JURNAL GAUSSIAN, Volume 2, Nomor 4, Tahun 2013, Halaman 395-404
Online di: http://ejournal-s1.undip.ac.id/index.php/gaussian

REGRESI *ROBUST* MM-*ESTIMATOR* UNTUK PENANGANAN PENCILAN PADA REGRESI LINIER BERGANDA

Sherly Candraningtyas[1], Diah Safitri[2*)], Dwi Ispriyanti[3]
[1]Mahasiswa Jurusan Statistika FSM UNDIP
[2,3]Staf Pengajar Jurusan Statistika FSM UNDIP

ABSTRACT

The multiple linear regression model is used to study the relationship between a dependent variable and more than one independent variables. Estimation method which is the most frequently be used to analyze regression is Ordinary Least Squares (OLS). OLS for linear regression models is known to be very sensitive to outliers. Robust regression is an important method for analyzing data contaminated by outliers. This paper will discuss the robust regression MM-estimator. This estimation is a combined estimation method which has a high breakdown value (LTS-estimator or S-estimator) and M-estimator. Generally, there are three steps for MM-estimator: estimation of regression parameters initial using LTS-estimators, residual and robust scale using M-estimator, and the final estimation parameter using M-estimator. The purpose of writing this paper are to detect outliers using DFFITS and determine the multiple linear regression equations containing outliers using robust regression MM-estimator. The data used is the generated data from software Minitab 14.0. Based on the analysis results can be concluded that data 21st, 27th, 34th are outliers and equation of multiple linear regression using robust regression MM-estimators is $\hat{y} = 3.97 + 0.392\,x_1 - 0.810\,x_2 - 0.263\,x_3 + 0.968\,x_4$.

Keywords : Multiple Linear Regression, Ordinary Least Square (OLS), Outliers, Robust Regression LTS-Estimator, Robust Regression M-Estimator, Robust Regression MM-Estimator

1. PENDAHULUAN

Menurut Greene (1951), model regresi linier berganda digunakan untuk mempelajari hubungan antara sebuah variabel tak bebas dan lebih dari satu variabel bebas. Berdasarkan Montgomery and Peck (1992), metode estimasi yang digunakan untuk menganalisis regresi adalah metode kuadrat terkecil. Menurut Gujarati (2007), metode kuadrat terkecil mempunyai asumsi-asumsi tertentu. Asumsi tersebut yaitu memiliki parameter-parameter yang bersifat linier, error mempunyai nilai rata-rata sebesar nol, homoskedastisitas, tidak terjadi autokorelasi, tidak terjadi multikolinieritas, dan error berdistribusi normal.

Menurut Willems and Aelst (2005), metode kuadrat terkecil untuk model regresi linier dikenal sangat sensitif terhadap pencilan pada data. Berdasarkan Draper and Smith (1992), pencilan merupakan suatu keganjilan dan menandakan suatu titik data yang sama sekali tidak sama dengan data yang lain. Penolakan pencilan yang begitu saja bukanlah langkah yang bijaksana. Adakalanya pencilan dapat memberikan informasi yang tidak bisa diberikan oleh titik data lainnya.

Menurut Chen (2002), regresi *robust* adalah metode yang penting untuk menganalisis data yang terkontaminasi oleh pencilan . Regresi robust terdiri dari 5 metode estimasi, antara lain: M-*estimator*, *Least Median Square* (LMS)-*estimator*, *Least Trimmed Square* (LTS)-*estimator*, S-*estimator*, dan MM-*estimator*. Metode estimasi MM dikenalkan oleh Yohai pada tahun 1987. Estimasi ini merupakan gabungan metode estimasi yang mempunyai nilai *breakdown* yang tinggi (LTS-*estimator* atau S-*estimator*) dan M-*estimator*.

Adapun tujuan dari penyusunan tugas akhir ini adalah

1. Mendeteksi pencilan pada data bangkitan yang digunakan dalam tugas akhir ini dengan menggunakan *DFFITS*.
2. Menentukan persamaan regresi linier berganda yang mengandung pencilan dengan menggunakan metode regresi *robust* MM-*estimator*.

