KEKEKARAN REGRESI LINIER GANDA DENGAN ESTIMASI MM (*METHOD OF MOMENT*) DALAM MENGATASI PENCILAN

SKRIPSI

Diajukan kepada Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Negeri Yogyakarta
untuk memenuhi sebagian persyaratan guna memperoleh gelar Sarjana Sains

Oleh
Lina Dewi Kurniawati
NIM. 07305141009

PROGRAM STUDI MATEMATIKA
JURUSAN PENDIDIKAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI YOGYAKARTA
2011

PERSETUJUAN

KEKEKARAN REGRESI LINIER GANDA DENGAN ESTIMASI MM (*METHOD OF MOMENT*) DALAM MENGATASI PENCILAN

Oleh:

Lina Dewi Kurniawati

07305141009

Telah disetujui dan disyahkan pada tanggal

20 Juni 2011

untuk diujikan di depan dewan penguji skripsi

Program Studi Matematika

Jurusan Pendidikan Matematika

Fakultas Matematika dan Ilmu Pengetahuan Alam

Universitas Negeri Yogyakarta

Menyetujui,

Pembimbing

Endang Listyani, M. S

NIP. 19591115 198601 2 001

SKRIPSI

KEKEKARAN REGRESI LINIER GANDA DENGAN ESTIMASI MM (*METHOD OF MOMENT*) DALAM MENGATASI PENCILAN

Oleh:
Lina Dewi Kurniawati
07305141009

Telah diujikan di depan dewan penguji skripsi Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Yogyakarta pada tanggal 6 juli 2011 dan dinyatakan telah memenuhi persyaratan guna memperoleh gelar Sarjana Sains.

DEWAN PENGUJI

Nama	Jabatan	Tandatangan	Tanggal
Endang Listyani, M.S.	Ketua Penguji	....	28/7/2011
Kismiantini, M.Si.	Sekretaris Penguji	....	28/7/2011
M. Susanti, M.Si.	Penguji Utama	....	25/7/2011
Dr. Heri Retnowati	Penguji Pendamping	....	14/7/2011

Yogyakarta, 29 Juli 2011
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Negeri Yogyakarta
Dekan,

KEMENTERIAN PENDIDIKAN NASIONAL UNIVERSITAS NEGERI YOGYAKARTA FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

Dr. Ariswan
NIP. 19590914 198803 1 003

PERNYATAAN

Yang bertanda tangan dibawah ini:

Nama : Lina Dewi Kurniawati

NIM : 07305141009

Prodi/ Jurusan : Matematika/ Pendidikan Matematika

Fakultas : Matematika dan Ilmu Pengetahuan Alam

Judul Skripsi : **Kekekaran Regresi Linier Ganda dengan Estimasi MM (*Method of Moment*) dalam Mengatasi Pencilan**

Dengan ini saya menyatakan bahwa skripsi ini adalah hasil pekerjaan saya sendiri. Sepanjang pengetahuan saya tidak terdapat karya atau pendapat yang ditulis atau diterbitkan orang lain kecuali sebagai acuan atau kutipan dengan mengikuti tata penulisan karya ilmiah yang telah lazim.

Yogyakarta,
Yang menyatakan,

Lina Dewi Kurniawati
NIM. 07305141009

MOTO

"Allah tidak akan membebani seseorang melainkan sesuai dengan kemampuannya…"
(QS. Al-Baqarah: 286)

"Sesungguhnya bersama kesulitan itu pasti ada kemudahan. Maka apabila kamu telah selesai (urusan dunia), bersungguh-sungguhlah (dalam beribadah)"
(QS. Al-Insyiroh: 6-7)

"Barang siapa menempuh jalan untuk memperoleh ilmu, maka Allah akan memudahkan baginya jalan menuju surga"
(H. R Muslim dari Abi Hurairah)

"Pertemuan yang menakutkan, selalu menyembunyikan hadiah-hadiah terbaik. Dan keajaiban mudah-mudahan diturunkan bagi yang memberanikan diri untuk memikul beban yang lebih besar daripada kemampuannya"
(Mario Teguh)

PERSEMBAHAN

Skripsi ini kupersembahkan khusus untuk:

- *Kedua Orangtuaku tercinta yang selalu mendoakan yang terbaik untukku*
- *Adikku tersayang Nelli Dwi Astuti yang selalu memotivasiku*
- *Penyemangatku Devriyadi Saputra S. yang selalu memberi kasih sayang, membantuku, dan selalu menemaniku saat suka maupun duka*
- *Seluruh keluargaku yang selalu mendukung dan mendoakanku.*
- *Teman-teman S.O.V: Anna, Nawang, Riza, Retno, Aziza, Dhita, Susi, Ika, dan Fifi yang selalu memberiku semangat. Terimakasih kebersamaanya selama ini*

KATA PENGANTAR

Puji dan syukur kehadirat Allah SWT yang telah melimpahkan rahmat dan hidayah-Nya, sehingga penulis dapat menyelesaikan penyusunan skripsi ini dengan baik dan lancar.

Penulis dapat menyelesaikan skripsi ini tidak lepas dari bantuan berbagai pihak. Dalam kesempatan ini penulis mengucapkan terima kasih kepada:

1. Bapak Dr. Ariswan selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Yogyakarta, yang telah memberikan ijin dalam penulisan skripsi ini.
2. Bapak Dr. Hartono selaku Kajurdik Pendidikan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Yogyakarta, yang telah memberikan ijin dalam menulis skripsi ini.
3. Ibu Atmini Dhoruri, M.S., selaku Kaprodi Matematika, yang telah membantu demi kelancaran penulisan.
4. Ibu Endang Listyani, M.S., selaku dosen pembimbing, yang telah memberikan bimbingan, saran dan pengarahan dalam penulisan skripsi ini.
5. Kedua orang tua dan seluruh keluarga yang selalu mendoakan, memberi motivasi, dan semangat sehingga skripsi ini dapat diselesaikan dengan baik.
6. Anna, Nawang, Riza, Retno, Aziza, Dhita, Ika, Susi, Fifi dan semua teman-teman Matematika 2007 yang selalu memberi bantuan, semangat dan dukungannya selama ini.

7. Semua pihak yang secara langsung atau tidak langsung telah memberikan bantuan dan saran yang tidak dapat penulis sebutkan satu-persatu.

Penulis berharap semoga skripsi ini bermanfaat tidak hanya bagi penulis tetapi juga bagi pembaca.

Yogyakarta,

Penyusun

KEKEKARAN REGRESI LINIER GANDA DENGAN ESTIMASI MM (*METHOD OF MOMENT*) DALAM MENGATASI PENCILAN

Oleh :
Lina Dewi Kurniawati
NIM. 07305141009

ABSTRAK

Tujuan penulisan ini adalah menunjukkan langkah-langkah dalam menduga parameter regresi dengan estimasi MM (*Method of Moment*) dan menunjukkan penerapan estimasi MM dalam regresi linier berganda.

Regresi *robust* merupakan metode regresi yang digunakan ketika ada beberapa *outlier* pada model. Adanya *outlier* menyebabkan estimasi koefisien regresi yang diperoleh tidak tepat. Metode estimasi MM merupakan gabungan dari metode estimasi S (*high breakdown*) dan metode estimasi M. Model regresi yang akan diestimasi yaitu regresi linier berganda, yang berbentuk $Y_i = \beta_0 + \beta_1 X_{i1} + ... + \beta_k X_{ik} + \varepsilon_i$. Langkah pertama dalam metode estimasi MM yaitu mencari estimator S, kemudian menetapkan parameter-parameter regresi menggunakan metode estimasi M. Sebelum mengestimasi dengan MM, data diidentifikasi terlebih dahulu dengan diagram pencar dan DfFITS (*Difference fitted value FITS*) untuk mengetahui apakah data tersebut mengandung pencilan. Setelah data dianalisis dan terdeteksi adanya pencilan kemudian diestimasi dengan metode MM untuk mendapatkan model regresinya. Pada kasus ini dalam mengestimasi parameter regresi dengan software SAS 9.1.

