^2 \tag{4}$$

Koefisien Determinasi

Koefisien determinasi berguna untuk mengetahui kontribusi model terhadap variasi data yang ada atau besarnya pengaruh semua variabel independen terhadap variabel dependen.

Secara umum R^2 dapat didefinisikan sebagai berikut (Widarjono, 2007) :

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad i = 1, 2, \dots, n \tag{5}$$

Data Pencilan (Outlier)

Secara umum data pencilan atau dikenal juga dengan data ekstrim adalah data yang berbeda dari data-data lainnya pada suatu hasil, kemungkinan nilainya terlalu besar (lebih besar dari observasi pada umumnya) atau terlalu kecil. Pencilan merupakan suatu keganjilan dan menandakan suatu titik yang khas dibandingkan pengamatan lain. Oleh karena itu pencilan perlu diperiksa secara seksama, kemungkinan saja alasan keganjilan dapat diketahui dan berpengaruh besar terhadap koefisien regresi (Soemartini, 2007).

Dampak Data Pencilan

Keberadaan pencilan dalam data hasil akan mengganggu dalam proses analisis data dan harus dihindari dalam banyak hal. Dalam kaitannya dengan analisis regresi, data pencilan dapat menyebabkan hal-hal berikut (Soemartini, 2007):

1. Residual yang besar dari model yang terbentuk atau $E[\varepsilon] \neq 0$.
2. Variansi pada data tersebut menjadi lebih besar.
3. Taksiran interval memiliki rentang yang lebar.

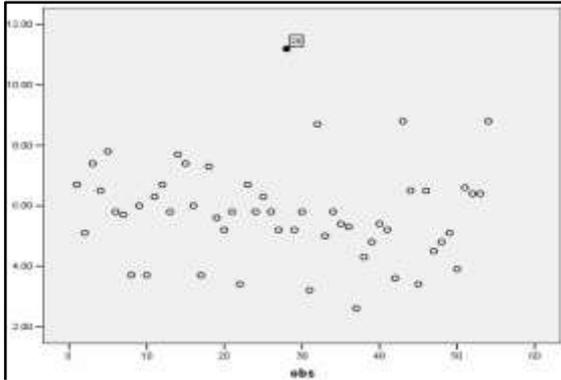
Identifikasi Pencilan

Dalam statistika, data pencilan harus dilihat terhadap posisi dan sebaran data yang lainnya sehingga akan dievaluasi apakah terdapat data pencilan atau tidak. Dalam statistika terdapat beberapa cara untuk mengetahui ada tidaknya data pencilan diantaranya sebagai berikut.

Metode Grafis

Untuk melihat apakah terdapat pencilan pada data, dapat dilakukan dengan memplot antara data dengan observasi ke- i ($i = 1, 2, 3, \dots, n$) seperti Gambar 1.

Berdasarkan Gambar 1 terdapat salah satu data, yakni observasi ke-28 yang mengindikasikan merupakan pencilan.



Gambar 1. Contoh Scatter-Plot dari Data dengan Observasi ke-*i*

Pendeteksian Pencilan dengan Uji Studentized Delete Residual (TRES)

Uji TRES digunakan untuk mendeteksi pencilan pada variabel *Y*. Statistik uji yang digunakan pada uji TRES adalah (Wulansari, 2012) :

$$TRES = \varepsilon_i \left(\frac{n - p - 2}{JKG(1 - h_{ii}) - \varepsilon_i^2} \right)^{\frac{1}{2}} \quad (6)$$

