Metode Regresi Robust Dengan Estimasi-M pada Regresi Linier Berganda (Studi Kasus : Indeks Harga Konsumen Kota Tarakan)

Robust Regression Method to m-estimation on Multiple Linear Regression (Case Study: Consumer Price Index Tarakan City)

Al Ghazali[1], Desi Yuniarti[1], Memi Nor Hayati[1]
[1]Program Studi Statistika FMIPA Universitas Mulawarman
Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Mulawarman
E-mail: al_ghazali@ymail.com

Abstract

Ordinary Least Square is one method to find the values of the parameters estimated on regression. One of the robust regression estimation that mostly used to find the estimate is the M-estimation which was introduced by Huber. M-*estimation method is similar to the OLS, the difference is only in giving the same weighting. From the analysis results obtained with the robust regression equation m-estimation is:* $\hat{Y} = 21{,}965 + 0{,}420\,X_1 + 0{,}189\,X_2 + 0{,}199\,X_3$. *Based on the partial significance testing can be concluded that all the independent variables (groceries (X_1), clothing (X_2), and education (X_3)) CPI effect on the dependent variable (Y).*

Keywords: Ordinary Least Squares, robust regression, outliers, Different fitted FITS value, m-estimation.

Pendahuluan

Istilah regresi pertama kali diperkenalkan oleh Sir Francis Galton pada tahun 1877 dalam penelitian biogenisisnya. Regresi berguna dalam menelaah hubungan sepasang variabel atau lebih. Salah satu cara untuk mencari estimasi koefisien parameter dengan dengan menggunakan Metode Kuadrat Terkecil atau *Ordinary Least Square* (OLS) (Sembiring, 1995).

Analisis regresi merupakan analisis statistik yang bertujuan untuk memodelkan hubungan antara variabel terikat dengan variabel bebas. Model regresi yang baik memerlukan data yang baik pula. Suatu data dikatakan baik apabila data tersebut berada di sekitar garis regresi. Kenyataannya, terkadang terdapat data yang terletak jauh dari garis regresi atau pola data keseluruhan. Data tersebut dikenal dengan istilah pencilan atau *outlier*. Pencilan merupakan suatu keganjilan dan menandakan suatu titik data yang sama sekali tidak tipikal dibanding data lainnya (Draper dan Smith,1992).

Salah satu metode untuk mengatasi pencilan adalah regresi *robust*. Regresi *robust* merupakan metode regresi yang digunakan ketika distribusi dari residual tidak normal dan atau mengandung beberapa pencilan yang berpengaruh pada model (Ryan, 1997).

Gunandi (2011) pernah melakukan penelitian tentang regresi *robust* dengan judul, Regresi *Robus*t dengan Metode M-Estimator dan Aplikasinya pada Regresi Linier Sederhana. Oleh karena itu, penulis melakukan penelitian regresi *robust* dengan estimasi-M dan aplikasinya pada regresi linier berganda.

Berdasarkan latar belakang tersebut, penulis tertarik memilih judul “Metode Regresi *Robust* dengan Estimasi-M pada Regresi Linier Berganda (Studi Kasus: Data Indeks Harga Konsumen Kota Tarakan)”.

Regresi Linier Berganda

Regresi berganda adalah regresi dengan dua atau lebih variabel $X_1, X_2, X_3....X_k$ sebagai variabel bebas dan variabel Y sebagai variabel tak bebas, Nilai-nilai koefisien atau taksiran parameter regresi berganda dapat diperoleh dengan metode OLS Model regresi linier berganda dapat dilihat pada persamaan :

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + ... + \beta_k X_{ik} + \varepsilon_i \qquad (1)$$

Estimasi Parameter Model Regresi Berganda

Pada regresi berganda untuk k variabel bebas penaksiran dari β dinyatakan dengan $\hat{\beta}$. Menurut Metode OLS penaksiran tersebut dapat diperoleh dengan meminimumkan bentuk kuadrat. Sehingga estimasi OLS untuk $\hat{\beta}$ adalah (Gujarati, 1999) :

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^{\mathbf{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathbf{T}}\mathbf{y} \qquad (2)$$

Pengujian Signifikansi Parameter

Pengujian signifikansi parameter pada regresi linier berganda (Sudjana, 2002)

1. Pengujian secara Simultan (Uji-*F*)
2. Pengujian secara Parsial (Uji-*t*)

Asumsi-asumsi Pada Regresi Linier Berganda

Asumsi utama yang mendasari pendugaa koefisien regresi dengan menggunakan metode OLS adalah (Widarjono, 2007):

1. Normalitas Residual
2. Non multikolineritas
3. Non Autokorelasi
4. Heteroskedastisitas

Pencilan

Menurut Montgomery dan 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.