Contoh kasus pertama yaitu mengenai hubungan antara gaji tahunan matematikawan dengan indeks mutu publikasi, lama pengalaman, dan indeks keberhasilan dalam memperoleh dukungan dana. Pada kasus ini, ada 1 observasi yang merupakan pencilan. Pada kasus kedua mengenai hubungan antara berat jenis kayu pinus dengan serat kayu pinus, kecepatan tumbuh, kelembaban tanah, penyerapan cahaya pada kayu pinus, dan kadar air pada kayu. Pada contoh kedua ada 2 observasi yang merupakan pencilan. Hasil pada kedua contoh tersebut menunjukkan bahwa metode estimasi MM dapat mengestimasi parameter pada data yang terdapat pencilan tanpa menghapus pencilan tersebut, tetapi hanya menurunkan bobot dari pencilan tersebut. Berbeda dengan metode kuadrat terkecil, apabila data terdeteksi adanya pencilan, untuk mendapatkan model regresi yang baik data pencilan tersebut dihapus. Padahal menghapus data bukan tindakan yang baik, dengan menghapus sebagian data berarti mengubah data aslinya sehingga kebenaran hasil prediksi masih dipertanyakan.

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

DAFTAR LAMPIRAN

BAB I
PENDAHULUAN

A. Latar Belakang

Analisis regresi merupakan suatu analisis statistik yang mempelajari hubungan antara variabel dependen dengan variabel independen. Di dalam analisis regresi, hubungan yang sebenarnya tidak dapat diketahui secara pasti tetapi model hubungan tersebut dapat diestimasi berdasarkan data pengamatan. Model regresi linier yang memuat beberapa variabel independen dan satu variabel dependen adalah model regresi linier berganda. Bentuk model regresi linier berganda adalah $Y_i = \beta_0 + \beta_1 X_{i1} + ... + \beta_k X_{ik} + \varepsilon_i, i = 1, 2, ..., n$ dengan Y_i adalah variabel dependen pada pengamatan ke-i, X_{ik} adalah variabel independen pada pengamatan ke-*i* dan parameter ke-*k* dan $\beta_0, \beta_1, ..., \beta_k$ adalah parameter regresi yang tidak diketahui nilainya dan akan dicari nilai estimasinya.

Dalam menentukan estimator terbaik sangat dipengaruhi oleh penggunaan metode. Metode yang biasa digunakan untuk mengestimasi parameter regresi antara lain adalah Metode Kuadrat Terkecil (MKT). Metode ini tidak dapat bekerja dengan baik apabila terdapat data pencilan (*outlier*). Pencilan adalah pengamatan yang jauh dari pusat data yang mungkin berpengaruh besar terhadap koefisien regresi. Untuk mengatasi kelemahan metode kuadrat terkecil tersebut dapat dilakukan dengan dua cara yaitu:

1. Mengeluarkan pencilan yang dapat dideteksi dengan *DfFITS* (*Difference fitted value* FITS), *Cook's Distance, DfBETA(s)* (*Difference fitted value*

Beta), setelah itu tetap menggunakan metode kuadrat terkecil (Soemartini, 2007:10).

2. Tetap menggunakan seluruh data, tetapi dengan memberikan bobot yang kecil untuk data pencilan, metode ini dikenal dengan nama metode **regresi *robust*** (Soemartini, 2007: 12).

Regresi *robust* merupakan metode regresi yang digunakan ketika ada beberapa pencilan pada model. Metode ini merupakan alat penting untuk menganalisa data yang dipengaruhi oleh pencilan sehingga dihasilkan model yang *robust* atau kekar terhadap pencilan. Suatu estimator yang kekar adalah relatif tidak terpengaruh oleh perubahan besar pada bagian kecil data atau perubahan kecil pada bagian besar data.

Metode estimasi dalam regresi *robus*t diantaranya estimasi M (*Maximum Likelihood type*), LTS (*Least Trimmed Square*), estimasi MM (*Method of Moment*), dan estimasi S (*Scale*) (Colin Chen, 2002:1). Estimasi M adalah metode yang paling sederhana dan paling banyak digunakan yang mempunyai nilai efisiensi yang tinggi, sedangkan estimasi S, LTS, LMS adalah estimasi dengan nilai *breakdown* tinggi , tetapi estimasi S mempunyai nilai *breakdown* yang paling tinggi diantara ketiganya. *Breakdown point* adalah proporsi minimal dari banyaknya pencilan dibandingkan seluruh data pengamatan. Estimasi MM merupakan metode yang baik untuk menanggulangi pencilan dan dapat menghasilkan estimator yang *robust* (kekar) dan juga dapat menghasilkan *breakdown point* yang tinggi dengan efisiensi tinggi. Metode estimasi MM dikenalkan oleh Yohai (1987), metode ini mempertahankan ke-*robust*-an dari metode estimasi S, serta efisiensi dari metode estimasi M. Metode ini memadukan metode estimasi *high breakdown* (estimasi S) dan metode estimasi M. Diharapkan melalui metode regresi *robust* estimasi MM dapat diperoleh estimator yang baik sehingga menghasilkan model yang lebih baik dari model hasil MKT (Metode Kuadrat Terkecil). Oleh karena itu penulis mengangkat judul "Kekekaran Regresi Linier Ganda dengan Estimasi MM (*Method of Moment*) dalam Mengatasi Pencilan", untuk dijadikan salah satu referensi dalam mengestimasi parameter pada data yang mengandung pencilan.

B. Pembatasan Masalah

Dalam penulisan skripsi ini, penulis memberikan pembatasan masalah pada penentuan estimator dengan metode MM untuk mengatasi pencilan pada model regresi linier berganda.

C. Rumusan Masalah

Berdasarkan uraian latar belakang, maka rumusan masalah yang akan dibahas adalah:

1. Bagaimana mengestimasi parameter pada model regresi linier ganda dengan regresi *robust* estimasi MM.
2. Bagaimana penerapan regresi *robust* estimasi MM dalam regresi linier berganda.

D. Tujuan

Tujuan dalam penulisan skripsi ini adalah:

1. Menunjukkan langkah-langkah dalam mengestimasi parameter regresi *robust* dengan estimasi MM.

2. Menunjukkan penerapan estimasi MM dalam regresi linier berganda.

E. Manfaat

Manfaat dari penulisan ini adalah:

1. Bagi Penulis

   Menambah wawasan serta pengetahuan dalam bidang statistik khususnya mengenai metode regresi *robust* dengan estimasi MM.

2. Bagi Jurusan Matematika FMIPA UNY

   Menambah kelengkapan koleksi pustakadan menjadi dasar pertimbangan untuk penelitian-penelitian selanjutnya.

3. Bagi Mahasiswa

   Sebagai acuan untuk penulisan karya ilmiah selanjutnya khusunya mengenai regresi *robust*.

BAB II
LANDASAN TEORI

Teori yang diperlukan untuk mendukung pada bab pembahasan diantaranya adalah konsep dasar statistika, model regresi linier berganda, metode kuadrat terkecil, metode maksimum likelihood, *breakdown point*, pencilan (*outlier*), *Goodness of FIT*, parameter lokasi dan skala, dan fungsi obyektif.

A. Konsep Dasar Statistika

Pada sub bab ini, diberikan pengertian tentang variabel random, variabel random diskret, dan variabel random kontinu.

1.1 Variabel Random

Definisi 2.1 (Bain dan Engelhardt, 1992: 53) variabel random X adalah suatu fungsi dengan daerah asal S dan daerah hasil bilangan real R sedemikian sehingga $X(e) = x$ dengan $e \in S$ dan $x \in R$. Huruf besar seperti X, Y, Z digunakan untuk menotasikan variabel random. Sedangkan huruf kecil seperti x, y, z digunakan untuk menotasikan nilai yang mungkin dari setiap hasil observasi pada ruang sampel. Variabel random terbagi menjadi dua yaitu:

1.1.1 Variabel Random Diskret

Definisi 2.2 (Bain dan Engelhardt, 1992: 56) jika himpunan semua nilai yang mungkin dari variabel random X adalah himpunan terhitung (*countable*), $\{x_1, x_2, \ldots, x_n\}$ atau $\{x_1, x_2, \ldots\}$, maka X disebut variabel random diskret. Fungsi

$$f(x) = P[X=x],\ x = x_1, x_2, \ldots$$

menyatakan nilai peluang dengan setiap nilai x yang mungkin dinamakan fungsi densitas peluang diskret.

Sebuah fungsi *f(x)* disebut fungsi densitas peluang diskret jika dan hanya jika memenuhi:

i. $f(x_i) \geq 0, \forall x_i$

ii. $\sum_{\forall x_i} f(x_i) = 1$

Definisi 2.3 (Bain dan Engelhardt, 1992: 58) fungsi distribusi kumulatif dari variabel random *X* didefinisikan oleh $F(x) = P[X \leq x]$ dengan *x* bilangan real.