2. TINJAUAN PUSTAKA

2.1 Regresi Linier Berganda

Menurut Montgomery and Peck (1992), secara umum variabel tak bebas *y* kemungkinan berhubungan dengan *k* variabel bebas. Model ini:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \varepsilon_i \qquad (1)$$

y_i adalah variabel tak bebas pada pengamatan ke-i, $x_{i1}, x_{i2}, \ldots, x_{ik}$ adalah nilai variabel bebas pada pengamatan ke-*i* dan parameter ke-*k*, $\beta_0, \beta_1, \ldots, \beta_k$ adalah parameter regresi, dan ε_i adalah error pengamatan ke-*i*.

2.2 Metode Kuadrat Terkecil Untuk Regresi Linier Berganda

Menurut Gujarati (1997), dalam kasus *k*-variabel bebas, penaksir metode kuadrat terkecil diperoleh dengan meminimumkan

$$\sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - \beta_0 - \sum_{j=1}^{k} \beta_j x_{ij})^2 \qquad (2)$$

di mana $\sum_{i=1}^{n} \varepsilon_i^2$ adalah jumlah kuadrat error. Dalam notasi skalar, metode kuadrat terkecil tercapai dalam menaksir $\beta_0, \beta_1, \ldots, \beta_k$ sehingga $\sum_{i=1}^{n} \varepsilon_i^2$ sekecil mungkin. Ini dicapai dengan menurunkan persamaan (2) secara parsial terhadap $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k$ dan menyamakan hasil yang diperoleh dengan nol.

Berdasarkan Montgomery and Peck (1992), dalam notasi matriks, metode kuadrat terkecil sama dengan meminimumkan $\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}$.

Dengan persamaan :
$$\boldsymbol{\varepsilon} = \boldsymbol{y} - \boldsymbol{X\beta} \qquad (3)$$

Oleh karena itu,
$$\begin{aligned}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} &= (\boldsymbol{y} - \boldsymbol{X\beta})'(\boldsymbol{y} - \boldsymbol{X\beta}) \\ &= \boldsymbol{y}'\boldsymbol{y} - 2\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{y} + \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X\beta} \qquad (4)\end{aligned}$$

$\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{y}$ adalah matrik 1 x 1, atau suatu skalar. Transposenya $(\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{y})' = \boldsymbol{y}'\boldsymbol{X\beta}$ adalah skalar. Estimator metode kuadrat terkecil harus memenuhi :

$$\frac{\partial \sum_{i=1}^{n} \varepsilon_i^2}{\partial \boldsymbol{\beta}} = \frac{\partial \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}}{\partial \boldsymbol{\beta}} = -2\boldsymbol{X}'\boldsymbol{y} + \boldsymbol{2X}'\boldsymbol{X}\widehat{\boldsymbol{\beta}} = \boldsymbol{0} \qquad (5)$$

Bila disederhanakan menjadi:
$$\boldsymbol{X}'\boldsymbol{X}\widehat{\boldsymbol{\beta}} = \boldsymbol{X}'\boldsymbol{y} \qquad (6)$$

Untuk menyelesaikan persamaan (6) kalikan keduanya dengan inverse dari $\boldsymbol{X'X}$. Jadi estimator kuadrat terkecil dari $\boldsymbol{\beta}$ adalah $\widehat{\boldsymbol{\beta}} = (\boldsymbol{X'X})^{-1}\boldsymbol{X'y}$ (7)

2.3 Uji Asumsi Model Regresi Linier Berganda

2.3.1 Uji Normalitas

Menurut Gujarati (2007), menjelaskan bahwa pada regresi linear klasik diasumsikan bahwa tiap ε_i didistribusikan normal dengan lambang: $\varepsilon_i \sim N(0, \sigma^2)$.
Uji statistik yang dapat digunakan untuk menguji normalitas residual adalah uji statistik Kolmogorov-Smirnov (K-S).