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

1. Residual yang besar dari model yang terbentuk atau $E(e) \neq 0$
2. Varians pada data tersebut menjadi lebih besar.
3. Taksiran interval memiliki rentang yang lebar.

Pendeteksian pencilan dengan Menggunakan *scatter plot* metode ini yaitu mudah dipahami karena menampilkan data secara grafis (gambar) dan tanpa melibatkan perhitungan yang rumit. Selain menggunakan *scatter plot* pendeteksian juga menggunakan uji *Different fitted value FITS* (*DfFITS*).

Hipotesis yang digunakan dalam pengujian *DfFITS* adalah

H_0 : Pencilan ke – i tidak berpengaruh, i=0,1,2,3...n

H_1 : Pencilan ke – i berpengaruh

Diasumsikan bahwa $f_i = Y_i - \hat{Y}_i$ untuk i=0,1,2,3...n dengan $s(f_i)$ sebagai galat baku, maka *DfFITS* didefinisikan sebagai berikut:

$$DfFITS = \frac{f_i}{s(f_i)} = e_i\left[\frac{n-k-1}{JKG(1-h_{ii})-e_i^2}\right]^{\frac{1}{2}}\left[\frac{h_{ii}}{1-h_{ii}}\right]^{\frac{1}{2}} \quad (3)$$

Nilai h_{ii} didapatkan dengan rumus (Montgomery dan Peck,1982) :

$$\mathbf{H} = \mathbf{X}(\mathbf{X}^{\mathbf{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathbf{T}} \quad (4)$$

Regresi *Robust*

Regresi *Robust* merupakan metode yang digunakan ketika ada beberapa pencilan pada model. Metode ini merupakan alat penting untuk menganalisis data yang dipengaruhi oleh pencilan sehingga dihasilkan model *robust* atau kekar terhadap pencilan. Suatu estimati yang *robust* adalah relatif tidak terpengaruh oleh perubahan kecil pada bagian besar data (Ryan,1997).

Regresi *Robust* Estimasi-M

Salah satu estimasi regresi *robust* yang paling penting dan paling luas digunakan adalah estimasi-M yang diperkenalkan oleh Huber. Pada prinsipnya estimasi-M merupakan estimasi yang meminimumkan suatu fungsi objektif ρ

$$\min_{\beta}\sum_{i=1}^{n}\rho(e_i) = \min_{\beta}\sum_{i=1}^{n}\rho\left(y_i - \sum_{j=0}^{k}X_{ij}\beta_j\right) \quad (5)$$

$$\sum_{i=1}^{n}\rho(u_i) = \sum_{i=1}^{n}\rho\left(\frac{e_i}{\hat{\sigma}}\right) \quad (6)$$

dimana $\rho(u_i)$ adalah fungsi simetris dari residual atau fungsi yang memberikan kontribusi pada masing-masing residual pada fungsi objektif. Pada umumnya suatu estimasi skala *robust* perlu diestimasi dan $\hat{\sigma}$ adalah skala estimasi *robust*. Untuk mencari $\hat{\sigma}$ pada regresi *robust* dengan estimasi-M yang sering digunakan persamaan :

$$\hat{\sigma}_l = \frac{median|e_i - median(e_i)|}{0{,}6745} \quad l = 1,2,3...n \quad (7)$$

Penyelesaian Koefiesien Regresi *Robust* Estimasi-M

Untuk meminimumkan fungsi objektif ρ turunan parsial pertama dari ρ terhadap β_j dimana j = 0,1,..,k, harus disama dengankan 0. Sehingga akan menghasilkan suatu syarat perlu untuk minimum. Ini menghasilkan sistem persamaan (Huber, 1981):

$$\frac{\partial\rho}{\partial\beta} = 0 \quad (8)$$

$$\sum_{i=1}^{n}X_{ij}\psi\left(\frac{e_i}{\hat{\sigma}}\right) = \sum_{i=1}^{n}X_{ij}\psi\left(\frac{y_i - \sum_{j=0}^{k}X_{ij}\beta_j}{\hat{\sigma}}\right) = 0 \quad (9)$$

dimana $\psi = \rho'$ dan X_{ji} adalah observasi ke-i pada parameter ke-j dan $X_{i0} = 1$ didefinisikan fungsi pembobot:

$$W(u_i) = \frac{\psi\left(\frac{e_i}{\hat{\sigma}}\right)}{\left(\frac{e_i}{\hat{\sigma}}\right)} \quad (10)$$

