

## **KAJIAN ESTIMASI-M IRLS MENGGUNAKAN FUNGSI PEMBOBOT HUBER DAN BISQUARE TUKEY PADA DATA KETAHANAN PANGAN DI JAWA TENGAH**

**Elen Dwi Pradewi<sup>1</sup>, Sudarno<sup>2</sup>**

<sup>1</sup>Alumni Program Studi Statistika FSM Universitas Diponegoro

<sup>2</sup>Staf Pengajar Program Studi Statistika FSM Universitas Diponegoro

### **Abstract**

Ordinary Least Squares (OLS) is one method of parameter estimation in regression analysis. However, the presence of outliers can cause estimation of regression coefficients obtained are not exact. Act of throwing away an outlier is not a wise move, because sometimes outliers provide significant information. Therefore, robust regression methods are needed to data contain outliers. This paper will use robust regression estimation method by M-estimation. This estimation use Iteratively Reweighted Least Squares (IRLS) method with weighting function by Huber and Tukey Bisquare. IRLS is applied to the case of food security in Central Java in 2007 that is influenced by the stock of rice, harvested area, average production, price of rice and the amount of consumption. The purpose of this writing is to compare goodness of M-estimation IRLS using Huber and Tukey Bisquare function in estimating the model parameters of food security in Central Java in 2007. Based on the research results can be concluded that the M-estimation by the Tukey Bisquare is better recommended than Huber function. This can be seen by value results of Mean Square Error and determination coefficient.

**Keywords:** Robust Regression, Iteratively Reweighted Least Squares, Outlier, Huber Weighting Function, Tukey Bisquare Weighting Function

### **1. Pendahuluan**

Analisis regresi merupakan salah satu teknik analisis statistika yang paling banyak digunakan. Analisis regresi baik yang linear maupun yang nonlinear. Pada kejadian sehari-hari terdapat hubungan sebab akibat yang muncul, baik yang terjadi pada bidang sains, sosial, industri maupun bisnis. Kejadian-kejadian tersebut dapat dimodelkan dalam bentuk fungsi regresi. Secara umum, analisis regresi berkenaan dengan studi ketergantungan suatu variabel dependen (tak bebas) pada satu atau lebih variabel independen (bebas), dengan maksud ketergantungan model itu dapat dipergunakan sebagai alat prediksi kejadian untuk waktu yang akan datang<sup>[1,3]</sup>.

Salah satu tujuan dalam analisis regresi adalah mengestimasi koefisien regresi dalam model regresi. Model regresi merupakan suatu cara formal untuk mengekspresikan dua unsur penting suatu hubungan statistik, yaitu kecenderungan berubahnya variabel tak bebas secara sistematis sejalan dengan berubahnya variabel bebas dan berpencarnya titik-titik di sekitar kurva taksiran model itu. Metode yang biasa digunakan untuk mengestimasi koefisien regresi yaitu metode kuadrat terkecil. Namun, metode ini mensyaratkan bahwa distribusi data harus memenuhi asumsi klasik dari regresi, yaitu linear dalam parameter, galat berdistribusi normal dengan rata-rata nol dan ragam konstan, antar galat tidak berkorelasi, antar prediktor tidak bermultikolinearitas. Jika data tidak memenuhi salah satu asumsi klasik regresi, maka penduga metode kuadrat terkecil tidak lagi efisien<sup>[8]</sup>.

Pencilan merupakan data yang nilainya atau letaknya berbeda sangat signifikan terhadap kumpulan data lainnya. Ada banyak cara pendeteksiannya, dan penyajiannya dapat berupa grafik atau nilai. Adanya pencilan dalam data dapat mengakibatkan estimasi

koefisien regresi yang diperoleh tidak tepat. Hal ini dapat ditunjukkan dengan nilai standar error yang besar apabila menggunakan metode kuadrat terkecil. Namun demikian tindakan membuang begitu saja suatu pencilan bukanlah tindakan yang bijaksana karena ada kalanya pencilan memberikan informasi yang cukup berarti. Oleh karena itu, diperlukan suatu estimasi yang bersifat *robust* atau tahan terhadap pencilan yang dikenal dengan regresi *robust*. Suatu estimasi yang tahan adalah relatif tidak terpengaruh oleh perubahan besar pada bagian kecil data atau perubahan kecil pada bagian besar data<sup>[5,6]</sup>.