Metode yang digunakan dalam mengidentifikasi pencilan terhadap variabel *X* adalah Nilai Pengaruh (*Leverage Point*). Nilai Pengaruh (h_{ii}) dari pengamatan (X_i, Y_i) menunjukkan besarnya peranan Y_i terhadap \hat{Y}_i dan didefinisikan sebagai (Wulansari, 2012) :

$$H = X^T (X^T \cdot X)^{-1} X \quad (7)$$

Tindakan Terhadap Pencilan

Bila ternyata hasil indentifikasi menunjukan adanya pencilan, maka data cenderung tidak normal, para peneliti seringkali melakukan transformasi pada data dengan bermacam-macam transformasi, namun dengan mentransformasi data kadangkala tidak membuat data tersebut menjadi lebih baik. Sering pula para peneliti membuang observasi ke-*i* pada data tersebut yang dianggap sebagai pencilan. Meskipun pencilan identik dengan data yang tidak bagus, akan tetapi itu merupakan bagian terpenting dari suatu data, karena setiap observasi akan mengandung suatu informasi. Untuk itu maka diperlukan suatu analisis yang *robust* terhadap data pencilan (Soemartini, 2007).

Regresi Robust

Regresi *robust* merupakan metode regresi yang digunakan ketika distribusi residual tidak normal. Distribusi residual tidak normal pada umumnya terjadi karena data terkontaminasi oleh pencilan. Metode ini merupakan alat penting untuk menganalisis data yang dipengaruhi pencilan sehingga dihasilkan model

yang *robust* atau resisten terhadap pencilan. Suatu estimasi yang resisten artinya estimasi ini relatif tidak terpengaruh oleh perubahan besar pada bagian kecil data atau perubahan kecil pada bagian besar data (Sugi, 2002).

Prosedur *robust* ditunjukkan untuk mengakomodasi adanya keanehan data, sekaligus meniadakan indentitas adanya data pencilan dan juga bersifat otomatis dalam menanggulangi data pencilan. Analisis regresi *robust* ini tidak membuat residual model menjadi normal namun model yang dihasilkan oleh metode ini memiliki keakuratan yang lebih tinggi dari model yang dihasilkan oleh model MKT. Ketika menggunakan alat-alat analisis, pada umumnya langkah pertama adalah mencoba menghapus pencilan kemudian data yang sudah bagus dengan menggunakan metode kuadrat terkecil, tetapi analisis *robust* mencocokkan model regresi dengan sebagian data dan kemudian mengatasi titik-titik pencilan sehingga memperoleh estimasi parameter yang *robust*. Jadi regresi *robust* ini tidak membuang bagian dari data melainkan menemukan model cocok dari sebagian data (Soemartini, 2007).

Estimasi-S

Estimasi-S pertama kali diperkenalkan oleh Rousseeuw dan Yohai (1984) merupakan estimasi *robust* yang dapat mencapai *breakdown point* hingga 50%. *Breakdown point* adalah ukuran umum proporsi dari pencilan yang dapat ditangani sebelum pengamatan tersebut mempengaruhi model. Karena estimasi-S dapat mencapai *breakdown point* hingga 50% maka estimasi-S dapat mengatasi setengah dari pencilan dan memberikan pengaruh yang baik bagi pengamatan lainnya. estimasi parameter regresi *robust* dengan IRLS, untuk $l+1$ iterasi adalah (Susanti, 2013) :

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

Energi Listrik

Energi listrik adalah energi utama yang dibutuhkan bagi peralatan listrik atau energi yang tersimpan dalam arus listrik dengan satuan *ampere* (A) dan tegangan listrik dengan satuan *Volt* (V) dengan ketentuan kebutuhan konsumsi daya listrik dengan satuan *Watt* (W) yang diantaranya untuk menggerakkan motor, lampu penerangan, memanaskan, mendinginkan ataupun untuk menggerakkan kembali suatu peralatan mekanik untuk menghasilkan bentuk energi yang lain (Sulasno, 2013).

Energi listrik merupakan salah satu komponen terpenting untuk menunjang pembangunan suatu bangsa. Peningkatan pembangunan, penambahan jumlah penduduk dan peningkatan taraf hidup menyebabkan laju

konsumsi energi listrik semakin meningkat(Sulasno, 2013).

Proses penyaluran energi listrik dimulai dengan membangkitkan tenaga listrik di pusat-pusat pembangkit listrik seperti PLTA, PLTD dan PLTN. Dari pusat pembangkit, listrik disalurkan ke Gardu Induk Tegangan Extra Tinggi (GITET) atau Gardu Induk (GI). Kemudian dari Gardu Induk disalurkan ke pelanggan yang telah terpasang atau terdaftar(Sulasno, 2013).