Menurut Fox (2002) fungsi obyektif digunakan untuk mendapatkan nilai fungsi pembobot pada regresi *robust*. Fungsi pembobot pada regresi *robust* estimasi-M yang sering digunakan adalah fungsi pembobot Huber. Kriteria fungsi pembobot Huber sebagai berikut:

$$W(u_i) = \begin{cases} 1, & |u_i| \le c \\ \dfrac{c}{u_i}, & |u_i| > c \end{cases} \qquad (11)$$

u_i merupakan residual ke-*i*, sedangkan nilai *c* dinyatakan dengan *tuning constant*. *Tuning constant* dalam regresi *robust* menentukan *kerobustan* penaksir terhadap pencilan dan efisiensi penaksir dalam ketidakadaan pencilan. Jika diambil α = 5%, maka estimasi-M Huber akan efektif digunakan nilai *c* = 1,345 (Fox, 2002).

Menurut Montgomery dan Peck (1992), estimasi parameter pada regresi *robust* estimasi-M 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 $\hat{\sigma}_l$ adalah skala estimlasi *robust*. Kemudian *p* = 1+*k* ditulis sistem persamaan :

$$\sum_{i=1}^{n} X_{ij} W_i \left(y_i - \sum_{j=0}^{k} X_{ij}\beta_j \right) = 0 \qquad (12)$$

W adalah matriks diagonal berukuran *n* x *n* dengan elemen-elemen diagonalnya $W_{1,1}, W_{1,2}, W_{1,3} \ldots W_{1,n}$. Jadi estimasi parameter regresi *robust* dengan IRLS, untuk *l+1* iterasi adalah:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^{\mathbf{T}}\mathbf{W}^{\mathbf{l}}\mathbf{X})^{-1}\mathbf{X}^{\mathbf{T}}\mathbf{W}^{\mathbf{l}}\mathbf{y} \qquad l = 0,1,2,3... \qquad (13)$$

Indeks Harga Konsumen

Indeks Harga Konsumen (IHK) merupakan salah satu indikator ekonomi penting yang dapat memberikan informasi mengenai perkembangan harga barang dan jasa yang dibayar oleh konsumen atau masyarakat, khususnya masyarakat perkotaan. IHK mengukur perubahan pengeluaran/biaya barang dan jasa (komoditas) yang biasa dibeli oleh mayoritas rumah tangga dari waktu ke waktu. Dengan kualitas dan kuantitas paket komoditas yang dianggap konstan pada tahun dasar. Indeks tersebut semata-mata mencerminkan perubahan harga dan didesain sebagai ukuran dari dampak perubahan harga pada pembelian barang dan jasa (BPS, 2014).

Metode Penelitian

Variabel penelitian yang digunakan ada 4 yaitu variabel *Y* merupakan data indeks harga konsumen kota tarakan, sedangkan variabel X_1 (bahan makanan), X_2 (sandang) dan X_3 (pendidikan).

Adapun teknik analisis data dalam penelitian ini adalah:

1) Analisis Deskriptif
2) Penentuan variabel dan persamaan model regresi awal
3) Estimasi parameter model untuk data IHK Kota Tarakan 2010-214
4) Pengujian signifikansi parameter
5) Persamaan model Regresi yang terbaik
6) Pengujian Asumsi analisis regresi linier berganda
7) Pendekteksian pencilan
8) Permodelan regresi linier berganda menggunakan metode regresi *robust* dengan estimasi-m

Hasil dan Pembahasan

Hasil analisis deskriptif pada data IHK Kota Tarakan tahun 2010-2014.