Sebuah fungsi *F(x)* disebut fungsi distribusi kumulatif dari variabel random *X* jika dan hanya jika memenuhi:

i. $\lim_{x \to -\infty} F(x) = 0$

ii. $\lim_{x \to \infty} F(x) = 1$

iii. $\lim_{h \to 0^+} F(x+h) = F(x)$

iv. a < b maka F(a) $\leq$ F(b)

Definisi 2.4 (Bain dan Engelhardt, 1992: 61) jika *X* adalah variabel random diskret dengan fungsi densitas peluang *f(x)*, maka nilai ekspektasi dari *X* didefinisikan sebagai:

$$E(X) = \sum_x x.f(x)$$

1.1.2 Variabel Random Kontinu

Definisi 2.5 (Bain dan Engelhardt, 1992: 64) variabel random X disebut variabel random kontinu jika terdapat fungsi *f(x)* yang disebut fungsi densitas peluang dari *X*, sehingga fungsi distribusi kumulatif dapat dituliskan sebagai:

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

Sebuah fungsi *f(x)* disebut fungsi densitas peluang dari variabel random kontinu *X* jika dan hanya jika memenuhi:

i. $f(x) \geq 0, \forall x$

ii. $\int_{-\infty}^{\infty} f(x)dx = 1$

Definisi 2.6 (Bain dan Engelhardt, 1992: 67) jika X adalah variabel random kontinu dengan fungsi densitas peluang *f(x)*, maka nilai ekspektasi dari *X* didefinisikan sebagai:

$$E(X) = \int_{-\infty}^{\infty} x.f(x)dx$$

B. Model Regresi Linier Berganda

Analisis regresi merupakan alat statistik yang bermanfaat untuk mengetahui hubungan antara dua variabel atau lebih sehingga salah satu variabel dapat diduga dari variabel lainnya. Dalam analisis regresi ini dapat diketahui bentuk dan pola hubungan yang ada dan juga dapat dilakukan prediksi berdasarkan nilai variabel yang sudah diketahui.

Analisis regresi digambarkan dalam model regresi yaitu suatu cara untuk mengekspresikan dua unsur penting suatu hubungan statistik, yaitu kecenderungan berubahnya variabel dependen (Y) sejalan dengan berubahnya variabel independen (X) dan berpencarnya titik-titik di sekitar kurva hubungan statistik itu. Jika analisis regresi dilakukan untuk satu varibel tidak bebas (Y) dengan lebih dari satu variabel bebas (X) maka regresi ini dinamakan regresi linier berganda dengan model

$$Y_i = \beta_0 + \beta_1 X_{i1} + ... + \beta_k X_{ik} + \varepsilon_i,\ \ i = 1, 2, ..., n \qquad (2.1)$$

dengan Y_i adalah variabel dependen pada pengamatan ke-i, X_{ik} adalah variabel independen pada pengamatan ke-i, dan $\beta_0, \beta_1, ..., \beta_k$ adalah parameter regresi yang tidak diketahui nilainya dan akan dicari nilai estimasinya, ε_i adalah galat yang berdistribusi normal dengan mean nol dan variansi σ^2 atau $\varepsilon_i \sim N\left(0, \sigma^2\right)$.

C. Metode Kuadrat Terkecil

Parameter $\beta_0, \beta_1, ..., \beta_k$ tidak diketahui, sehingga perlu diestimasi. Estimasi parameter yang biasa digunakan adalah metode kuadrat terkecil yaitu meminimumkan jumlah kuadrat galat. Dari persamaan (2.1) dapat ditulis:

$$S(\beta_j) = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 X_{i1} - ... - \beta_k X_{ik})^2 \qquad (2.2)$$

Untuk mencari nilai-nilai β dengan meminimumkan jumlah kuadrat galat, dicari turunan dari $S(\beta_j)$ secara parsial terhadap β_j , j = 0, 1, 2, …, k dan disama dengan nol, sehingga diperoleh:

$$\begin{aligned}
\frac{\partial S}{\partial \beta_0} &= -2\sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{i1} - ... - \beta_k x_{ik}) = 0, \\
\frac{\partial S}{\partial \beta_1} &= -2\sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{i1} - ... - \beta_k x_{ik}) x_{i1} = 0, \\
\frac{\partial S}{\partial \beta_2} &= -2\sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{i1} - ... - \beta_k x_{ik}) x_{i2} = 0, \qquad (2.3) \\
\vdots \quad & \quad \vdots \qquad\qquad \vdots \qquad\qquad\qquad \vdots \qquad \vdots
\end{aligned}$$

$$\frac{\partial S}{\partial \beta_k} = -2\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_{i1} - \ldots - \beta_k x_{ik})x_{ik} = 0,$$

Persamaan (2.3) menghasilkan persamaan normal sebagai berikut:

$$\begin{aligned}
& n\widehat{\beta}_0 + \widehat{\beta}_1 \sum_{i=1}^{n} x_{i1} + \widehat{\beta}_2 \sum_{i=1}^{n} x_{i2} + \ldots + \widehat{\beta}_k \sum_{i=1}^{n} x_{ik} = \sum_{i=1}^{n} y_i \\
& \widehat{\beta}_0 \sum_{i=1}^{n} x_{i1} + \widehat{\beta}_1 \sum_{i=1}^{n} x_{i1}^2 + \widehat{\beta}_2 \sum_{i=1}^{n} x_{i1}x_{i2} + \ldots + \widehat{\beta}_k \sum_{i=1}^{n} x_{i1}x_{ik} = \sum_{i=1}^{n} x_{i1}y_i \\
& \widehat{\beta}_0 \sum_{i=1}^{n} x_{i2} + \widehat{\beta}_1 \sum_{i=1}^{n} x_{i1}x_{i2} + \widehat{\beta}_2 \sum_{i=1}^{n} x_{i2}^2 + \ldots + \widehat{\beta}_k \sum_{i=1}^{n} x_{i2}x_{ik} = \sum_{i=1}^{n} x_{i2}y_i \\
& \vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad \vdots \\
& \widehat{\beta}_0 \sum_{i=1}^{n} x_{ik} + \widehat{\beta}_1 \sum_{i=1}^{n} x_{i1}x_{ik} + \widehat{\beta}_2 \sum_{i=1}^{n} x_{i2}x_{ik} + \ldots + \widehat{\beta}_k \sum_{i=1}^{n} x_{ik}^2 = \sum_{i=1}^{n} x_{ik}y_i
\end{aligned} \tag{2.4}$$

Jika disusun dalam bentuk matriks maka persamaan (2.4) menjadi:

$$\boldsymbol{X}'\boldsymbol{X}\widehat{\boldsymbol{\beta}} = \boldsymbol{X}'\boldsymbol{Y} \tag{2.5}$$

dengan,

$$\boldsymbol{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \boldsymbol{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ 1 & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}, \widehat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{bmatrix}$$

$$\boldsymbol{X}'\boldsymbol{X} = \begin{bmatrix} n & \sum_{i=1}^{n} x_{i1} & \cdots & \sum_{i=1}^{n} x_{ik} \\ \sum_{i=1}^{n} x_{i1} & \sum_{i=1}^{n} {x_{i1}}^2 & \cdots & \sum_{i=1}^{n} x_{i1}x_{ik} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} x_{ik} & \sum_{i=1}^{n} x_{i1}x_{ik} & \cdots & \sum_{i=1}^{n} {x_{ik}}^2 \end{bmatrix}$$

$$\boldsymbol{X'Y} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_{11} & x_{21} & \dots & x_{n1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1k} & x_{2k} & \dots & x_{nk} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_{i1} y_i \\ \vdots \\ \sum_{i=1}^{n} x_{ik} y_i \end{bmatrix}$$

Untuk menyelesaikan persamaan (2.5), kalikan kedua ruas dengan invers dari (**X'X**). Sehingga estimator kuadrat terkecil dari β adalah

$$(\boldsymbol{X'X})^{-1}\boldsymbol{X'X\hat{\beta}} = (\boldsymbol{X'X})^{-1}\boldsymbol{X'Y}$$

$$\underset{p\times 1}{\boldsymbol{\hat{\beta}}} = \underset{p\times p}{(\boldsymbol{X'X})^{-1}} \; \underset{p\times 1}{\boldsymbol{X'Y}}$$

Sifat-sifat estimator kuadrat terkecil (Gujarati, 2004: 79):

1. Linier

   Estimator bersifat linier yaitu merupakan fungsi linier dari variabel random.