Metodologi Penelitian

Variabel dalam penelitian ini adalah variabel dependen yang dinotasikan dengan Y adalah data penjualan energi listrik (MWh) kategori rumah tangga Provinsi Kalimantan Timur. Variabel independen yang dinotasikan X adalah jumlah pelanggan energi listrik kategori rumah tangga Provinsi Kalimantan Timur.

Hasil dan Pembahasan

a. Analisis Deskriptif

Berikut merupakan hasil analisis deskriptif yang mendeskripsikan keadaan sekelompok data yaitu berupa banyaknya data, rata-rata, median, modus, standar deviasi, total, nilai maksimum dan nilai minimum pada data.

Tabel 2 Hasil Analisis Statistik Deskriptif

	Penjualan Energi Listrik (MWh)	Jumlah Pelanggan Energi Listrik
Banyak Data	36,000	36,000
Rata-rata	137.981,289	155.851,694
Median	67.470,470	28.146,000
Modus	22,138,900	11.402,000
Standar Deviasi	142.624,488	52.730,364
Nilai Minimum	22.138,900	11.402,000
Nilai Maksimum	529.538,720	205.166,000
Total	4.967.326,420	2.010.661,000

b. Model Awal Regresi

Model regesi pada penelitian ini, yaitu :

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

Dimana :

Y : Banyaknya Energi Listrik yang Terjual Kategori Rumah Tangga(MWh)

X : Jumlah Pelanggan Energi Listrik Kategori Rumah Tangga

ε : Nilai residual, dimana $\varepsilon \sim N(0, \sigma^2)$

c. Estimasi Parameter Model Regresi dengan Metode Kuadrat Terkecil

Parameter model regresi (β_0, β_1) diestimasi menggunakan model MKT sehingga diperoleh model regresi dengan metode MKT adalah :

$$\hat{Y} = -11615,887 + 2,678 X$$

Berdasarkan model yang diperoleh, maka setiap penambahan satu rumah pelanggan energi listrik akan menaikkan penjualan energi listrik kategori rumah tangga di Provinsi Kalimantan Timur sebesar 2,678 MWh. Apabila jumlah pelanggan energi listrik kategori rumah tangga adalah nol maka penjualan energi listrik kategori rumah tangga berkurang sebesar 11615,887 MWh.

d. Pengujian Signifikansi Parameter

Uji Parameter Secara Simultan (Uji F)

Berdasarkan hasil analisis diketahui nilai $|F_{hitung}| (1721,754) > F_{tabel(0,05;1;34)} (4,130)$, maka H_0 ditolak atau dapat dilihat dari nilai $Sig. (0,000) < \alpha (0,05)$, maka H_0 ditolak. Sehingga dapat disimpulkan model regresi sudah tepat digunakan untuk memprediksi penjualan energi listrik.

Uji Parameter Secara Parsial (Uji t)

Berdasarkan hasil analisis diketahui nilai $|t_{hitung}| (\beta_0(2,358))(\beta_1(41,494)) > t_{tabel(0,025;33)} (2,035)$, maka H_0 ditolak. atau dapat dilihat dari nilai $Sig.(\beta_0 (0,024)) (\beta_1(0,000)) < \alpha (0,05)$, maka H_0 ditolak. Sehingga dapat ditarik kesimpulan bahwa terdapat pengaruh variabel konstan dan pengaruh jumlah pelanggan energi listrik kategori rumah tanggayang signifikan terhadap penjualan energi listrik kategori rumah tangga.

e. Uji Asumsi Klasik Regresi Linier Sederhana

Berdasarkan hasil analisis uji residual kenormalan dengan Jarque-Beradiketahui nilai $JB (4,592) < \chi^2_{(0,05;2)} (5,991)$ dan $Sig. (0,101) > \alpha (0,05)$, maka H_0 gagal ditolak. Sehingga dapat disimpulkan bahwa residual berdistribusi normal.