2.3.2 Uji Multikolinieritas

Menurut Montgomery and Peck (1992), untuk mendeteksi ada atau tidaknya multikolinieritas di dalam model regresi, dapat dilihat pada nilai VIF (*variance inflation factors*).

$$VIF_j = \frac{1}{(1-R_j^2)} \tag{8}$$

dengan R_j^2 adalah nilai koefisien determinasi yang diperoleh dari meregresikan variabel bebas x_j dengan variabel bebas lainnya. Nilai VIF>10 menunjukkan multikolinearitas yang kuat.

2.3.3 Uji Heteroskedastisitas

Menurut Gujarati (1997), lambang homoskedastisitas adalah

$$E(\varepsilon_i^2) = \sigma^2 \qquad i = 1,2,\ldots,n$$

Ada tidaknya heteroskedastisitas dapat dilakukan Uji Glejser. Glejser telah menemukan bahwa: $|e_i| = \beta_1 x_j + \varepsilon_i$ (9)
Dasar keputusannya adalah jika variabel bebas signifikan secara statistik mempengaruhi variabel tak bebas, maka ada indikasi terjadi heteroskedastisitas.

2.3.4 Uji Autokorelasi

Menurut Gujarati (1997), dengan menggunakan lambang: $E(\varepsilon_i \varepsilon_j) = 0 \qquad i \neq j$.
Secara sederhana dapat dikatakan model klasik mengasumsikan bahwa error yang berhubungan dengan pengamatan tidak dipengaruhi oleh error yang berhubungan dengan pengamatan lain yang manapun. Salah satu cara yang dapat digunakan untuk mendeteksi masalah autokorelasi adalah dengan uji Durbin Watson.

2.3.5 Uji Linieritas

Berdasarkan Gujarati (2007), uji ini dimaksudkan bahwa model regresi memiliki parameter-parameter yang bersifat linier dan model ditentukan secara tepat. Menurut Suliyanto (2011), asumsi linieritas terpenuhi jika plot antara nilai error terstandarisasi dengan nilai prediksi terstandarisasi tidak membentuk suatu pola tertentu (acak).

2.3.6 Asumsi yang lain menurut Gujarati (2007), residual mempunyai nilai rata-rata sebesar nol, E(e_i) = 0.

2.4 Goodness Of Fit

Menurut Gujarati (1997), R^2 yang dikenal dengan koefisien determinasi majemuk, merupakan besaran yang memberikan informasi untuk mengetahui dalam model regresi berganda proporsi variasi dalam y yang dijelaskan oleh variabel x_j secara bersama-sama. Rumus koefisien determinasi dapat ditunjukkan sebagai berikut:

$$R^2 = \frac{SS_R}{S_{yy}} \tag{10}$$

dengan (Montgomery and Peck, 1992):

SS_R adalah *Regression Sum of Square* $= \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y} - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n}$ (11)

S_{yy} adalah *Total Sum of Square* $= \boldsymbol{y}'\boldsymbol{y} - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n}$ (12)

SS_E adalah *Residual Sum of Square of* $= S_{yy} - SS_R = \boldsymbol{y}'\boldsymbol{y} - \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y}$ (13)

R^2 terletak antara 0 dan 1.

2.5 Uji Kecocokan Model (Uji-F)

Menurut Montgomery and Peck (1992), uji ini digunakan untuk menguji signifikansi regresi jika ada hubungan linier antara variabel tak bebas y dan variabel bebas $x_1, x_2, \ldots, x_k$. Adapun langkah-langkah analisisnya adalah:

Hipotesis :

$H_0 : \beta_1 = \beta_2 = \cdots = \beta_k = 0$

H_1 : tidak semua $\beta_j (j = 1, \ldots, k) = 0$

Tingkat Signifikansi: $\alpha = 0.05$

Statistik Uji : $F_{hitung} = \frac{SS_R/k}{SS_E/(n-k-1)} = \frac{MS_R}{MS_E}$ (14)

Kriteria Uji :

H_0 ditolak jika $F_{hitung} > F_{(\alpha,k,n-k-1)}$

2.6 Pencilan (*Outlier*)

Menurut Montgomery and Peck (1992), pencilan adalah suatu pengamatan yang ekstrim. Residual yang nilai mutlaknya jauh lebih besar daripada yang lain dan bisa jadi terletak tiga atau empat simpangan baku dari rata-ratanya adalah yang menyebabkan data sebagai pencilan. Pencilan adalah titik-titik data yang tidak setipe dengan titik data yang lainnya.