Tabel 1. Analisis Deskriptif

Variabel	Rataan	Standar Deviasi	Variansi
IHK	145,54	18,3	334,87
Bahan Makanan	170,96	29,8	888,57
Sandang	130,25	12,3	151,09
Pendidikan	136,15	21,71	471,64

Berdasarkan Tabel 1 analisis deskriptif data IHK didapatkan hasil bahwa rata-rata nilai harga indeks untuk IHK sebesar 145,54, rata-rata harga indeks bahan makanan, sandang dan pendidikan yang dikonsumsi oleh konsumen masing-masing sebesar 170,96, 130,25, dan 136,15. Standar deviasi untuk IHK sebesar 18,30. Indeks harga Bahan makanan memiliki standar deviasi sebesar 29,80, indeks harga sandang memiliki standar deviasi sebesar 12,30 dan indeks harga pendidikan memiliki standar deviasi sebesar 21,71. IHK, indeks harga bahan makanan, indeks harga sandang dan indeks harga pendidikan memiliki variansi masing-masing sebesar 334,87, 888,57, 151,09 dan 471,64.

Diperoleh estimasi model regresi untuk data IHK Kota Tarakan Tahun 2010-2014:

$$\hat{Y} = 18{,}428 + 0{,}398\,X_1 + 0{,}247\,X_2 + 0{,}198\,X_3$$

Sebelum dianalisis lebih lanjut maka dilakukan pengujian secara simultan (uji *F*), pengujian variabel bebas secara parsial (uji *t*), dan dilanjutkan dengan pengujian asumsi.

Pengujian signifikansi secara simultan (uji-*F*), untuk mengetaui ada tidaknya pengaruh variabel bebas terhadap variabel terikat.

Hipotesis:

H_0: $\beta_0 = \beta_1 = \beta_2 = \beta_3 = 0$

H_1: minimal ada satu $\beta_j \neq 0$ dimana j =0,1,2,3

Dari hasil diperoleh nilai *p-value* = 0,000 < α = 0,05, dengan demikian menolak H_0, artinya minimal ada satu β_j yang berpengaruh terhadap IHK.

Pengujian signifikansi secara parsial (uji-*t*) bertujuan untuk melihat ada tidaknya pengaruh secara individual variabel bebas

Tabel 2. Hasil Pengujian Parameter Secara Parsial

Paramater	***Estimate***	***Standard Error* (SE)**	***p-value***
Konstanta	18,428	3,417	0
Bahan Makanan	0,398	0,019	0
Sandang	0,247	0,054	0
Pendidikan	0,198	0,017	0

Hipotesis:

$H_0 : \beta_j = 0$ untuk j = 0,1,2,3

$H_1 : \beta_j \neq 0$ untuk j = 0,1,2,3

Dari Tabel 2 diperoleh bahwa nilai *p-value* untuk masing-masing parameter adalah sebesar 0,000 < α = 0,05 maka dapat diambil keputusan bahwa menolak H_0 dan dapat disimpulkan bahwa konstanta, bahan makanan, sandang dan pendidikan berpengaruh terhadap IHK.

Berdasarkan uji signifikansi parameter yang berpengaruh, model persamaaan regresi merupakan persamaan model regresi yang terbaik.

$$\hat{Y} = 18{,}428 + 0{,}398\,X_1 + 0{,}247\,X_2 + 0{,}198\,X_3$$

Pengujian asumsi pada regresi linier berganda. Normalitas residual dengan hipotesis sebagai berikut:

H_0 : Residual berdistribusi normal

H_1 : Residual tidak berdistribusi normal

Statistik uji yang digunakan adalah *Kolmogorov-Smirnov*, diperoleh nilai *p-value* statistik uji *Kolmogorov-Smirnov* adalah sebesar 0,02. Nilai *p-value* untuk *Kolmogorov-Smirnov* sebesar 0,02 < α = 0,05, dengan demikian keputusannya H_0 ditolak. Dapat disimpulkan bahwa data residual tidak berdistribusi normal.

Pendeteksian Multikolinieritas dilakukan dengan melihat nilai *Variance Inflation Factor* (VIF).

Tabel 3. Nilai VIF

No	**Variabel**	**VIF**
1	Bahan Makanan	6,9
2	Sandang	9,3
3	Pendidikan	3

Berdasarkan Tabel 3 disimpulkan bahwa tidak terdapat masalah multikolinieritas antara variabel bahan makanan, sandang dan pendidikan dikarenakan nilai VIF yang diperoleh lebih kecil dari 10.

Pengujian Autokorelasi dilakukan dengan menggunakan uji Durbin-Watson (DW) dengan prosedur:

Hipotesis:

H_0 : Tidak terdapat masalah autokorelasi dalam model regresi

H_1 : Terdapat masalah autokorelasi dalam model regresi

Statistik Uji:

Statistik uji yang digunakan adalah Uji DW dan diperoleh nilai DW adalah sebesar 1,730.