Berdasarkan hasil analisis uji autokorelasi (Durbin-Watson)dengan $\alpha = 5\%$, jumlah pengamatan 36 dan banyak parameter, maka diperoleh nilai d_L sebesar 1,354 dan d_U sebesar 1,587, sedangkan nilai d_W sebesar 2,331. Karena $d_U (1,587) \leq d_W (2,331) \leq 4-d_U (2,413)$ maka dapat disimpulkan bahwa tidak terjadi masalah autokorelasi.

Berdasarkan uji heteroskedastisitas metode *White* diketahuinilai $Sig. (0,005) < \alpha = 0,05$, maka diputuskan H_0 ditolak. Sehingga dapat ditarik kesimpulan Terjadi masalah heteroskedastisitas.

f. Koefisien Determinasi

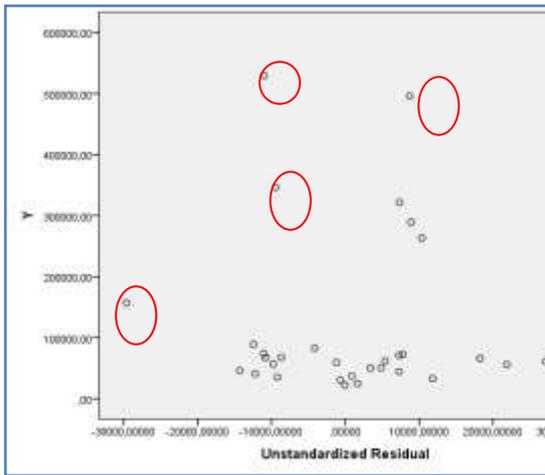
Diperoleh nilai koefisien determinasi adalah 98,1 %, artinya variasi yang terjadi terhadap penjualan energi listrik disebabkan oleh jumlah pelanggan energi listrik. Sedangkan residualnya

sebesar 1,9 % disebabkan variabel lain yang tidak dianalisis.

Nilai koefisien determinasi sebesar 0,981 dapat dikatakan sangat baik karena nilai yang mendekati satu berarti variabel-variabel dependent memberikan hampir semua informasi yang dibutuhkan untuk memprediksi variasi variabel independen.

g. Deteksi Pencilan Metode Grafis

Untuk melihat apakah terdapat pencilan pada data, dapat dilakukan dengan memplot antara data dengan observasi ke-*i* (*i* = 1, 2, 3,..., *n*).



Gambar 2. Scatterplot Antara Penjualan Energi Listrik (Y) dengan Nilai Residual

Gambar 2 memperlihatkan bahwa ada data yang terletak jauh dari kumpulan data. Data tersebut yang disebut dengan pencilan (*outlier*).

Uji TRES

Mendeteksi pencilan dapat diketahui dengan menggunakan uji TRES, untuk mengetahui pencilan terhadap Y yaitu TRES dengan menarik kesimpulan menolak H₀ apabila nilai TRES > t_(α/2; n-k-1) dan untuk mengetahui pencilan terhadap X yaitu h_{ii} dengan menarik kesimpulan menolak H₀ apabila nilai h_{ii} > 2k/n. Diperoleh nilai h_{ii} sebagai berikut :

Pengamatan	h _{ii}	Pengamatan	h _{ii}	Pengamatan	h _{ii}
1	0,032	13	0,040	25	0,056
2	0,043	14	0,037	26	0,114
3	0,036	15	0,039	27	0,036
4	0,036	16	0,071	28	0,040
5	0,033	17	0,200	29	0,048
6	0,037	18	0,032	30	0,031
7	0,094	19	0,035	31	0,044
8	0,257	20	0,047	32	0,043
9	0,030	21	0,035	33	0,045
10	0,034	22	0,044	34	0,046
11	0,048	23	0,040	35	0,087
12	0,030	24	0,041	36	0,038

Berdasarkan Tabel 3 diketahui nilai h_{ii} pada pengamatan 8, 17 dan 26 > 2k/n (0,111) sehingga H₀ ditolak, sehingga dapat disimpulkan bahwa terdapat pencilan pada data penjualan dan pelanggan energi listrik kategori rumah tangga Provinsi Kalimantan Timur dari tahun 2010 sampai tahun 2013 yaitu pengamatan ke 8, 17 dan 26 merupakan pencilan.