Beberapa metode estimasi dalam regresi *robust* diantaranya adalah estimasi-M, *Least Trimmed Squares* (LTS), *Least Median Squares* (LMS), estimasi *S* dan estimasi *MM* (*Method of Moment*)<sup>[2]</sup>. Estimasi-M merupakan suatu metode regresi *robust* yang terkenal dan paling luas digunakan daripada metode regresi *robust* yang lain, karena hasilnya lebih teliti. Metode *Iteratively Reweighted Least Squares* (IRLS) merupakan salah satu metode iterasi pada estimasi-M yang memerlukan 3 fungsi pembobot, yaitu metode kuadrat terkecil, fungsi Huber dan fungsi Bisquare Tukey. Fungsi pembobot yang akan dipilih adalah fungsi pembobot Huber dan fungsi pembobot Bisquare Tukey.

Masalah yang akan dibahas dalam tulisan ini adalah pengujian ketidakpenuhan asumsi klasik, cara pendeteksian pencilan dan penaksiran model pada data ketahanan pangan Jawa Tengah tahun 2007 menggunakan metode refresi *robust* hingga didapat taksiran model terbaiknya. Tujuan dari penulisan ini adalah untuk menggunakan regresi *robust* estimasi-M IRLS dengan fungsi pembobot Huber dan Tukey Bisquare pada kumpulan data yang terdapat pencilan.

## 2. Tinjauan Pustaka

### 2.1 Cook's Distance

Metode untuk mendeteksi adanya *pencilan* adalah menggunakan *Cook's Distance*, yaitu

$$D_i = \left( \frac{(\mathbf{b} - \mathbf{b}_{(i)})' \mathbf{X}' \mathbf{X} (\mathbf{b} - \mathbf{b}_{(i)})}{p \text{ MSE}} \right)$$

dengan:

$\mathbf{b}$  = vektor koefisien regresi dengan semua pengamatan ikut dalam pembentukan model regresi

$\mathbf{b}_{(i)}$  = vektor koefisien regresi yang tidak mengikutsertakan pengamatan ke-*i* dalam pembentukan model regresi

$\mathbf{X}$  = matriks variable predictor ukuran (*n* x *p*)

*p* = banyaknya parameter

MSE = *Mean Square Error* dari semua pengamatan dalam pembuatan model regresi

$$= \frac{\sum_{i=1}^n e_i^2}{n}$$

Suatu kumpulan data pengamatan dikatakan mengandung pencilan jika nilai pengamatan

$D_i > \frac{p}{n}$ , dengan *n* menyatakan ukuran sampel<sup>[4]</sup>.

## 2.2 Estimasi-M IRLS

Salah satu kelas estimasi *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  $\rho$

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

Fungsi  $\rho$  merupakan representasi pembobot dari residual. Untuk memperoleh suatu skala invariant dari estimator ini, biasanya dilakukan dengan menyelesaikan persamaan

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

dengan  $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$  merupakan nilai estimasi-M dari  $\beta_0, \beta_1, \dots, \beta_k$  yang meminimumkan

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

dimana  $\rho(u_i)$  adalah fungsi simetris dari residual atau fungsi yang memberikan kontribusi pada masing-masing residual pada fungsi objektif<sup>[2]</sup>.

Pada umumnya, suatu estimasi skala *robust* perlu diestimasi. Pilihan estimasi yang populer untuk  $s$  adalah

$$s = \frac{\text{median} \{ |e_i - \text{median}(e_i)| \}}{0,6745}$$

Pemilihan konstanta 0,6745 membuat sedemikian hingga  $s$  merupakan suatu estimator yang mendekati tak bias dari  $\sigma$ , jika  $n$  besar dan error berdistribusi normal<sup>[4]</sup>.