Kriteria Penolakan:

H_0 ditolak jika nilai 0< DW < d_L

H_0 diterima jika nilai d_u< DW < 4-d_u

Kesimpulan:

Nilai DW = 1,784, nilai d_L = 1,4064 dan d_U = 1,6708

Dilihat pada tabel Durbin-Watson berdasarkan nilai $k = 3$, $n = 60$ dan $\alpha = 0{,}05$. Sehingga d_u (1,4064) < DW(1,730) <4 – d_u (2,3292) maka dapat disimpulkan menerima H_0, jadi tidak terdapat masalah autokorelasi dalam model regresi.

Uji Heteroskedastisitas, adapun prosedurnya pengujiannya, hipotesis:

H_0 : Tidak terdapat masalah heteroskedastsitas dalam model regresi

H_1 : Terdapat masalah heteroskedastsitas dalam model regresi

Statistik uji yang digunakan adalah Uji *Glejser*. Nilai *p-value* untuk harga indeks bahan makanan (X_1), sandang (X_2) dan pendidikan (X_3) sebesar adalah 0,149 > 0,05. Dengan demikian, dapat disimpulkan menerima H_0 atau tidak terdapat masalah heteroskedastisitas pada model regresi.

Pendeteksian pencilan dengan menggunakan *scatter plot*, didapatkan hasil *scatter plot* antara IHK (*Y*) dengan nilai residual, yang dapat dilihat pada Gambar 1.

Pendeteksian pencilan menggunakan *DfFITS*.

Hipotesis pengujian *DfFITS* :

H_0 : Pencilan ke – *i* tidak berpengaruh, dimana *i*=1,2,3... 60

H_1 : Pencilan ke – *i* berpengaruh

Nilai kritis untuk pengujian *DfFITS*:

$$2\sqrt{\frac{k+1}{n}} = 2\sqrt{\frac{3+1}{60}} = 0{,}50$$

Adapun kriteria pengujian yang melandasi keputusan adalah:

$$DfFITS = \begin{cases} \leq 0{,}50, & H_0\ diterima \\ < 0{,}50, & H_0\ ditolak \end{cases}$$

Hasil perhitungan *DfFITS* 7, ke-1 sebagai sebagai berikut :

$$DfFITS = e_i\left[\frac{n-k-1}{JKG(1-h_{11}) - e_i^2}\right]^{\frac{1}{2}}\left[\frac{h_{11}}{1-h_{11}}\right]^{\frac{1}{2}}$$

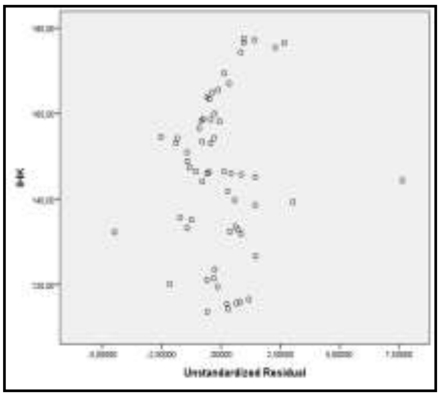
DfFITS= -0,25.

Tabel 4. Hasil Lengkap *DfFITS*

No	*DfFITS*	*|DfFITS|*	No	*DfFITS*	*|DfFITS|*
1	-0,2506	0,25063	31	-0,0969	0,09689
2	-0,3426	0,34264	32	-0,081	0,08102
3	-0,24	0,24	33	-0,0999	0,09986
4	0,06185	0,06185	34	-0,0619	0,06189
5	0,14313	0,14313	35	-0,0086	0,00856
6	0,12665	0,12665	36	-0,0324	0,03239
7	0,10453	0,10453	37	-0,0577	0,05774
8	-0,9027	**0,90271**	38	-0,0731	0,07308
9	0,12203	0,12203	39	-0,0505	0,05047
10	0,27772	0,27772	40	-0,0172	0,0172
11	0,54757	**0,54757**	41	0,05729	0,05729
12	0,03269	0,03269	42	0,01933	0,01933
13	-0,0966	0,09659	43	0,18329	0,18329
14	-0,0947	0,0947	44	0,30589	0,30589
15	-0,0735	0,07348	45	0,19051	0,19051
16	0,04938	0,04938	46	0,17219	0,17219
17	1,93056	**1,93056**	47	0,36498	0,36498
18	0,15475	0,15475	48	0,43606	0,43606
19	0,09396	0,09396	49	-0,1231	0,1231
20	0,0212	0,0212	50	0,06384	0,06384
21	-0,2468	0,24678	51	0,05279	0,05279
22	-0,2511	0,25105	52	0,13791	0,13791
23	-0,24	0,23999	53	0,17369	0,17369
24	-0,1885	0,18847	54	0,28472	0,28472
25	-0,1934	0,19341	55	-0,0377	0,03769
26	-0,257	0,25703	56	-0,3603	0,36034
27	-0,1945	0,19451	57	-0,1051	0,10511
28	-0,0896	0,08958	58	-0,0601	0,06006
29	-0,0528	0,05282	59	-0,0429	0,04288
30	-0,0357	0,03573	60	0,22577	0,22577