Menurut Draper and Smith (1992), adakalanya pencilan memberikan informasi yang tidak bisa diberikan oleh titik data lainnya. Berdasarkan Montgomery and Peck (1992), sebagai kaidah umum, pencilan baru ditolak jika setelah ditelusuri ternyata merupakan akibat dari kesalahan-kesalahan seperti memasukkan ukuran atau analisis yang salah, ketidaktepatan pencatatan data, dan terjadi kerusakan alat pengukuran. Bila ternyata bukan akibat dari kesalahan-kesalahan semacam itu, penyelidikan yang seksama harus dilakukan. Menghapus data tersebut untuk “memperbaiki persamaan yang cocok” dapat berbahaya, tindakan tersebut dapat menimbulkan kesalahan ketelitian dalam mengestimasi atau memprediksi.

2.6.1 Pengujian Pencilan dengan Uji *DFFITS*

Menurut Montgomery and Peck (1992), rumus $DFFITS_i$ sebagai berikut:

$$(DFFITS_i) = t_i \left(\frac{h_{ii}}{1-h_{ii}}\right)^{\frac{1}{2}} \tag{15}$$

dimana t_i adalah *R-student* (*studentized deleted residual*) untuk kasus ke-*i*, dengan

$$t_i = \frac{e_i}{\sqrt{S_{(i)}^2(1-h_{ii})}} \quad , \text{i} = 1, 2, \ldots, \text{n} \tag{16}$$

$$s_{(i)}^2 = \frac{(n-p)MS_E - e_i^2/(1-h_{ii})}{n-p-1} \tag{17}$$

$$MS_E = \frac{\boldsymbol{y'y} - \boldsymbol{\hat{\beta}'X'y}}{n-k-1} \tag{18}$$

dimana e_i adalah residual ke-i, h_{ii} adalah elemen-elemen diagonal dari matriks hat **H**. ***H*** adalah matriks n x n

$$\boldsymbol{H} = \boldsymbol{X}(\boldsymbol{X'X})^{-1}\boldsymbol{X'} \tag{19}$$

Disebut pencilan jika nilai $|DFFITS| > 1$ untuk gugus data yang berukuran kecil sampai sedang dan nilai $|DFFITS| > 2\sqrt{p/n}$ untuk gugus data yang berukuran besar, dengan $p = k + 1$, dan n adalah banyaknya observasi (Neter, 1997).

2.7 Fungsi Obyektif

Menurut Fox (2002), fungsi obyektif adalah fungsi yang digunakan untuk mencari fungsi pembobot pada regresi *robust*. Fungsi pembobot yang digunakan :

Tabel 1. Fungsi Obyektif dan Fungsi Pembobot untuk Estimasi Kuadrat Terkecil, Huber, dan Tukey Bisquare

Metode	Fungsi Obyektif	Fungsi Pembobot	Interval
Kuadrat Terkecil	$\rho(u_i) = \frac{1}{2}u_i^2$	$w(u_i) = 1$	$\lvert u_i \rvert < \infty$
Huber	$\rho(u_i) = \begin{cases} \frac{1}{2}u_i^2 \\ \lvert u_i \rvert c - \frac{1}{2}c^2 \end{cases}$	$w(u_i) = \begin{cases} 1 \\ \frac{c}{\lvert u_i \rvert} \end{cases}$	$\lvert u_i \rvert \le c$ $\lvert u_i \rvert > c$
Tukey Bisquare	$\rho(u_i) = \begin{cases} \frac{c^2}{6}\left\{1 - \left[1 - (\frac{u_i}{c})^2\right]^3\right\} \\ \frac{c^2}{6} \end{cases}$	$w(u_i) = \begin{cases} \left[1 - \left(\frac{u_i}{c}\right)^2\right]^2 \\ 0 \end{cases}$	$\lvert u_i \rvert \le c$ $\lvert u_i \rvert > c$

Dimana nilai *c* pada metode Huber dan Tukey Bisquare disebut *tunning constant*. Telah diketahui bahwa *tunning constan* untuk metode Tukey Bisquare adalah $c = 4.685$. Penulis menggunakan fungsi pembobot Tukey Bisquare dalam analisis regresi *robust* MM-*estimator*.