## 2.3 Penyelesaian untuk Koefisien Regresi

Untuk meminimumkan persamaan (1), turunan parsial pertama dari  $\rho$  terhadap  $\beta_j$ ,  $j = 0, 1, \dots, k$ , harus disamakan dengan 0. Sehingga akan menghasilkan suatu syarat perlu untuk minimum. Ini menghasilkan sistem persamaan (2):

$$\frac{\partial \rho}{\partial \beta_j} = 0$$

$$\sum_{i=1}^n x_{ij} \psi \left( \frac{y_i - \sum_{j=0}^k x_{ij} \beta_j}{s} \right) = 0, \quad j = 0, 1, \dots, k \quad (2)$$

dengan  $\psi = \rho'$  dan  $x_{ij}$  adalah observasi ke- $i$  pada regresor ke- $j$  dan  $x_{i0} = 1$ .

Didefinisikan suatu fungsi bobot

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

dan misal  $w_i = w(u_i)$ . Maka persamaan (2) dapat ditulis sebagai

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

Pada umumnya, fungsi  $\psi$  tidak linear dan persamaan (2) harus diselesaikan dengan metode iterasi.

Estimasi koefisien regresi dengan estimasi-M dilakukan dengan estimasi kuadrat terkecil dengan pembobot iteratif. Prosedur estimasi ini membutuhkan proses iterasi dimana  $w_i$  akan berubah pada tiap iterasinya sehingga diperoleh  $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$ . Prosedur tersebut dinamakan *Iteratively Reweighted Least Squares* (IRLS). Untuk menggunakan IRLS, anggap bahwa suatu estimasi awal  $\hat{\beta}^0$  ada dan  $s$  adalah suatu estimasi skala. Untuk parameter dengan  $p$  adalah jumlah parameter yang akan diestimasi, maka

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

dengan

$$w_i^0 = \begin{cases} \frac{\psi \left( \frac{y_i - \sum_{j=0}^k x_{ij} \hat{\beta}_j^0}{s} \right)}{y_i - \sum_{j=0}^k x_{ij} \hat{\beta}_j^0}, & \text{jika } y_i \neq \sum_{j=0}^k x_{ij} \hat{\beta}_j^0 \\ 1, & \text{jika } y_i = \sum_{j=0}^k x_{ij} \hat{\beta}_j^0 \end{cases} \quad (3)$$

Untuk kasus regresi berganda perhitungan parameternya dapat diperoleh dari persamaan matriks

$$\mathbf{X}' \mathbf{W}^0 \mathbf{X} \hat{\beta} = \mathbf{X}' \mathbf{W}^0 \mathbf{Y}$$

$\mathbf{W}^0$  adalah matriks diagonal berukuran  $(n \times n)$  dari bobot dengan elemen-elemen diagonal  $w_1^0, w_2^0, \dots, w_n^0$  diberikan oleh persamaan (3). Maka dari itu, estimator satu langkah adalah

$$\hat{\beta}^1 = (\mathbf{X}' \mathbf{W}^0 \mathbf{X})^{-1} \mathbf{X}' \mathbf{W}^0 \mathbf{Y}$$

Pada langkah selanjutnya, dihitung kembali bobot dari  $w_i = w(u_i)$  tetapi menggunakan  $\hat{\beta}^1$  sebagai pengganti  $\hat{\beta}^0$ , dan seterusnya. Perhitungan iterasi ini dihentikan bila perubahan yang terjadi pada koefisien regresi yaitu selisih antara  $\hat{\beta}^{l+1}$  dengan  $\hat{\beta}^l$  lebih kecil dari 0,1%, dengan  $l = 0, 1, \dots$ . Estimasi regresi robust dengan estimasi-M IRLS dapat ditulis

$$\hat{\beta}^{l+1} = (\mathbf{X}' \mathbf{W}^l \mathbf{X})^{-1} (\mathbf{X}' \mathbf{W}^l \mathbf{Y})$$

Estimasi kuadrat terkecil dapat digunakan sebagai nilai permulaan,  $\hat{\beta}^0$ . Selanjutnya, untuk  $\hat{\beta}^2$  dapat dituliskan sebagai berikut

$$\hat{\beta}^2 = (\mathbf{X}'\mathbf{W}^1\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}^1\mathbf{Y}$$

## 2.4 Fungsi Pembobot

Fungsi pembobot dalam estimasi-M bergantung pada residual dan konstanta tertentu. Fungsi pembobot yang digunakan adalah MKT, Huber dan Bisquare Tukey.