Berdasarkan Tabel 4. data pengamatan ke-5,11 dan 17 dengan nilai pencilan sebesar 0,90271, 0,54757, 1,93056 memiliki nilai yang lebih besar

dari kriteria uji sebesar 0,50. Maka menolak H_0 sehingga disimpulkan data ke-5, 11 dan 17 merupakan data pencilan yang berpengaruh terhadap model regresi. Apabila terdapat pencilan pada data maka akan dilanjutkan analisis dengan metode regresi *robust* dengan estimasi-M.



Gambar 1. *Scatter Plot* Antar IHK dengan Residual

Untuk mendapatkan hasil regresi *Robust* penduga estimasi-M, maka pengerjaannya dilakukan dengan IRLS. Adapun langkah-langkah pengerjaannya sebagai berikut:

1. Menghitung parameter regresi awal dengan menggunakan metode OLS sehingga didapatkan nilai β_j dan mendapatkan nilai $e_i = Y_i - \hat{Y}_i$ untuk mencari nilai pembobot awal. Didapatkan model persamaan regresi awal dengan metode OLS.
$$\hat{Y} = 18{,}428 + 0{,}398\,X_1 + 0{,}247\,X_2 + 0{,}198\,X_3$$
2. Mencari nilai $\hat{\sigma}_l$
$$\hat{\sigma}_1 = \frac{median|e_i - median(e_i)|}{0{,}6745}$$
$$\hat{\sigma}_1 = \frac{0{,}7953}{0{,}6745} = 1{,}1790$$
3. Menghitung nilai u_i
$$u_1 = \frac{e_1}{\hat{\sigma}_1} = \frac{-1{,}4249}{1{,}1790} = -1{,}2084$$
$$\vdots$$
$$u_{60} = \frac{e_{60}}{\hat{\sigma}_1} = \frac{1{,}45205}{1{,}1790} = 1{,}2314$$
4. Mencari nilai pembobot $W(u_i)$ sebagai nilai pembobot awal dengan dengan menggunakan kriteria fungsi pembobot Hubber.
$$W_{1,1} = 1$$
$$\vdots$$
$$W_{1,60} = 1$$
5. Mencari nilai parameter regresi *robust* dengan estimasi-M
$$\hat{Y} = 21{,}543 + 0{,}417\,X_1 + 0{,}194\,X_2 + 0{,}201\,X_3$$

Model persamaan regresi *robust* diatas merupakan model regresi *robust* untuk iterasi-1. Pada iterasi-1 belum diperoleh nilai yang konvergen. Iterasi akan berhenti sampai didapatkan nilai $\hat{\beta}_j$ yang konvergen yaitu selisih nilai $\hat{\beta}_j^{l+1}$ dan $\hat{\beta}_j^{l}$ mendekati 0.

Tabel 6. Persamaan Model Regresi

Iterasi	Persamaan Model Regresi
OLS	$\hat{Y} = 18{,}428 + 0{,}398X_1 + 0{,}247X_2 + 0{,}198X_3$
1	$\hat{Y} = 21{,}543 + 0{,}417X_1 + 0{,}194X_2 + 0{,}201X_3$
2	$\hat{Y} = 21{,}941 + 0{,}419X_1 + 0{,}189X_2 + 0{,}200X_3$
3	$\hat{Y} = 21{,}881 + 0{,}419X_1 + 0{,}190X_2 + 0{,}200X_3$
4	$\hat{Y} = 21{,}959 + 0{,}420X_1 + 0{,}189X_2 + 0{,}199X_3$
5	$\hat{Y} = 21{,}963 + 0{,}420X_1 + 0{,}189X_2 + 0{,}199X_3$
6	$\hat{Y} = 21{,}964 + 0{,}420X_1 + 0{,}189X_2 + 0{,}199X_3$
7	$\hat{Y} = 21{,}965 + 0{,}420X_1 + 0{,}189X_2 + 0{,}199X_3$
8	$\hat{Y} = 21{,}965 + 0{,}420X_1 + 0{,}189X_2 + 0{,}199X_3$