Diperoleh nilai TRES sebagai berikut :

Pengamatan	TRES	t _(0,05/2; 36-2-1)	Keputusan
1	0,453	2,035	H ₀ diterima
2	0,137	2,035	H ₀ diterima
3	2,425	2,035	H₀ ditolak
4	0,487	2,035	H ₀ diterima
5	0,443	2,035	H ₀ diterima
6	0,620	2,035	H ₀ diterima
7	0,363	2,035	H ₀ diterima
8	0,477	2,035	H ₀ diterima
9	0,521	2,035	H ₀ diterima
10	0,335	2,035	H ₀ diterima
11	0,690	2,035	H ₀ diterima
12	1,408	2,035	H ₀ diterima
13	1,020	2,035	H ₀ diterima
14	0,036	2,035	H ₀ diterima
15	0,517	2,035	H ₀ diterima
16	0,491	2,035	H ₀ diterima
17	0,619	2,035	H ₀ diterima
18	0,120	2,035	H ₀ diterima
19	0,393	2,035	H ₀ diterima
20	0,176	2,035	H ₀ diterima
21	2,884	2,035	H₀ ditolak
22	1,499	2,035	H ₀ diterima
23	0,264	2,035	H ₀ diterima
24	0,365	2,035	H ₀ diterima
25	0,563	2,035	H ₀ diterima
26	1,992	2,035	H ₀ diterima
27	0,459	2,035	H ₀ diterima
28	0,338	2,035	H ₀ diterima
29	0,090	2,035	H ₀ diterima
30	2,449	2,035	H₀ ditolak
31	1,214	2,035	H ₀ diterima
32	0,459	2,035	H ₀ diterima
33	0,059	2,035	H ₀ diterima
34	0,633	2,035	H ₀ diterima
35	1,908	2,035	H ₀ diterima
36	0,365	2,035	H ₀ diterima

Berdasarkan Tabel 4 diketahui nilai TRES pada pengamatan 3, 21 dan 30 > t_(0,05/2; 36-2-1) (2,035) sehingga H₀ ditolak, sehingga dapat disimpulkan bahwa terdapat pencilan pada data penjualan dan pelanggan energi listrik kategori rumah tangga Provinsi Kalimantan Timur dari tahun 2010 sampai tahun 2013 yaitu pengamatan ke 3, 21 dan 30 merupakan pencilan.

Karena data X dan Y mengandung pencilan, maka dilakukan penaksiran parameter model regresi dengan metode regresi *robust* estimasi-S untuk mengatasi masalah pencilan dan menghasilkan model yang lebih baik.

h. Model Regresi Robust dengan Estimasi-S

Proses penghitungan estimasi-S yang iteratif dimulai dengan menentukan estimasi awal koefisien regresi, yang diperoleh dari MKT yaitu $\hat{\beta} = (-11615,887 ; 2,678)$ kemudian berdasarkan algoritma estimasi-S, dihitung nilai \hat{Y}_i dan residual ϵ_i . Proses iterasi menggunakan MKT terboboti dilanjutkan dengan menghitung residualan dan pembobot $w(u_i)$ yang baru dan dilakukan pendugaan parameter secara berulang-ulang sampai konvergen. Kekonvergenan tercapai jika koefisien regresi sudah sama dengan koefisien regresi sebelumnya.