1. Fungsi pembobot Metode Kuadrat Terkecil

$$w_{LS}(u_i) = 1$$

2. Fungsi pembobot Metode Huber

$$w_H(u_i) = \begin{cases} 1, & \text{untuk } |u_i| \leq r \\ \frac{r}{|u_i|}, & \text{untuk } |u_i| > r \end{cases}$$

3. Metode Bisquare Tukey

$$w_B(u_i) = \begin{cases} \left(1 - \left(\frac{u_i}{r}\right)^2\right)^2, & \text{untuk } |u_i| \leq r \\ 0, & \text{untuk } |u_i| > r \end{cases}$$

Pada persamaan di atas  $u_i$  merupakan residual ke- $i$ , sedangkan nilai  $r$  dinyatakan dengan *tuning constant*. *Tuning constant* dalam regresi robust menentukan kerobusan penaksir terhadap pencilan dan efisiensi penaksir dalam ketidakadaan pencilan. Jika diambil  $\alpha = 5\%$ , maka estimasi-M Huber akan efektif digunakan bilamana  $r = 1,345$  sedangkan pada Bisquare Tukey bilamana  $r = 4,685$ . Permasalahan dalam estimasi regresi robust adalah perlu dilakukan pemilihan *tuning constant* agar estimasi yang diperoleh lebih spesifik dan meminimumkan jumlah kuadrat residual<sup>[4]</sup>.

## 3. Hasil dan Pembahasan

### 3.1 Sumber Data

Pada tulisan ini akan dilakukan estimasi regresi robust pada model rasio ketersediaan beras di Jawa Tengah pada tahun 2007. Data yang digunakan adalah data sekunder yang bersumber dari Badan Pusat Statistik Jawa Tengah. Adapun untuk hubungan variable rasio ketersediaan beras di Jawa Tengah dipengaruhi oleh stok beras, luas areal panen, rata-rata produksi padi, rata-rata harga beras, dan jumlah konsumsi beras, yang disajikan pada Tabel 1 berikut.