Berdasarkan Tabel 4.6 didapatkan hasil iterasi yang konvergen, yaitu iterasi ke-7 dan iterasi ke-8 mendekati atau sama dengan 0. Jadi persamaaan model regresi *robust* dengan estimasi-M pada iterasi ke-7, dengan persamaan model regresi *robust* sebagai berikut :

$$\hat{Y} = 21{,}965 + 0{,}420\,X_1 + 0{,}189\,X_2 + 0{,}199\,X_3$$

Interpretasi pada persamaan adalah jika bahan makanan (X_1), sandang (X_2), dan pendidikan (X_3) sama dengan 0, maka harga rata-rata IHK kota Tarakan sebesar 21,965. Setiap penambahan satuan indeks harga bahan makanan (X_1) akan meningkatkan harga rata-rata IHK kota Tarakan sebesar 0,420, apabila sandang (X_2), dan pendidikan (X_3) tetap. Setiap penambahan satuan harga indeks sandang (X_2) akan meningkatkan harga rata-rata IHK kota Tarakan sebesar 0,189 apabila bahan makanan (X_1), dan pendidikan (X_3) tetap. Setiap penambahan satuan harga indeks pendidikan (X_3) akan meningkatkan harga rata-rata IHK kota Tarakan sebesar 0,199 apabila makanan (X_1), dan sandang (X_2) tetap.

Berdasarkan hasil pengujian normalitas residual menggunakan metode regresi robust diperoleh bahwa residual telah berdistribusi normal. Dapat disimpulkan bahwa data residual berdistribusi normal.

Diketahui bahwa nilai R^2 adalah sebesar 0,997. Nilai ini menunjukkan bahwa 99,6 % variasi IHK dipengaruhi oleh variasi bahan makanan, sandang dan pendidikan. Sedangkan 0,4% dipengaruhi oleh faktor lain yang tidak dimasukkan dalam model.

Terdapat perbedaan nilai R^2 pada regresi metode OLS dengan regresi metode *robust*. Pada metode OLS nilai R^2 sebesar 99,2% sedangkan pada metode regresi *robust* nilai R^2 sebesar 99,6%.

Kesimpulan

Berdasarkan hasil penelitian dan pembahasan, maka diperoleh kesimpulan sebagai berikut:

1. Pada data Indeks Harga Konsumen (IHK) kota Tarakan pada tahun 2010-2014 terdapat satu pencilan yaitu, data pada pengamatan ke-5,11 dan 17.
2. Model analisis regresi *robust* dengan estimasi-M pada data IHK kota Tarakan tahun 2010-2014 adalah :

$$\hat{Y} = 21{,}965 + 0{,}420\,X_1 + 0{,}189\,X_2 + 0{,}199\,X_3$$

3. Faktor-faktor yang mempengaruhi data Indeks Harga Konsumen (IHK) Kota Tarakan tahun 2010-2014 yaitu bahan makanan, sandang dan pendidikan.

Daftar Pustaka

Badan Pusat Statistik. 2010-2014. Publikasi *Resmi Badan Pusat Statistik Provinsi Kalimantan Timur*.

Draper, N dan H. Smith. 1992. *Analisis Regresi Terapan, Terjemahan Edisi Kedua*. Jakarta: PT. Gramedia Pustaka Utama.

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

Gujarati, D. 1999. *Ekonometrika Dasar*. Jakarta: Erlangga.

Gunandi, M. 2011. *Regresi Robust dengan Metode M-Estimator dan Aplikasinya pada Regresi Linier Sederhana*: Skripsi (S1), FMIPA Universitas Mulawarman.

Huber, P.J. 1981. *Robust Statistic.* Canada: John Wiley & Sons Inc.

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

Ryan, T.P. 1997. *Modern Regression Methods*. A Wiley-Interscience Publication: New York.

Sembiring, R.K., 1995. *AnalisisRegresi*. Bandung: ITB.

Soermatini. 2007. *Pencilan (Outlier)*. Jatinanggor: Universitas Padjajaran.

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

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