Adapun Hasil perhitungan setiap iterasi pada estimasi-S sebagai berikut :

Tabel 5 Nilai $\hat{\beta}^s$ setiap iterasi pada Estimasi-S

Tahapan	$\hat{\beta}^s$
MKT	[-1011615,887 ; 2,678]
Iterasi 1	[-10536,706 ; 2,680]
Iterasi 2	[-10109,782 ; 2,681]
Iterasi 3	[-9938,088 ; 2,682]
Iterasi 4	[-9864,730 ; 2,682]
Iterasi 5	[-9839,181 ; 2,682]
Iterasi 6	[-9828,899 ; 2,682]
Iterasi 7	[-9825,179 ; 2,682]
Iterasi 8	[-9823,283 ; 2,683]
Iterasi 9	[-9822,525 ; 2,683]
Iterasi 10	[-9822,223 ; 2,683]
Iterasi 11	[-9822,100 ; 2,683]
Iterasi 12	[-9822,050 ; 2,683]
Iterasi 13	[-9822,029 ; 2,683]
Iterasi 14	[-9822,021 ; 2,683]
Iterasi 15	[-9822,018 ; 2,683]
Iterasi 16	[-9822,016 ; 2,683]
Iterasi 17	[-9822,016 ; 2,683]

Berdasarkan Tabel5 terlihat bahwa koefisien regresi sudah konvergen pada iterasi ke-17 karena selisih nilai $\hat{\beta}_j^{s+1}$ dan $\hat{\beta}_j^l$ pada iterasi 17 sudah 0, maka diperoleh estimasi parameternya adalah (-9822,16 ; 2,683) dan dapat dituliskan dalam model regresi yaitu :

$$\hat{Y} = -9822,16 + 2,683X$$

Berdasarkan model yang diperoleh, maka setiap penambahan satu rumah pelanggan energi listrik akan menaikkan penjualan energi listrik kategori rumah tangga di Provinsi Kalimantan Timur sebesar 2,683 MWh. Apabila jumlah pelanggan energi listrik kategori rumah tangga adalah nol maka penjualan energi listrik kateggori rumah tangga berkurang sebesar 9822,16 MWh.

i. Pengujian Signifikansi Parameter Dengan Pembobot

Uji Parameter Secara Simultan (Uji F)

Berdasarkan hasil analisis diketahui nilai $|F_{hitung}| (7387,952) > F_{tabel(0,05;1;34)} (4,130)$, maka H_0 ditolak atau dapat dilihat dari nilai Sig.

$(0,000) < \alpha (0,05)$, maka H_0 ditolak. Sehingga dapat disimpulkan bahwa model regresi sudah tepat digunakan untuk memprediksi penjualan energi listrik

Uji Parameter Secara Parsial (Uji t)

Berdasarkan hasil analisis diketahui nilai $|t_{hitung}| (\beta_0 (4,471))(\beta_1(85,953)) > t_{tabel(0,025;33)} (2,035)$, maka H_0 ditolak atau dapat dilihat dari nilai Sig. $(\beta_0 (0,000)) (\beta_1(0,000)) < \alpha (0,05)$, maka H_0 ditolak. Sehingga dapat ditarik kesimpulan bahwa terdapat pengaruh variabel konstan dan pengaruh jumlah pelanggan energi listrik kategori rumah tanggayang signifikan terhadap penjualan energi listrik kategori rumah tangga.

j. Uji Asumsi Klasik Regresi Linier Sederhana Dengan Pembobot

Berdasarkan hasil analisis uji residual kenormalan dengan Jarque-Bera diketahui nilai JB $(2,894) < \chi^2_{(0,05;2)} (5,991)$ dan Sig. $(0,235) > \alpha (0,05)$, maka H_0 gagal ditolak. Sehingga dapat ditarik kesimpulan bahwa residual berdistribusi normal.

Berdasarkan hasil analisis uji autokorelasi (Durbin-Watson) dengan $\alpha = 5\%$, jumlah pengamatan 36 dan variabel bebas 1, maka diperoleh nilai d_L sebesar 1,354 dan d_U sebesar 1,587, sedangkan nilai d_W sebesar 1,651. Karena $d_U (1,587) \leq d_W (1,651) \leq 4-d_U (2,413)$ maka dapat disimpulkan bahwa tidak terjadi masalah autokorelasi.