**Tabel 1.** Data Ketahanan Pangan di Jawa Tengah Tahun 2007<sup>[7]</sup>

Kab/Kota	Produksi (ton)	Stok (ton)	Luas (Ha)	RataProd (ton/ha)	Harga (Rp/ton)	Jmlkons (ton)	Rasio
Cilacap	622,442	20.974.997	111.725	5,571	4.510.000	183.418	3,39
Banyumas	351,340	21.348.038	64.989	5,406	4.519.000	169.045	2,08
Purbalingga	188,644	13.006.013	35.590	5,300	4.354.000	92.871	2,03
Banjarnegara	145,025	13.902.979	27.132	5,345	3.950.000	97.648	1,49
Kebumen	360,331	16.287.499	67.959	5,302	3.981.000	136.584	2,64
Purworejo	284,618	8.192.333	52.729	5,398	4.283.000	81.291	3,50
Wonosobo	156,034	10.553.640	29.793	5,237	4.908.000	85.252	1,83
Magelang	280,093	14.826.134	53.481	5,237	4.574.000	131.224	2,13
Boyolali	225,248	11.705.054	41.717	5,399	4.827.000	105.394	2,14
Klaten	327,522	14.769.888	58.505	5,598	4.923.000	127.560	2,57
Sukoharjo	267,230	9.032.458	46.176	5,787	4.572.000	92.617	2,89
Wonogiri	269,556	10.626.106	54.622	4,935	4.748.000	110.754	2,43
Karanganyar	243,685	8.375.952	42.826	5,690	4.857.000	91.017	2,68
Sragen	493,681	10.782.643	90.833	5,435	2.562.000	96.936	5,09
Grobogan	571,485	20.731.709	101.994	5,603	4.938.000	149.884	3,81
Blora	320,851	11.108.962	63.513	5,052	5.036.000	94.005	3,41
Rembang	132,025	8.961.480	26.895	4,909	4.896.000	64.735	2,04
Pati	385,164	15.549.600	76.608	5,028	4.948.000	131.941	2,92
Kudus	127,543	4.373.232	24.992	5,103	5.652.000	87.556	1,46
Jepara	198,981	9.978.677	38.020	5,234	5.326.000	121.320	1,64
Demak	502,407	14.459.342	91.516	5,490	4.835.000	115.868	4,34
Semarang	170,787	9.225.898	32.862	5,197	5.396.000	101.747	1,68
Temanggung	177,551	7.588.949	32.624	5,442	4.643.000	79.195	2,24
Kendal	214,111	9.074.419	40.063	5,344	4.712.000	106.006	2,02
Batang	207,477	9.762.173	40.265	5,153	4.720.000	76.716	2,70
Pekalongan	223,888	11.454.328	44.457	5,036	4.841.000	95.397	2,35
Pemalang	357,467	14.188.973	70.694	5,057	4.865.000	153.561	2,33
Tegal	298,062	11.056.435	55.898	5,332	4.566.000	159.362	1,87
Brebes	458,518	28.133.467	84.696	5,414	4.299.000	200.681	2,28
Kota Magelang	2,513	848.309	484	5,192	4.640.000	14.936	0,17
Kota Surakarta	1,782	3.264.077	347	5,138	4.784.000	58.483	0,03
Kota Salatiga	7,134	1.124.184	1.385	5,151	5.121.000	19.740	0,36
Kota Semarang	24,689	10.172.117	5.046	4,893	4.711.000	168.216	0,15
Kota Pekalongan	11,835	2.819.165	2.315	5,112	4.729.000	30.887	0,38
Kota Tegal	7,135	1.964.309	1.347	5,297	4.934.000	27.104	0,26

### 3.2 Identifikasi Variabel

Berikut adalah data-data yang diperlukan dalam penelitian.

#### 1. Variabel Tak Bebas

Variabel tak bebas berupa rasio ketersediaan beras di Jawa Tengah pada tahun 2007. Nilai rasio diperoleh dari perbandingan antara produksi dan konsumsi beras di tiap daerah.

#### 2. Variabel Bebas

Beberapa variabel bebas yang digunakan dalam penelitian ini berupa data faktor-faktor yang diduga mempengaruhi rasio ketahanan pangan di Jawa Tengah, meliputi:

1. Stok beras ( $X_1$ )
2. Luas panen padi ( $X_2$ )
3. Rata-rata produksi ( $X_3$ )
4. Harga beras ( $X_4$ )
5. Jumlah konsumsi ( $X_5$ )

### 3.3 Metode Kuadrat Terkecil

Setelah diolah menggunakan bantuan *Software Minitab 14* diperoleh taksiran model regresi dengan metode kuadrat terkecil yaitu

$$\hat{y}_i = 1,32 - 0,000045 x_{i1} + 0,000048 x_{i2} + 0,0188 x_{i3} - 0,000237 x_{i4} - 0,000007 x_{i5}$$

Selanjutnya akan dilakukan pengujian asumsi klasik yang bertujuan untuk melihat apakah taksiran model yang diperoleh terdapat penyimpangan klasik atau tidak. Ternyata setelah dilakukan uji asumsi klasik yang meliputi uji asumsi: normalitas, nonmultikolinearitas adalah dipenuhi, tetapi untuk asumsi nonautokorelasi dan homoskedastisitas tidak dipenuhi. Sehingga perlu dilakukan penanganan lebih lanjut agar diperoleh estimasi regresi yang tepat, yaitu dengan menggunakan estimasi-M pada regresi robust dengan fungsi pembobot Huber dan Bisquare Tukey. Hal ini dimungkinkan terdapat pencilan pada data tersebut. Oleh karena itu, untuk langkah selanjutnya perlu dilakukan pendeteksian pencilan.