Persamaan:

$$\hat{\beta} = \frac{\sum_{i=1}^{n} Y_i (x_i - \bar{x})}{\sum_{i=1}^{n} (x_i - \bar{x})^2} = \sum_{i=1}^{n} k_i Y_i$$

dengan,

$k_i = \dfrac{(x_i - \bar{x})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$, untuk i = 1, 2, …, n.

Menunjukkan bahwa $\hat{\beta}$ adalah estimator linier karena merupakan fungsi linier dari Y.

Sifat-sifat k_i:

a. Karena X_i diasumsikan nonstokastik, sehingga k_i merupakan nonstokastik juga.

b. $\sum k_i = 0$

Bukti:

$$\sum k_i = \sum \frac{(x_i - \overline{x})}{\sum (x_i - \overline{x})^2}$$

$$= \frac{1}{\sum (x_i - \overline{x})^2} \sum (x_i - \overline{x})$$

$$= \frac{1}{\sum (x_i - \overline{x})^2} \left(\sum x_i - n\overline{x} \right)$$

$$= \frac{1}{\sum (x_i - \overline{x})^2} \left(n\overline{x} - n\overline{x} \right) = 0$$

c. $\sum k_i^2 = \dfrac{1}{\sum (x_i - \overline{x})^2}$

Bukti:

$$\sum k_i^2 = \frac{\sum (x_i - \overline{x})^2}{\left(\sum (x_i - \overline{x})^2 \right)^2}$$

$$= \frac{1}{\left(\sum (x_i - \overline{x})^2 \right)^2} \sum (x_i - \overline{x})^2$$

$$= \frac{1}{\sum (x_i - \overline{x})^2}$$

d. $\sum k_i x_i = 1$

Bukti:

$$\sum k_i x_i = \frac{\sum (x_i - \bar{x})}{\sum (x_i - \bar{x})^2} x_i$$

$$= \frac{1}{\sum (x_i - \bar{x})^2} \sum (x_i - \bar{x}) x_i$$

$$= \frac{1}{\sum (x_i - \bar{x})^2} \sum (x_i^2 - x_i \bar{x})$$

$$= \frac{1}{\sum (x_i - \bar{x})^2} \left(\sum x_i^2 - \bar{x} \sum x_i \right)$$

$$= \frac{1}{\sum x_i^2 - n\bar{x}^2} \left(\sum x_i^2 - n\bar{x}^2 \right) = 1$$

2. Takbias

Takbias yaitu ekspektasi dari estimator $\hat{\beta}$ atau $E(\hat{\beta}) = \beta$.

Dari persamaan $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ subtitusikan kepersamaan $\hat{\beta} = \sum_{i=1}^{n} k_i Y_i$, dengan menggunakan sifat-sifat k_i sehingga diperoleh:

$$\hat{\beta} = \sum k_i (\beta_0 + \beta_1 X_i + \varepsilon_i)$$

$$= \beta_0 \sum k_i + \beta_1 \sum x_i k_i + \sum k_i \varepsilon_i$$

$$= \beta_1 + \sum k_i \varepsilon_i$$

maka,

$$\begin{aligned} E\left(\hat{\beta}\right) &= E\left(\beta + \sum k_i \varepsilon_i\right) \\ &= E\left(\beta\right) + \sum k_i E\left(\varepsilon_i\right) \\ &= \beta + \sum k_i (0) \\ &= \beta \end{aligned}$$

Terbukti estimator kuadrat terkecil bersifat takbias.

3. Memiliki variansi minimum

Suatu estimator takbias dengan variansi terkecil diketahui sebagai suatu estimator efisien. Dengan menggunakan definisi variansi, akan ditunjukkan bahwa estimasi kuadrat terkecil menghasilkan variansi minimum.

$$
\begin{aligned}
\operatorname{var}\left(\widehat{\beta}\right) &= E\left[\widehat{\beta} - E\left(\widehat{\beta}\right)\right]^2 \\
&= E\left(\widehat{\beta} - \beta\right)^2 \\
&= E\left(\beta + \sum k_i \varepsilon_i - \beta\right)^2 \\
&= E\left(\sum k_i \varepsilon_i\right)^2 \\
&= E\left(k_1^2 \varepsilon_1^2 + k_2^2 \varepsilon_2^2 + \ldots + k_n^2 \varepsilon_n^2 + 2k_1 k_2 \varepsilon_1 \varepsilon_2 + \ldots + 2k_{n-1} k_n \varepsilon_{n-1} \varepsilon_n\right) \\
&= E\left(k_1^2 \varepsilon_1^2 + k_2^2 \varepsilon_2^2 + \ldots + k_n^2 \varepsilon_n^2\right) + E\left(2k_1 k_2 \varepsilon_1 \varepsilon_2 + \ldots + 2k_{n-1} k_n \varepsilon_{n-1} \varepsilon_n\right) \\
&= \sum k_i^2 E\left[\varepsilon_i^2\right] + 2\sum k_i k_j E\left[\varepsilon_i \varepsilon_j\right] \\
&= \sum k_i^2 E\left[\varepsilon_i^2\right]
\end{aligned}
$$

Karena $E\left[\varepsilon_i^2\right] = \sigma^2$ untuk setiap *i* dan $E\left[\varepsilon_i \varepsilon_j\right] = 0, i \neq 0$ sehingga,

$$
\begin{aligned}
\operatorname{var}\left(\widehat{\beta}\right) &= \sigma^2 \sum k_i^2 \\
&= \frac{\sigma^2}{\sum \left(x_i - \overline{x}\right)^2} \quad \text{(menggunakan definisi } k_i^2\text{)}
\end{aligned}
$$

Misalkan suatu estimator linier alternatif β sebagai berikut:

$$\beta^* = \sum v_i Y_i$$

dimana v_i tidak perlu sama dengan k_i, sehingga:

$$
\begin{aligned}
E(\beta^*) &= \sum v_i E(Y_i) \\
&= \sum v_i (\beta_0 + \beta_1 X_i) \\
&= \beta_0 \sum v_i + \beta_1 \sum v_i X_i
\end{aligned}
$$

Supaya β^* tidak bias, maka $\sum v_i = 0$ dan $\sum v_i X_i = 1$, dan dapat ditulis:

$$
\begin{aligned}
\operatorname{var}(\beta^*) &= \operatorname{var} \sum v_i Y_i \\
&= \sum v_i^2 \operatorname{var} Y_i \\
&= \sigma^2 \sum v_i^2 \\
&= \sigma^2 \sum \left(v_i - \frac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} + \frac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} \right)^2 \\
&= \sigma^2 \sum \left(v_i - \frac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} \right)^2 + \sigma^2 \frac{\sum (x_i - \bar{x})^2}{\left(\sum (x_i - \bar{x})^2\right)^2} + 2\sigma^2 \sum \left(v_i - \frac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} \right) \left(\frac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} \right) \\
&= \sigma^2 \sum \left(v_i - \frac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} \right)^2 + \sigma^2 \left(\frac{1}{\sum (x_i - \bar{x})^2} \right)
\end{aligned}
$$

Karena hasil terakhir dari persamaan diatas adalah konstan, variansi dari β^* dapat diminimumkan hanya dengan memanipulasi $\operatorname{var} \sum v_i Y_i$, jika dimisalkan $v_i = \dfrac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2}$, persamaan diatas menjadi:

$$
\operatorname{var}(\beta^*) = \frac{\sigma^2}{\sum (x_i - \bar{x})^2} = \operatorname{var}\left(\hat{\beta}\right)
$$

Secara singkat dengan $v_i = k_i$, variansi dari estimator linier β^* sama dengan variansi dari estimator kuadrat terkecil $\hat{\beta}$.

D. Pencilan (*Outlier*)

Pencilan adalah pengamatan yang jauh dari pusat data yang mungkin berpengaruh besar terhadap koefesien regresi. Pencilan dapat muncul karena kesalahan dalam memasukkan data, kesalahan pengukuran, analisis, atau kesalahan-kesalahan lain.