Berdasarkan uji heteroskedastisitas metode Whitenilai Sig. $(0,434) > \alpha = 0,05$, maka diputuskan H_0 gagal ditolak. Sehingga dapat disimpulkan bahwa tidak terjadi masalah heteroskedastisitas.

Berdasarkan pengujian asumsi klasik pada model data sekunder penjualan dan pelanggan energi listrik kategori rumah tangga Provinsi Kalimantan Timur dari tahun 2010 samapai tahun 2013 menggunakan analisis regresi *robust* dengan metode Estimasi-S diperoleh bahwa semua asumsi klasik analisi regresi sederhana telah terpenuhi.

k. Koefisien Determinasi Dengan Pembobot

Berdasarkan hasil analisis diperoleh nilai koefisien determinasi adalah 99,6 %, artinya variasi yang terjadi terhadap penjualan energi listrik disebabkan oleh jumlah pelanggan energi listrik. Sedangkan residualnya sebesar 0,4 % disebabkan variabel lain yang tidak dianalisis.

Nilai koefisien determinsai sebesar 0,996 dapat dikatakan sangat baik karena nilai yang mendekati satu berarti variabel-variabel dependent memberikan hampir semua informasi yang dibutuhkan untuk memprediksi variasi variabel independen.

Kesimpulan

Adapun kesimpulan dari penelitian ini adalah :

1. Data penjualan energi listrik (Y) kategori rumah tangga Provinsi Kalimantan Timur dari tahun 2010 sampai tahun 2013 mengandung pencilan padapengamatan ke 16, 21 dan 30. Data jumlah pelanggan energi listrik (X) kategori rumah tangga Provinsi Kalimantan Timur dari tahun 2010 sampai tahun 2013 mengandung pencilan pada pengamatan ke 8, 17, dan 26.
2. Model regresi *robust* dengan metode estimasi-S untuk data penjualan dan jumlah pelanggan energi listrik kategori rumah tangga Provinsi Kalimantan Timur dari tahun 2010 sampai tahun 2013 adalah :

$$\hat{Y} = -9822,16 + 2,683X$$
3. Dengan menggunakan analisis regresi *robust* estimasi-S diketahui bahwa terdapat pengaruh pelanggan energi listrik terhadap penjualan energi listrik.

Daftar Pustaka

Chen, Colin. 2002. *Robust Regression and Outlier Detection with the ROBUSTREG Procedur*. Paper 256-27.

Nawari. 2010. *Analisis Regresi dengan Ms. Excel 2007 dan SPSS 17*. Jakarta: PT. Elex Media Komputindo.

FMIPA Universitas Sebalas Maret Surakarta

Salibian dan Yohai. 2006. A Fast Algoritma For S-Regression Estimates. *Journal Of Computational and Graphical Statistics*, Vol 15. No 2. Hal 414-427.

Sembiring, RK. 1995. *Analisis Regresi*. Bandung : ITB Bandung.

Soemartini. 2007. *Pencilan (Outlier)*. Jatinangor: Universitas Padjajaran.

Sudjana. 2002. *Metode Statistika*. Bandung: Tarsito.

Sugi. 2002. The Robust Regression and Outlier Detection With The Rosbustreg. *Statistics and Data Analysis*, Paper 265-267.

Sulasno. 2001. *Distribusi Tenaga Listrik*. Semarang: Badan Penerbit Universitas Diponegoro

Suliyanto. 2011. *Ekonometrika Terapan Teori & Aplikasi dengan SPSS*. Cv. Andi Offset: Yogyakarta.

Susanti, Yuliana. 2013. *Optimasi Model Regresi Robust Untuk Memprediksi Produksi Kedelai Di Indonesia*. Prosiding ISBN : 978-979-16353-9-4. FMIPA

Widarjono, Agus. 2007. *Ekonometrika Teori dan Aplikasi untuk Ekonomi dan Bisnis*. Ekonisia: Yogyakarta.

Wulansari, Yurista. 2012. "Regresi Robust Dengan Generalized S-Estimation (Estimasi-S) Pada Penjualan Tenaga Listrik Di Jawa Tengah Tahun 2010".