2.8 Regresi *Robust*

Menurut Chen (2002), regresi *robust* adalah sebuah alat yang penting untuk menganalisis data yang terkontaminasi pencilan. Tujuan utama regresi *robust* adalah untuk memberikan hasil yang stabil karena kehadiran pencilan.

2.8.1 *Robust* M-*Estimator*

Menurut Montgomery and Peck (1992), pada umumnya regresi *robust* M-*estimator* dilakukan dengan meminimumkan fungsi obyektif:

$$min_\beta \sum_{i=1}^{n} \rho(e_i) = min_\beta \sum_{i=1}^{n} \rho\left(y_i - \sum_{j=0}^{k} x_{ij}\beta_j\right) \quad (20)$$

Untuk mendapatkan skala invariant pada *estimator* ini, biasanya dengan menyelesaikan persamaan

$$min_\beta \sum_{i=1}^{n} \rho(u_i) = min_\beta \sum_{i=1}^{n} \rho(\tfrac{e_i}{s}) = min_\beta \sum_{i=1}^{n} \rho\left(\frac{y_i - \sum_{j=0}^{k} x_{ij}\beta_j}{s}\right) \quad (21)$$

dimana *s* adalah skala estimasi robust. Estimasi *s* yang sering digunakan adalah

$$s = median|e_i - median(e_i)|/0.6745 \quad (22)$$

Konstan 0.6745 membuat *s* mendekati estimator tak bias dari σ, dan error berdistribusi normal.

Estimasi Parameter β

Menurut Montgomery and Peck (1992), untuk meminimumkan persamaan (21), turunan parsial pertama dari ρ terhadap β_j $(j = 0,1,\dots,k)$ disamakan dengan 0, sehingga menghasilkan suatu kondisi perlu untuk minimum. Ini memberikan $p = k + 1$ sistem persamaan : $\frac{\partial \rho}{\partial \beta_j} = 0$

$$\sum_{i=1}^{n} x_{ij}\psi\left(\frac{y_i - \sum_{j=0}^{k} x_{ij}\beta_j}{s}\right) = 0, \qquad j = 0,1,\dots,k \quad (23)$$

dimana $\psi = \rho'$ dan x_{ij} adalah pengamatan ke-i pada parameter ke-j dan $x_{i0} = 1$. Didefinisikan fungsi pembobot:

$$w(u_i) = \frac{\psi\left(\frac{y_i - \sum_{j=0}^{k} x_{ij}\beta_j}{s}\right)}{\left(\frac{y_i - \sum_{j=0}^{k} x_{ij}\beta_j}{s}\right)} \quad (24)$$

Dan $w_i = w(u_i)$. Kemudian estimasi persamaan (23) dapat ditulis:

$$\sum_{i=1}^{n} x_{ij} w_i\left(y_i - \sum_{j=0}^{k} x_{ij}\beta_j\right) = 0 \qquad , j = 0, 1, \dots, k \quad (25)$$

Menurut Montgomery and Peck (1992), estimasi parameter pada regresi *robust* M-*estimator* dilakukan dengan estimasi *Iteratively Reweighted Least Square* (IRLS). Iterasi ini membutuhkan proses iterasi dimana nilai w_i akan berubah nilainya di setiap iterasi. Untuk menggunakan IRLS, dianggap $\hat{\beta}_0$ ada dan *s* adalah skala estimasi *robust*. Kemudian ditulis $p = k + 1$ sistem persamaan:

$$\sum_{i=1}^{n} x_{ij} w_i^{(0)} \left(y_i - \sum_{j=0}^{k} x_{ij} \hat{\beta}_j^{(0)} \right) = 0 \qquad , j = 0, 1, ..., \qquad (26)$$