Keberadaan pencilan akan mengganggu dalam proses analisis data dan harus dihindari dalam banyak hal. Dalam kaitannya dengan analisis regresi, pencilan dapat menyebabkan hal-hal berikut (Soemartini, 2007: 7):

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

Pada analisis regresi, terdapat 3 tipe pencilan (*outlier*) yang berpengaruh terhadap estimasi kuadrat terkecil yaitu sebagai berikut (Soemartini, 2007:14):

a. Pencilan vertical (*vertical outlier*)

   Merupakan pengamatan yang terpencil pada variabel dependen (Y), tetapi tidak terpencil pada variabel independen (X). Dalam estimasi kuadrat terkecil, pencilan vertikal sangat berpengaruh khususnya pada estimasi intersep.

b. *Good leverage point*

   Merupakan pengamatan yang terpencil pada variabel X tetapi terletak dekat dengan garis regresi, yang berarti bahwa pengamatan x_i menjauh tetapi y_i cocok dengan garis regresi. *Good leverage* ini tidak berpengaruh terhadap estimasi kuadrat terkecil, tetapi berpengaruh terhadap inferensi statistik karena dapat meningkatkan estimasi standar eror.

c. *Bad leverage point*

   Merupakan pengamatan yang terpencil pada variabel prediktor (X) dan terletak jauh dari garis regresi. *Bad leverage* ini berpengaruh signifikan terhadap estimasi kuadrat terkecil, baik terhadap intersep maupun *slope* dari persamaan regresi.

Metode yang digunakan untuk mengidentifikasi adanya *outlier* yang berpengaruh dalam koefisien regresi antara lain:

1. Metode Grafis

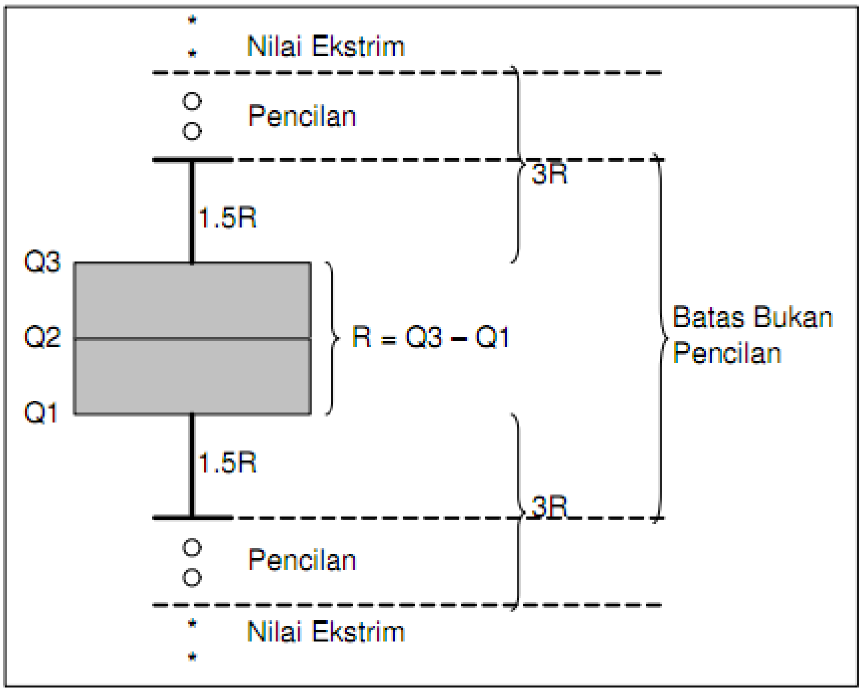
Keuntungan dari metode ini yaitu mudah dipahami karena menampilkan data secara grafis (gambar) dan tanpa melibatkan perhitungan yang rumit. Sedangkan kelemahan metode ini yaitu keputusan yang meperlihatkan data tersebut merupakan pencilan atau tidak bergantung pada kebijakan (*judgement*) peneliti, karena hanya mengandalkan visualisasi gambar.

a. Diagram Pencar (*Scatter Plot*)

Metode ini dilakukan dengan cara memplot data dengan observasi ke-i (i = 1, 2, …, n). Selain itu, jika sudah didapatkan model regresi maka dapat dilakukan dengan cara memplot antara residual (e) dengan nilai prediksi Y ($\widehat{Y}$). Jika terdapat satu atau beberapa data yang terletak jauh dari pola kumpulan data keseluruhan maka hal ini mengindikasikan adanya *outlier*.

b. *Box Plot*

Metode ini mempergunakan nilai kuartil dan jangkauan untuk mendeteksi pencilan. Kuartil 1, 2, dan 3 akan membagi data yang telah diurutkan sebelumnya menjadi empat bagian. Jangkauan (IQR, *Interquartile Range*) didefinisikan sebagai selisih kuartil 1 terhadap kuartil 3, atau IQR = Q3 – Q1. Data-data yang merupakan pencilan yaitu nilai yang kurang dari 1.5*IQR terhadap kuartil 1 dan nilai yang lebih dari 1.5*IQR terhadap kuartil 3.



Gambar 2.1 Skema Identifikasi Data Pencilan Dengan IQR Atau Box Plot

2. Metode *DfFITS* (*Difference fitted value FITS*) atau *Standardized DfFITS*

Metode ini menampilkan nilai perubahan dalam harga yang diprediksi bilamana *case* tertentu dikeluarkan, yang sudah distandarkan.

Perhitungan *DfFITS* adalah sebagai berikut:

$$(DfFITS)_i = t_i \left(\frac{h_{ii}}{1 - h_{ii}} \right)^{\frac{1}{2}}$$

dimana t_i adalah *studentized deleted residual* untuk kasus ke-*i* dan h_{ii} adalah nilai *leverage* untuk kasus ke-*i*.

dengan,

$$t_i = e_i \sqrt{\frac{n - p - 1}{JKG(1 - h_{ii}) - e_i^2}} \ ,$$

e_i adalah residual ke-i dan JKG adalah jumlah kuadrat galat.

Matriks topi:

$$\underset{n \times n}{\boldsymbol{H}} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$$

Elemen diagonal *hii* dalam matriks topi dapat diperoleh langsung dari:

$$h_{ii} = \underset{p \times 1}{\boldsymbol{X_i}} \; \underset{p \times p}{(\boldsymbol{X}'\boldsymbol{X})^{-1}} \; \underset{1 \times p}{\boldsymbol{X}_i'}$$

Suatu data yang mempunyai nilai *absolute DfFITS* lebih besar dari $2\sqrt{p/n}$ maka diidentifikasikan sebagai *outlier*, dengan *p* banyaknya variabel independen dan *n* banyaknya observasi (Montgomery dan Peck, 1982: 184).

E. *Goodness of FIT*

Ketepatan fungsi regresi sampel dalam menaksir nilai aktual dapat diukur dari *Goodness of FIT*nya. Nilai *Goodness of FIT* dapat diukur dari nilai koefisien determinasi (R^2). Koefisien determinasi pada intinya mengukur seberapa jauh kemampuan model dalam menerangkan variasi variabel dependen. Nilai koefisien determinasi adalah antara nol dan satu. Nilai R^2 yang kecil berarti kemampuan variabel-variabel independen dalam menjelaskan variasi variabel dependen amat terbatas, sedangkan jika nilai R^2 mendekati satu berarti variabel-variabel independen memberikan hampir semua informasi yang dibutuhkan untuk memprediksi variasi variabel dependen (Imam Ghozali, 2006: 87).

F. Parameter Lokasi dan Skala

Definisi parameter lokasi dan skala akan digunakan dalam membahas konsep regresi robust dengan estimasi M.

Definisi 2.7 (Bain dan Engelhardt, 1992: 124) parameter η adalah parameter lokasi untuk distribusi dari X jika fungsi distribusi kumulatifnya mempunyai bentuk

$$F(x;\eta) = F_0(x-\eta)$$

Dengan kata lain, fungsi densitas peluangnya berbentuk

$$f(x;\eta) = f_0(x-\eta)$$

Definisi 2.8 (Bain dan Engelhardt, 1992: 126) parameter θ disebut parameter skala untuk distribusi dari X jika fungsi distribusi kumulatifnya mempunyai bentuk

$$F(x;\theta) = F_0(\frac{x}{\theta})$$

Dengan kata lain, fungsi densitas peluangnya berbentuk

$$f(x;\theta) = \frac{1}{\theta} f_0(\frac{x}{\theta})$$

Definisi 2.9 (Bain dan Engelhardt, 1992: 126) parameter η dan θ > 0 disebut parameter lokasi-skala untuk distribusi dari X jika fungsi distribusi kumulatifnya mempunyai bentuk

$$F(x;\theta,\eta) = F_0(\frac{x-\eta}{\theta})$$

Dengan kata lain, fungsi densitas peluangnya berbentuk

$$f(x;\theta,\eta) = \frac{1}{\theta} f_0(\frac{x-\eta}{\theta})$$

G. Metode Maksimum Likelihood

Metode ini merupakan salah satu cara yang digunakan untuk mendapat estimator yang baik dari suatu parameter. Metode maksimum likelihood adalah suatu cara untuk mendapatkan estimator θ yang memaksimalkan fungsi likelihood.