### 3.4 Pendeteksian Pencilan

Untuk mendeteksi pencilan dari data dideteksi menggunakan metode *Cook's distance*. Berdasarkan hasil perhitungan didapat hasil bahwa data ke-1, yaitu Kabupaten Cilacap, data ke-14, yaitu Kabupaten Sragen dan data ke-33 yaitu Kota Semarang merupakan pencilan karena mempunyai nilai yang lebih besar dari  $6/35 = 0,17143$ , yaitu masing-masing 0,222554; 0,220494 dan 0,400533.

### 3.5 Estimasi-M IRLS

Adanya pencilan pada data ketahanan pangan tahun 2007 menyebabkan taksiran model regresi belum bisa dikatakan baik. Karena terdapat asumsi klasik yang tidak dipenuhi. Sehingga perlu dilakukan alternatif lain untuk mengestimasi parameter-parameter yang tidak peka terdapat pencilan yaitu dengan metode regresi robust dengan IRLS. Akan dipergunakan metode estimasi-M IRLS menggunakan fungsi pembobot Huber dan Bisquare Tukey. Prosesnya menggunakan cara iteratif.

#### 3.5.1 IRLS dengan Fungsi Huber

Proses iteratif dimulai dengan menentukan estimasi awal koefisien regresi yang diperoleh dengan metode kuadrat terkecil. Taksiran model yang diperoleh adalah

$$\hat{y}_i = 1,32 - 0,000045 x_{i1} + 0,000048 x_{i2} + 0,0188 x_{i3} - 0,000237 x_{i4} - 0,000007 x_{i5}$$

Selanjutnya, dihitung nilai residual  $e_i^{(0)} = y_i - \hat{y}_i$  dengan  $i = 1, 2, \dots, 35$ , dan dihitung pula nilai  $s^{(0)}$  dari 35 kabupaten/kota, yaitu

$$s^{(0)} = \frac{\text{median} \left\{ e_i^{(0)} - \text{median} \left( e_i^{(0)} \right) \right\}}{0,6745} = 0,44$$

Skala residual untuk semuanya menggunakan

$$u_i^{(0)} = \frac{e_i^{(0)}}{s^{(0)}}, \quad i = 1, 2, \dots, 35$$

Kriteria pemberian bobot pada fungsi Huber berdasarkan nilai skala residualnya

$$w_i^{(0)} = \begin{cases} 1, & |u_i^{(0)}| \leq 1,345 \\ \frac{1,345}{|u_i^{(0)}|}, & |u_i^{(0)}| > 1,345 \end{cases}, \quad i = 1, 2, \dots, 35$$

Pada iterasi pertama menggunakan WLS dengan pembobot  $w_i^{(0)}$  diperoleh taksiran model regresi linear sebagai berikut:

$$\hat{y}_i = 1,12 - 0,000048 x_{i1} + 0,000049 x_{i2} + 0,0231 x_{i3} - 0,000235 x_{i4} - 0,000007 x_{i5}$$

Nilai residual dari model iterasi 1 digunakan untuk iterasi ke-2. Iterasi akan terus berlanjut hingga diperoleh nilai  $\hat{\beta}$  yang konvergen atau sama dengan hasil iterasi sebelumnya. Ternyata hasil perhitungan  $\hat{\beta}$  untuk tiap iterasi berhenti pada iterasi ke-8, karena nilai  $\hat{\beta}$  yang baru sama dengan nilai  $\hat{\beta}$  sebelumnya, seperti yang tertulis di bawah ini.

**Tabel 2.** Hasil Iterasi Estimasi Parameter menggunakan Fungsi Huber

Iterasi	Huber					
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$
MKT	1,32	- 0,000045	0,000048	0,0188	- 0,000237	- 0,000007
1	1,12	- 0,000048	0,000049	0,0231	- 0,000235	- 0,000007
2	1,03	- 0,000048	0,000048	0,0253	- 0,000232	- 0,000007
3	1,03	- 0,000048	0,000048	0,0261	- 0,000233	- 0,000008
4	1,08	- 0,000048	0,000048	0,0260	- 0,000234	- 0,000008
5	1,14	- 0,000047	0,000048	0,0252	- 0,000235	- 0,000008
6	1,17	- 0,000047	0,000048	0,0248	- 0,000236	- 0,000008
7	1,17	- 0,000047	0,000048	0,0247	- 0,000236	- 0,000008
8	1,17	- 0,000047	0,000048	0,0247	- 0,000236	- 0,000008