Dalam bentuk matriks dapat ditulis:

$$\boldsymbol{X'} \ \boldsymbol{W^{(0)}} \ \boldsymbol{X} \quad \boldsymbol{\hat{\beta}_j} \quad = \quad \boldsymbol{X'} \ \boldsymbol{W^{(0)}} \quad \boldsymbol{y} \qquad (27)$$

dimana $\boldsymbol{W^{(0)}}$ adalah matriks diagonal berukuran n x n dengan elemen-elemen diagonal $w_1, w_2, \dots, w_n$. Jadi estimasi parameter regresi *robust* dengan IRLS, untuk *m+1* iterasi adalah:

$$\boldsymbol{\hat{\beta}_j^{(m+1)}} = (\boldsymbol{X'W^{(m)}X})^{-1}\boldsymbol{X'W^{(m)}y} \qquad (28)$$

Iterasi akan berhenti sampai didapatkan nilai $\boldsymbol{\hat{\beta}_j}$ yang konvergen yaitu selisih nilai $\boldsymbol{\hat{\beta}_j^{(m+1)}}$ dan $\boldsymbol{\hat{\beta}_j^{(m)}}$ mendekati 0.

2.8.2 *Robust* LTS (*Least Trimmed Square*) *-Estimator*

Menurut Willems and Aelst (2005), *estimator* LTS adalah meminimumkan fungsi objektif (jumlah kuadrat residual sebanyak *h*) $Q_{LTS} = \sum_{i=1}^{h} e_{(i:n)}^2$, dimana *h* dipilih antara $\frac{n}{2}$ dan n. *Estimator* LTS adalah:

$$\hat{\theta}_{Lts} = \min Q_{LTS} \qquad (29)$$

Keterangan, $\quad h = (\gamma . n) \ dan \ \gamma = (1 - \alpha)$

α adalah nilai breakdown (persentase banyaknya pencilan terhadap seluruh data (n)), e_i^2 adalah kuadrat residual yang diurutkan dari yang terkecil ke terbesar, dan n adalah jumlah pengamatan.

Pada LTS-*estimator* n-h diberikan bobot nol dalam mengestimasi parameter. LTS-*estimator* dapat untuk menangani nilai *breakdown* antara 0% sampai dengan 50% (Willems and Aelst, 2005).

2.8.3 *Robust* MM-*Estimator*

Menurut Chen (2002), kelas MM-*estimator* pertama kali dikenalkan oleh Yohai (1987) untuk regresi linier. Menurut Yohai (1987), MM-*estimator* didefinisikan melalui tiga langkah. Langkah pertama, estimasi parameter awal regresi dihitung dengan metode LTS-*estimator*. Tahap kedua, menghitung residual dan skala estimasi *robust* dengan menggunakan M-*estimator*. Terakhir, menghitung estimasi parameter akhir dengan M-*estimator*.

3. METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang akan dianalisis dalam skripsi ini berupa data bangkitan dari *software* Minitab 14.0.

3.2 Variabel Penelitian

Variabel yang digunakan adalah satu variabel tak bebas (y) dan empat variabel bebas (x_1, x_2, x_3, dan x_4).