Definisi 2.10 (Bain dan Engelhardt, 1992: 293) fungsi densitas peluang bersama dari n variabel random $X_1,...,X_n$ yang dipandang sebagai fungsi θ disebut fungsi likelihood. Untuk nilai x_1, …, x_n tertentu, fungsi likelihoodnya merupakan fungsi dari θ dan sering dinotasikan dengan $L(\theta)$.

Jika $X_1,...,X_n$ sampel random dengan fungsi densitas peluang $f(x;\theta)$ maka fungsi likelihood $L(\theta)$ didefinisikan sebagai:

$$L(\theta) = f(x_1;\theta)\cdots f(x_n;\theta), \text{ dengan } \theta \text{ parameter yang tidak diketahui}$$

Definisi 2.11 (Bain dan Engelhardt, 1992: 294) misal $L(\theta) = f(x_1;\theta)\cdots f(x_n;\theta)$, $\theta \in \Omega$ merupakan fungsi densitas peluang dari $X_1,...,X_n$.

Diberikan himpunan pengamatan $\{x_1,...,x_n\}$, suatu nilai $\widehat{\theta}$ dalam Ω yang memaksimumkan $L(\theta)$ disebut estimator maksimum likelihood (MLE) dari θ. $\widehat{\theta}$ adalah nilai dari θ yang memenuhi:

$$f(x_1,...,x_n;\widehat{\theta}) = \max_{\theta\in\Omega} f(x_1,...,x_n;\theta)$$

H. Fungsi Obyektif

Fungsi obyektif adalah fungsi yang digunakan untuk mencari fungsi pembobot pada regresi robust.

Fungsi pembobot yang digunakan antara lain adalah (Montgomery dan Peck, 1982: 369):

1. Fungsi pembobot yang disarankan oleh Huber memakai fungsi obyektif

$$\rho(e_i) = \begin{cases} \frac{1}{2}e_i^2, & |e_i| \le c \\ c\,|e_i| - \frac{1}{2}c^2, & |e_i| > c \end{cases}$$

dengan,

$$\psi(e_i) = \rho'(e_i) = \frac{\partial(\rho(e_i))}{\partial e_i} = \begin{cases} e_i, & |e_i| \le c \\ c, & e_i > c \\ -c, & e_i < -c \end{cases}$$

dan fungsi pembobot,

$$w_i = w(e_i) = \frac{\psi(e_i)}{e_i} = \begin{cases} 1, & |e_i| \le c \\ \dfrac{c}{|e_i|}, & |e_i| > c \end{cases}$$

2. Fungsi pembobot yang disarankan oleh Tukey memakai fungsi obyektif

$$\rho(e_i) = \begin{cases} \dfrac{c^2}{6}\left\{1 - \left[1 - \left(\dfrac{e_i}{c}\right)^2\right]^3\right\}, & |e_i| \le c \\ \dfrac{c^2}{6}, & |e_i| > c \end{cases}$$

dengan,

$$\psi(e_i) = \rho'(e_i) = \frac{\partial(\rho(e_i))}{\partial e_i} = \begin{cases} e_i\left[1 - \left(\dfrac{e_i}{c}\right)^2\right]^2, & |e_i| \le c \\ 0, & |e_i| > c \end{cases}$$

dan fungsi pembobot,

$$w_i = w(e_i) = \frac{\psi(e_i)}{e_i} = \begin{cases} \left[1 - \left(\dfrac{e_i}{c}\right)^2\right]^2, & |e_i| \le c \\ 0, & |e_i| > c \end{cases}$$

Secara ringkas, fungsi ρ dan fungsi pembobot dari estimator kuadrat terkecil, Huber, dan Tukey Bisquare dapat dilihat pada Tabel 2.1 (Fox, 2002: 3). Konstanta yang menghasilkan efisiensi tinggi dengan residual berdistribusi normal dan dapat memberikan perlindungan terhadap *outlier* yaitu konstanta dengan nilai c = 1,345 untuk fungsi pembobot Huber dan c = 4,685 untuk pembobot Tukey bisquare.

Tabel 2.1 Fungsi Obyektif dan Fungsi Pembobot untuk Kuadrat Terkecil, Huber, Dan Tukey Bisquare

Metode	Fungsi obyektif	Fungsi pembobot	Interval
Kuadrat terkecil	$\rho(e_i)=\frac{1}{2}e_i^2$	$w(e_i)=1$	$\lvert e_i\rvert<\infty$
Huber	$\rho(e_i)=\begin{cases}\frac{1}{2}e_i^2\\ c\lvert e_i\rvert-\frac{1}{2}c^2\end{cases}$	$w(e_i)=\begin{cases}1\\ \frac{c}{\lvert e_i\rvert}\end{cases}$	$\lvert e_i\rvert\le c$ $\lvert e_i\rvert>c$
Tukey bisquare	$\rho(e_i)=\begin{cases}\frac{c^2}{6}\left\{1-\left[1-\left(\frac{e_i}{c}\right)^2\right]^3\right\}\\ \frac{c^2}{6}\end{cases}$	$w(e_i)=\begin{cases}\left[1-\left(\frac{e_i}{c}\right)^2\right]^2\\ 0\end{cases}$	$\lvert e_i\rvert\le c$ $\lvert e_i\rvert>c$

I. Breakdown Point

Breakdown point adalah salah satu cara yang digunakan untuk mengukur ke-*robust*-an (kekekaran) suatu estimator (Yohai, 2003: 11). *Breakdown point* merupakan proporsi minimal dari banyaknya *outlier* dibandingkan seluruh data pengamatan. Regresi *robust* yang mempunyai *breakdown point* adalah regresi *robust* dengan metode estimasi S, LTS, LMS, dan MM. Metode estimasi MM mempunyai *breakdown point* 50%. *Breakdown point* 50% adalah *breakdown point* yang tinggi.

BAB III
PEMBAHASAN

Cara untuk mengestimasi parameter regresi adalah menggunakan metode kuadrat terkecil, tetapi apabila data tidak normal dan terkontaminasi pencilan maka metode ini tidak bekerja dengan baik. Metode yang lain yang dapat mengatasi pencilan adalah regresi *robust*.

A. Regresi *Robust*

Regresi *robust* merupakan metode yang dapat mengatasi pencilan tanpa menghapus data pencilan tersebut. Regresi robust bertindak sebagai penurun bobot data pencilan.

Dalam mendeteksi pencilan, metode regresi *robust* yang sering digunakan adalah Huber estimasi M, estimasi dengan nilai *breakdown* tinggi, dan gabungan dari dua metode tersebut.

Menurut Chen (2002:1) metode-metode estimasi dalam regresi robust diantaranya adalah:

1. Estimasi M (*Maximum likelihood type*) yang dikenalkan oleh Huber (1973) adalah metode yang sederhana baik dalam penghitungan maupun secara teoritis. Estimasi ini menganalisis data dengan mengasumsikan bahwa sebagian besar yang terdeteksi pencilan pada variabel independen.
2. Estimasi LTS (*Least Trimmed Squares*) adalah metode dengan *high breakdown point* yang dikenalkan oleh Rousseeuw (1984). *Breakdown point* adalah ukuran proporsi minimal dari banyaknya data yang terkontaminasi pencilan dibandingkan seluruh data pengamatan.

3. Estimasi S (*Scale*) juga merupakan metode dengan *high breakdown point* yang dikenalkan oleh Rousseeuw and Yohai (1984). Dengan nilai *breakdown* yang sama, metode ini mempunyai efisiensi yang lebih tinggi dibanding estimasi LTS.
4. Estimasi MM (*Method of Moment*), dikenalkan oleh Yohai (1987). Metode ini menggabungkan estimasi S (estimasi dengan *high breakdown point*) dan estimasi M.

B. Estimasi M (*Maximum likelihood type*)

Metode ini merupakan metode yang paling sederhana dan sering digunakan. Estimasi M akan menjaga ke-*robust*-an dengan mengatasi pencilan vertikal. Estimator M yang meminimumkan fungsi ρ (fungsi obyektif) dari residualnya (Montgomery dan peck, 1982: 367):

$$\min_{\beta} \sum_{i=1}^{n} \rho(e_i) = \min_{\beta} \sum_{i=1}^{n} \rho\left(y_i - \sum_{j=0}^{k} x_{ij}\beta_j \right) \qquad (3.1)$$

Dalam mengestimasi parameter regresi robust M metode iterasi diperlukan, karena residual tidak dapat dihitung sampai diperoleh model yang cocok dan parameter regresi juga tidak dapat dihitung tanpa mengetahui nilai residual. *Iteratively reweighted least squares* (IRLS) adalah metode iterasi yang banyak digunakan.