Jadi dengan menggunakan estimasi-M IRLS memakai fungsi pembobot Huber diperoleh taksiran model regresi linier sebagai berikut

$$\hat{y}_i = 1,17 - 0,000047 x_{i1} + 0,000048 x_{i2} + 0,0247 x_{i3} - 0,000236 x_{i4} - 0,000008 x_{i5}$$

### 3.5.2 IRLS dengan Fungsi Bisquare Tukey

Jika menggunakan metode IRLS dengan fungsi pembobot Bisquare Tukey. Proses iteratif dimulai dengan menentukan estimasi awal koefisien regresi yang diperoleh dengan metode kuadrat terkecil. Model yang diperoleh adalah

$$\hat{y}_i = 1,32 - 0,000045 x_{i1} + 0,000048 x_{i2} + 0,0188 x_{i3} - 0,000237 x_{i4} - 0,000007 x_{i5}$$

Selanjutnya, dihitung nilai residual  $e_i^{(0)} = y_i - \hat{y}_i$  dengan  $i = 1, 2, \dots, 35$ , dan dihitung pula nilai  $s^{(0)}$  dari 35 kabupaten/kota, yaitu

$$s^{(0)} = \frac{\text{median} \{e_i^{(0)} - \text{median}(e_i^{(0)})\}}{0,6745} = 0,44$$

Skala residual untuk semuanya menggunakan

$$u_i^{(0)} = \frac{e_i^{(0)}}{s^{(0)}}, \quad i = 1, 2, \dots, 35$$

Kriteria pemberian bobot pada fungsi Bisquare Tukey berdasarkan nilai skala residualnya

$$w_i^{(0)} = \begin{cases} \left[ 1 - \left( \frac{u_i^{(0)}}{4,685} \right)^2 \right]^2, & |u_i^{(0)}| \leq 4,685 \\ 0, & |u_i^{(0)}| > 4,685 \end{cases}, \quad i = 1, 2, \dots, 35$$

Pada iterasi pertama menggunakan WLS dengan pembobot  $w_i^{(0)}$  diperoleh taksiran model regresi linear sebagai berikut

$$\hat{y}_i = 1,06 - 0,000048 x_{i1} + 0,000048 x_{i2} + 0,0244 x_{i3} - 0,000234 x_{i4} - 0,000007 x_{i5}$$

Nilai residual dari model iterasi 1 akan digunakan untuk iterasi ke-2. Iterasi akan terus berlanjut hingga diperoleh nilai  $\hat{\beta}$  yang konvergen atau sama dengan hasil iterasi sebelumnya. Dalam kajian ini hasil perhitungan  $\hat{\beta}$  untuk setiap iterasi ternyata berhenti pada iterasi ke-8, yaitu dengan dihasilkannya nilai  $\hat{\beta}$  yang baru sama dengan nilai  $\hat{\beta}$  sebelumnya, seperti yang tertabelkan berikut ini.

**Tabel 3.** Hasil Iterasi Estimasi Parameter menggunakan Fungsi Bisquare Tukey

Iterasi	Bisquare Tukey					
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$
MKT	1,32	- 0,000045	0,000048	0,0188	- 0,000237	- 0,000007
1	1,06	- 0,000048	0,000049	0,0244	- 0,000234	- 0,000007
2	0,93	- 0,000049	0,000048	0,0269	- 0,000232	- 0,000007
3	0,91	- 0,000049	0,000048	0,0280	- 0,000232	- 0,000008
4	0,94	- 0,000049	0,000048	0,0282	- 0,000234	- 0,000008
5	0,99	- 0,000048	0,000048	0,0277	- 0,000235	- 0,000008
6	1,03	- 0,000048	0,000048	0,0270	- 0,000235	- 0,000008
7	1,04	- 0,000048	0,000048	0,0268	- 0,000236	- 0,000008
8	1,04	- 0,000048	0,000048	0,0268	- 0,000236	- 0,000008