3.3 Langkah-langkah Analisis

1. Menyiapkan data yang akan digunakan dalam penelitian
2. Dengan metode kuadrat terkecil didapatkan model regresi linier berganda, kemudian dilakukan uji kecocokan model (Uji-F)
3. Melakukan pendeteksian pencilan dengan $DFFITS_i$.
4. Estimasi parameter dengan regresi *robust* MM-*estimator*. Menurut Yohai (1987), prosedurnya adalah:
   a. Menghitung parameter $\hat{\beta}_j^0$ sebagai estimasi awal (inisial) regresi linier MM-*estimator*. $\hat{\beta}_j^0$ dihitung dengan menggunakan metode LTS-*estimator*. Menurut Willems and Aelst (2005), langkah-langkah perhitungan LTS-*estimator* sebagai berikut:
      1. Menggunakan $\hat{\beta}_j^0$ dari hasil estimasi parameter metode kuadrat terkecil sebagai parameter inisiasi LTS-*estimator*.
      2. Menentukan kuadrat residual (e_i^2), kemudian menghitung sejumlah $h_0 = (\gamma * n)$ pengamatan dengan nilai (e_i^2) terkecil. Menentukan $H_0 = \{h_0 \text{ observasi dengan } e_i^2 \text{ terkecil}\}$. Residual kuadrat yang terbesar $(n-h)$diberi bobot nol.
      3. Menghitung $Q_{LTS}^0 = \sum_{i=1}^{h_o} e_i^2$.
      4. Menghitung estimasi parameter $\hat{\beta}_{new}$ dari H_0 pengamatan.
      5. Menentukan kuadrat residual (e_i^2) yang bersesuaian dengan $(\hat{\beta}_{new})$, kemudian menghitung sejumlah h_{new} pengamatan dengan nilai (e_i^2) terkecil. Menentukan $H_{new} = \{\ h \text{ observasi dengan } e_i^2 \text{ terkecil}\}$. Residual kuadrat yang terbesar $(n-h)$ diberi bobot nol.
      6. Menghitung $Q_{new} = \sum_{i=1}^{h_{new}} e_i^2$.
      7. Ulangi langkah 4 sampai 6. Iterasi dihentikan jika nilai $Q_{LTS}^{m+1} \leq Q_{LTS}^m$ dan nilai parameternya sudah konvergen, artinya selisih $\hat{\beta}_j^{m+1}$ dan $\hat{\beta}_j^m$ mendekati 0.

   b. Menghitung residual e_i dan skala estimasi *robust* (s).
   c. Menghitung fungsi objektif Tukey Bisquare yaitu $\sum_{i=1}^n \rho_i^m(u)$.
   d. Menghitung fungsi pengaruh Tukey Bisquare.
   e. Menghitung fungsi pembobot Tukey Bisquare.
   f. Menghitung parameter regresi yang baru $(\hat{\beta}_j^m)$.
   g. Mengulang langkah (b) sampai (f) sehingga parameter regresi $\hat{\beta}_j$ konvergen, artinya selisih $\hat{\beta}_j^{m+1}$ dan $\hat{\beta}_j^m$ mendekati 0, maka iterasi dihentikan.

4. HASIL DAN PEMBAHASAN

4.1 Uji Kecocokan Model (Uji F)

Persamaan regresi linier berganda dengan pendekatan metode kuadrat terkecil adalah:

$$\hat{y} = 1.50 + 0.703\, x_1 - 0.767\, x_2 - 0.433\, x_3 + 1.06\, x_4$$

Kemudian dilakukan uji kecocokan model (Uji F). Didapatkan hasil F_{hitung} = 10.45 > $F_{(5\%,4,30)}$ = 2.689628, maka H_0 ditolak. Jadi model cocok, sehingga dapat dikatakan variabel bebas (x_1, x_2, x_3, dan x_4) secara bersama-sama berpengaruh terhadap variabel tak bebas.

4.2 Pendeteksian Pencilan

Karena jumlah data (n = 35) termasuk gugus data sedang, maka kriteria disebut pencilan jika nilai $|DFFITS_i| > 1$. Dari hasil di atas didapatkan data ke- 21, 27, dan 34 yang mempunyai nilai $|DFFITS_i|$ > 1, sehingga data tersebut merupakan pencilan.

4.3 Perhitungan Estimasi Parameter β dengan Regresi *Robust* MM-*Estimator*

Estimasi awal untuk mengestimasi parameter β dengan regresi *robust* MM-*Estimator* adalah menggunakan regresi *robust* LTS-*estimator*. Iterasi dihentikan jika nilai $Q_{LTS}^{m+1} \leq Q_{LTS}^{m}$ dan nilai parameternya sudah konvergen, artinya selisih $\hat{\beta}_j^{m+1}$ dan $\hat{\beta}_j^{m}$ mendekati 0. Iterasi dihentikan pada iterasi ke-4 karena nilai $Q_{LTS}^{4}(585.0823) \leq Q_{LTS}^{3}(585.0823)$ dan nilai parameternya konvergen, yaitu:

Tabel 2. Hasil Parameter Regresi *Robust* LTS-*Estimator*

$\hat{\beta}_j$	Iterasi ke-3	Iterasi ke-4	Hasil Selisih
$\hat{\beta}_0$	2.27	2.27	0
$\hat{\beta}_1$	0.675	0.675	0
$\hat{\beta}_2$	– 0.589	– 0.589	0
$\hat{\beta}_3$	– 0.348	– 0.348	0
$\hat{\beta}_4$	0.854	0.854	0

sehingga persamaan regresi :

$$\hat{y} = 2.27 + 0.675x_1 - 0.589x_2 - 0.348x_3 + 0.854x_4$$

Langkah selanjutnya menggunakan regresi *robust* M-*estimator* untuk mendapatkan penyelesaian dari regresi *robust* MM-*estimator*. Iterasi dihentikan jika nilai estimasi $\hat{\beta}_j$ konvergen. Iterasi dihentikan pada iterasi ke-27 karena $\hat{\beta}$ sudah konvergen artinya selisih $\hat{\beta}_j$ mendekati 0 yaitu:

Tabel 3. Hasil Parameter Regresi *Robust* MM-*Estimator*

$\hat{\beta}_j$	Iterasi ke-26	Iterasi ke-27	Hasil Selisih
$\hat{\beta}_0$	3.97	3.97	0
$\hat{\beta}_1$	0.392	0.392	0
$\hat{\beta}_2$	-0.810	-0.810	0
$\hat{\beta}_3$	-0.263	-0.263	0
$\hat{\beta}_4$	0.968	0.968	0

Didapatkan persamaan regresi linier berganda dengan menggunakan estimasi regresi *robust* MM-*estimator* :

$$\hat{y} = 3.97 + 0.392x_1 - 0.810x_2 - 0.263x_3 + 0.968x_4$$

5. PENUTUP

1. Pendeteksian pencilan pada data bangkitan yang dilakukan dengan menggunakan *DFFITS* menunjukkan bahwa data ke-21, 27, dan 34 sebagai pencilan.
2. Persamaan regresi linier berganda yang mengandung pencilan dengan menggunakan estimasi regresi *robust* MM-*estimator* yang diperoleh dalam tugas akhir ini adalah:

$$\hat{y} = 3.97 + 0.392x_1 - 0.810x_2 - 0.263x_3 + 0.968x_4$$

6. DAFTAR PUSTAKA

Chen, C. 2002. *Robust Regression and Outlier Detection with the ROBUSTREG Procedure*. Paper 265-27. North Carolina: SAS Institute.

Draper, N. R., Smith, H. 1992. *Analisis Regresi Terapan,* Edisi kedua. Ir. Bambang Sumantri, penerjemah. Jakarta: Gramedia Pustaka Utama. Terjemahan dari: *Applied Regression Analysis,* 2nd edition.

Fox, J. 2002. *Robust Regression.* Appendix to An R and S-Plus Companion to Applied Regression.

Greene, W. H. 1951. *Econometric Analysis*, 5th edition. New York: Pearson Education.

Gujarati, D. 1997. *Ekonometrika Dasar*. Dra. Ak. Sumarno Zain, MBA, penerjemah. Jakarta: Erlangga. Terjemahan dari: *Basic Econometrics.*

Gujarati, D. 2007. *Dasar-dasar Ekonometrika*, Edisi ketiga, jilid 1. Julius A. Mulyadi, S.E, penerjemah. Jakarta: Erlangga. Terjemahan dari: *Essentials of Econometrics*.

Montgomery, D. C., Peck, E. A. 1992. *Introduction to Linear Regression Analysis*, 2nd edition. New York: John Wiley & Sons, Inc.

Neter. J., Wasserman, W., Kutner M. H. 1997. *Model Linear Terapan*. Bambang Sumantri, penerjemah. Jurusan Statistika FMIPA-IPB. Terjemahan dari: *Applied Linear Model.*

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

Willems, G., Aelst, S. V. 2005. *Fast and Robust Bootstrap for LTS*. Computational Statistics & Data Analysis 48 (2005) 703-715.

Yohai, V. J. 1987. *High Breakdown-Point And High Efficiency Robust Estimates For Regression.* The Annals of Statistics. Vol. 15, No. 20, 642-656.