C. Estimasi S (*Scale*)

Jika data terkontaminasi pencilan pada variabel *X* (prediktor), estimasi M tidak dapat bekerja dengan baik. Estimasi M tidak dapat mengidentifikasi *bad observation* yang berarti tidak dapat membedakan *good leverage point* dan *bad* *leverage point*. Untuk mengatasi hal tersebut, estimasi *high breakdown* sangat diperlukan (Chen, 2002:5). Salah satu estimasi yang mempunyai nilai *high breakdown* adalah estimasi S. Bentuk estimator S adalah:

$$\tilde{\beta}_S = \arg\min_{\beta} \hat{\sigma}_S(e_1, e_2, ..., e_n) \qquad (3.2)$$

dimana $\hat{\sigma}_S$ adalah estimator skala *robust* yang memenuhi $\frac{1}{n}\sum_{i=1}^{n}\rho\left(\frac{e_i}{\hat{\sigma}_S}\right) = b$.

dengan b konstan yang didefinisikan $b = E[\Phi(\rho)]$, Φ adalah distribusi normal standar.

Estimator S mempunyai nilai *breakdown* tinggi yaitu 50%. Nilai *breakdown* dari estimator S dapat ditulis $\frac{b}{\max \rho(e)} = 0,5$.

D. Estimasi MM (*Method of Moment*)

Estimasi MM menggabungkan estimasi *high breakdown point* dan efisiensi statistik yang dikenalkan oleh Yohai (1987). Langkah pertama dalam estimasi ini adalah mencari estimator S, kemudian menetapkan parameter-parameter regresi menggunakan estimasi M. Estimasi S menjamin nilai *breakdown point* tinggi dan estimasi M membuat estimator mempunyai efisiensi tinggi. Pada umumnya digunakan fungsi Tukey Bisquare β baik pada estimasi S maupun estimasi M. Bentuk dari metode estimasi MM:

$$\tilde{\beta}_{MM} = \arg\min_{\beta} \sum_{i=1}^{n} \rho\left(\frac{e_i}{\hat{\sigma}_S}\right) = \arg\min_{\beta} \sum_{i=1}^{n} \rho\left(\frac{y_i - \sum_{j=0}^{k} x_{ij}\beta_j}{\hat{\sigma}_S}\right) \qquad (3.3)$$

Metode MM juga menggunakan IRLS (*Iteratively Reweighted Least Square*) untuk mencari estimasi parameter regresi.

Prosedur estimasi parameter pada model regresi linier ganda dengan regresi *robust* estimasi MM:

1. Menghitung estimator awal koefisien $\hat{\beta}_j^{(1)}$ dan residual $e_i^{(1)}$ dari regresi *robust* dengan *high breakdown point* (estimasi S) dengan bobot huber / bisquare (dilihat sebagai bentuk estimasi M).
2. Residual $e_i^{(1)}$ pada langkah pertama digunakan untuk menghitung skala estimasi $\hat{\sigma}_s$ dan dihitung pula pembobot awal $w_i^{(1)}$.
3. Residual $e_i^{(1)}$ dengan skala estimasi $\hat{\sigma}_s$ pada langkah kedua digunakan dalam iterasi awal sebagai penaksir WLS untuk menghitung koefisien regresi.

   $\sum_{i=1}^{n} w_i^{(1)} \left(\frac{e_i^{(1)}}{\hat{\sigma}_s} \right) x_i = 0$, $w_i^{(1)}$ merupakan pembobot Huber/bisquare.
4. Menghitung bobot baru $w_i^{(2)}$ dengan skala estimasi dari iterasi awal WLS.
5. Mengulang langkah 2, 3, 4 (dengan skala estimasi tetap konstan) sampai mendapatkan $\sum_{i=1}^{n} | e_i^{(m)} |$ konvergen (selisih $\hat{\beta}_j^{(m+1)}$ dan $\hat{\beta}_j^{(m)}$ mendekati 0, dengan m banyaknya iterasi).

E. Penyelesaian Untuk $\tilde{\beta}$

Untuk meminimumkan ρ (fungsi obyektif) dari residualnya, dicari turunan parsial pertama dari ρ terhadap β_j, $j = 0,1,2,...,k$ dan disama dengan 0. Ini memberikan $p = k + 1$ sistem persamaan

$$\sum_{i=1}^{n} x_{ij} \psi \left[\frac{y_i - \sum_{j=0}^{k} x_{ij} \beta_j}{\hat{\sigma}_s} \right] = 0 \qquad (3.4)$$

dengan $\psi = \rho'$ dan ψ merupakan fungsi *influence* yang digunakan dalam memperoleh bobot, x_{ij} adalah observasi ke-*i* pada regressor ke-*j* dan $x_{i0} = 1$.

Didefinisikan suatu fungsi pembobot:

$$w(e_i) = \frac{\psi \left(\frac{y_i - \sum_{j=0}^{k} x_{ij} \beta_j}{\hat{\sigma}_s} \right)}{\frac{y_i - \sum_{j=0}^{k} x_{ij} \beta_j}{\hat{\sigma}_s}} \qquad (3.5)$$

dan misal $w_i = w(e_i)$, maka persamaan (3.4) dapat ditulis :

$$\sum_{i=1}^{n} x_{ij} w_i \left[y_i - \sum_{j=0}^{k} x_{ij} \beta_j \right] = 0 , \; j = 0,1,2,...,k \qquad (3.6)$$

Persamaan (3.6) diselesaikan dengan IRLS, estimasi awal koefisien $\tilde{\beta}^{(1)}$ dan residual $e_i^{(1)}$ diambil dari regresi robust dengan *high breakdown point* (estimasi S), untuk bobot permulaan $w_i^{(1)} = w(e_i^{(1)})$, maka $p = k+1$ persamaan (3.6) ditulis:

$$\sum_{i=1}^{n} x_{ij} w_i^{(1)} \left[y_i - \sum_{j=0}^{k} x_{ij} \beta_j \right] = 0 \qquad (3.7)$$

dimana,

$$w_i^{(1)} = \begin{cases} \dfrac{\psi\left(\dfrac{y_i - \sum_{j=0}^{k} x_{ij} \widehat{\beta}_j}{\widehat{\sigma}_s} \right)}{\dfrac{y_i - \sum_{j=0}^{k} x_{ij} \widehat{\beta}_j}{\widehat{\sigma}_s}}, & \text{jika } y_i \neq \sum_{j=1}^{k} x_{ij} \widehat{\beta}_j^{(1)} \\ 1 & \text{, jika } y_i = \sum_{j=1}^{k} x_{ij} \widehat{\beta}_j^{(1)} \end{cases}$$

Untuk regresi berganda, persamaan (3.6) menjadi:

$$\sum_{i=1}^{n} x_i w_i^{(1)} \left[y_i - \sum_{j=0}^{k} x_{ij} \widehat{\beta}_j \right] = 0$$

$$\sum_{i=1}^{n} w_i^{(1)} x_i y_i - \sum_{i=1}^{n} \sum_{j=0}^{k} x_{ij}^{2} w_i^{(1)} \widehat{\beta}_j = 0$$

$$\sum_{i=1}^{n} \sum_{j=0}^{k} x_{ij}^{2} w_i^{(1)} \widehat{\beta}_j = \sum_{i=1}^{n} w_i^{(1)} x_i y_i$$

Dalam bentuk matriks dapat ditulis:

$$\underset{p \times n}{\boldsymbol{X}'} \; \underset{n \times n}{\boldsymbol{W}^{(1)}} \; \underset{n \times p}{\boldsymbol{X}} \; \underset{p \times 1}{\boldsymbol{\beta}} = \underset{p \times n}{\boldsymbol{X}'} \; \underset{n \times n}{\boldsymbol{W}^{(1)}} \; \underset{n \times 1}{\boldsymbol{y}}$$

Dimana $\mathbf{W}^{(1)}$ adalah matriks diagonal yang berukuran n x n dengan elemen diagonalnya $w_1, w_2, \ldots, w_n$ (n banyaknya observasi).