Jadi dengan menggunakan estimasi-M IRLS memakai fungsi pembobot Bisquare Tukey diperoleh taksiran model regresi linear

$$\hat{y}_i = 1,04 - 0,000048 x_{i1} + 0,000048 x_{i2} + 0,0268 x_{i3} - 0,000236 x_{i4} - 0,000008 x_{i5}$$

### 3.5.3 Model Terbaik

Kriteria yang dipakai untuk menentukan model regresi terbaik adalah menggunakan  $R^2_{adjusted}$  dan MSE. Hasil perbandingan antara  $R^2_{adjusted}$  dan MSE pada fungsi pembobot Huber dan Tukey Bisquares disajikan pada Tabel 4 berikut ini

**Tabel 4.** Hasil Kriteria Model Berdasarkan Fungsi Pembobot

Fungsi Pembobot	$R^2_{adjusted}$ (%)	MSE
Huber	85,7	0,17
Bisquare Tukey	86,1	0,1573

Berdasarkan hasil di atas, dapat dikatakan bahwa metode yang paling baik untuk mengestimasi model ketahanan pangan di Jawa Tengah pada tahun 2007 adalah menggunakan metode fungsi pembobot Bisquare Tukey, dengan tingkat kepercayaan model sebesar 86,1 % bahwa rasio ketersediaan beras dipengaruhi oleh stok beras, luas

panen, rata-rata produksi, harga beras, jumlah konsumsi. Sedangkan 13,9% dipengaruhi oleh faktor lain atau kesalahan yang bersifat random.

#### 4. Kesimpulan

Berdasarkan pembahasan yang telah dilakukan, dapat disimpulkan bahwa:

1. Pendeteksian pencilan yang dilakukan menggunakan *Cook's distance* menduga bahwa data ke-1, ke-14 dan data ke-33 masing-masing sebagai pencilan, yaitu kabupaten Cilacap, Kabupaten Sragen dan Kota Semarang.
2. Taksiran model regresi pada data ketahanan pangan di Jawa Tengah tahun 2007 adalah sebagai berikut
  - a) Menggunakan fungsi pembobot Huber
$$\hat{y}_i = 1,17 - 0,000047 x_{i1} + 0,000048 x_{i2} + 0,0247 x_{i3} - 0,000236 x_{i4} - 0,000008 x_{i5}$$
  - b) Menggunakan fungsi pembobot Bisquare Tukey
$$\hat{y}_i = 1,04 - 0,000048 x_{i1} + 0,000048 x_{i2} + 0,0268 x_{i3} - 0,000236 x_{i4} - 0,000008 x_{i5}$$
3. Berdasarkan nilai MSE dan koefisien determinasi dapat dikatakan bahwa metode estimasi-M IRLS dengan fungsi pembobot Bisquare Tukey lebih baik dipilih dari pada menggunakan fungsi pembobot Huber untuk taksiran model ketahanan pangan di Jawa Tengah tahun 2007.

#### DAFTAR PUSTAKA

1. Bain, L.J., and Engelhart, M., *Introduction to Probability and Mathematical Statistics*, Second Edition, Duxbury Press, Belmont, California, 1992.
2. Chen, C., *Robust Regression and Outlier Detection with the ROBUSTREG Procedure*, Presented at Sugi No. 27, Institute Inc., Cary NC, 1992.
3. Gujarati, D., *Basics Ekonometrics*, McGraw-Hill, Inc., 1995.
4. Montgomery, D.C. and Peck, E.A., *Introduction to Linier Regression Analysis*, John Wiley and Sons, New York, 1992.
5. Sembiring, R.K., *Analisis Regresi*, Penerbit ITB, Bandung, 1995.
6. Soemartini, *Pencilan (Outlier)*, FMIPA Universitas Padjadjaran, 2007.
7. [www.bps.go.id](http://www.bps.go.id), diakses pada tanggal 10 Mei 2011.
8. Yafee, R.A., *Robust Regression Analysis: Some Popular Statistical Package Options*, Academic Computing Services, 2002.