Estimator satu langkah dapat ditulis:

$$\boldsymbol{X}'\boldsymbol{W}^{(1)}\boldsymbol{X}\widetilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(2)} = \boldsymbol{X}'\boldsymbol{W}^{(1)}\boldsymbol{y}$$

$$\left(\boldsymbol{X}'\boldsymbol{W}^{(1)}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{W}^{(1)}\boldsymbol{X}\widetilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(2)} = \left(\boldsymbol{X}'\boldsymbol{W}^{(1)}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{W}^{(1)}\boldsymbol{y}$$

$$\underset{p\times 1}{\tilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\mathbf{2})}} = \underset{p\times p}{\left(\boldsymbol{X}'\boldsymbol{W}^{(\mathbf{1})}\boldsymbol{X}\right)^{-\mathbf{1}}} \underset{p\times 1}{\boldsymbol{X}'\boldsymbol{W}^{(\mathbf{1})}\boldsymbol{y}}$$

Pada langkah selanjutnya dihitung kembali bobot $w_i^{(2)}$ menggunakan $\hat{\beta}_j^{(2)}$ dan skala parameter $\hat{\sigma}_s$. Untuk $w^{(m)}$ bobot yang diberikan, dapat diperoleh estimator $\hat{\boldsymbol{\beta}}^{(\boldsymbol{m}+\mathbf{1})} = \left(\boldsymbol{X}'\boldsymbol{W}^{(\boldsymbol{m})}\boldsymbol{X}\right)^{-\mathbf{1}}\boldsymbol{X}'\boldsymbol{W}^{(\boldsymbol{m})}\boldsymbol{y}$ sampai $\sum_{i=0}^{n} |e_i^{(m)}|$ konvergen (selisih nilai $\hat{\beta}^{(m+1)}$ dan $\hat{\beta}^{(m)}$ mendekati 0), dengan *m* banyaknya iterasi.

F. Contoh Ilustrasi Kasus

Contoh Kasus I :

Sebagai contoh ilustrasi kasus I dalam pembahasan ini adalah data yang yang diambil dari buku Model Linier Terapan (Buku II Analisis Regresi Ganda) karangan J. Netter, W. Wasserman, M. H.Kutner yang diterjemahkan oleh Bambang Sumantri (1997: 39).

Data ini merupakan data penelitian yang dilakukan di sebuah yayasan ilmiah. Peneliti ingin mengevaluasi hubungan antara gaji tahunan matematikawan (Y, dalam ribuan dolar) dan indeks mutu publikasi (X_1), lama pengalaman (X_2, dalam tahun), dan indeks keberhasilan dalam memperoleh dukungan dana (X_3). Data ditunjukkan pada table 3.1.

Tabel 3.1. Data Gaji matematikawan

No.	X1	X2	X3	Y
1	3,5	9	6,1	33,2
2	5,3	20	6,4	40,3
3	5,1	18	7,4	38,7
4	5,8	33	6,7	46,8
5	4,2	31	7,5	41,4

No.	X1	X2	X3	Y
6	6	13	5,9	37,5
7	6,8	25	6	39
8	5,5	30	4	40,7
9	3,1	5	5,8	30,1
10	7,2	47	8,3	52,9
11	4,5	25	5	38,2
12	4,9	11	6,4	31,8
13	8	23	7,6	43,3
14	6,5	35	7	44,1
15	6,6	39	5	42,8
16	3,7	21	4,4	33,6
17	6,2	7	5,5	34,2
18	7	40	7	48
19	4	35	6	38
20	4,5	23	3,5	35,9
21	5,9	33	4,9	40,4
22	5,6	27	4,3	36,8
23	4,8	34	8	45,2
24	3,9	15	5	35,1

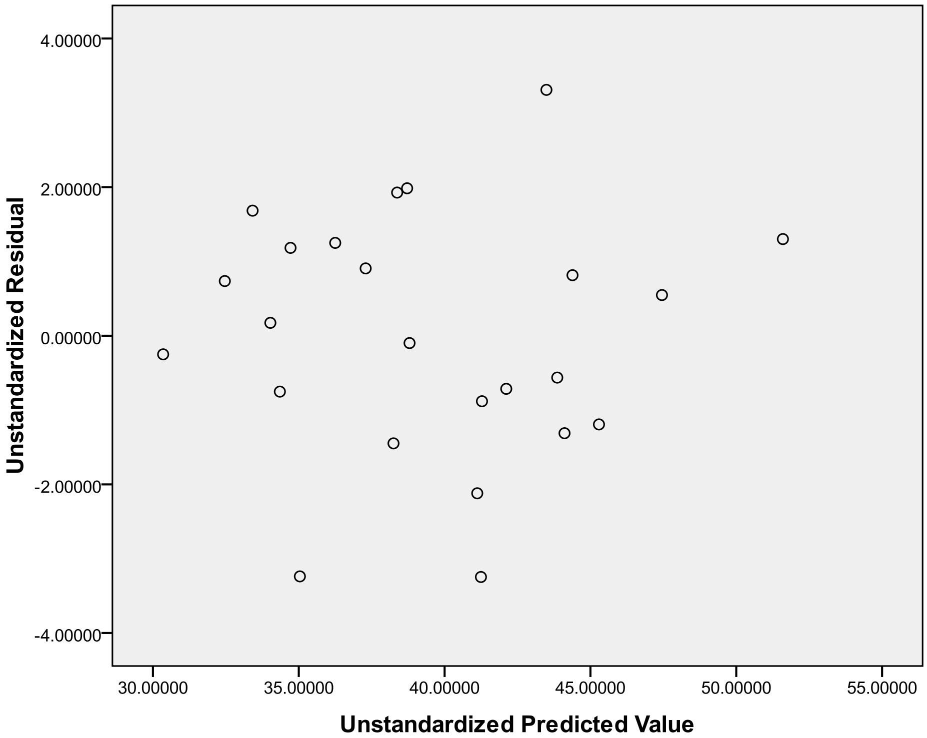
Untuk menganalisis data pada tabel 3.1 langkah pertama yaitu melakukan identifikasi pencilan untuk mengetahui apakah data mengandung pencilan atau tidak.

A. Identifikasi Pencilan

Mengidentifikasi suatu pencilan dapat dilakukan dengan cara sebagai berikut:

a. Diagram Pencar (*Scatter Plot*)

Berdasarkan *output* SPSS 17.0 didapat plot antara residual (e) dengan nilai prediksi Y ($\hat{Y}$) sebagai berikut:



Gambar 3.1. Scatterplot antara Residual (E) dengan Nilai Prediksi Y ($\widehat{Y}$)

Gambar 3.1 memperlihatkan bahwa ada beberapa data yang terletak jauh dari kumpulan data. Data tersebut yang disebut dengan pencilan (*outlier*). Untuk lebih jelasnya data mana yang teridentifikasi pencilan dapat dilihat pada hasil *DfFITS.*

b. *DfFITS*

Nilai *DfFITS* yang diidentifikasi sebagai pencilan adalah data yang nilai *DfFITS*-nya lebih besar dari $2\sqrt{p/n}$ = $2\sqrt{3/24}$ = 0,7071

Tabel 3.2. Nilai *DfFITS*

No.	DfFITS	|DfFITS|
1	0,16568	0,16568
2	0,12072	0,12072
3	-0,01495	0,01495
4	0,25089	0,25089

No.	**DfFITS**	**|DfFITS|**
5	-0,19438	0,19438
6	0,21395	0,21395
7	-0,27459	0,27459
8	0,43306	0,43306
9	-0,07933	0,07933
10	0,5265	0,5265
11	0,08223	0,08223
12	-0,47508	0,47508
13	-0,26549	0,26549
14	-0,12901	0,12901
15	-0,29873	0,29873
16	-0,13389	0,13389
17	0,06342	0,06342
18	0,094	0,094
19	-0,80108	**0,80108**
20	0,30735	0,30735
21	-0,11736	0,11736
22	-0,22757	0,22757
23	0,23619	0,23619
24	0,20849	0,20849

Berdasarkan tabel 3.2, data yang diindikasikan sebagai pencilan (yang dicetak tebal) yaitu data ke 19.

Data ternyata teridentifikasi pencilan, metode yang bisa digunakan untuk mengestimasi parameter yaitu regresi robust estimasi MM atau metode kuadrat terkecil dengan menghapus pencilan tersebut.

B. Estimasi dengan regresi *robust* MM

Estimasi MM adalah gabungan dari metode estimasi S dan estimasi M. Langkah pertama dalam metode estimasi MM yaitu mencari estimator S, kemudian menetapkan parameter-parameter regresi menggunakan metode estimasi M. Dengan bantuan software SAS 9.1 didapat:

Tabel 3.3. Hasil Estimasi Regresi Robust MM

The ROBUSTREG Procedure

Model Information

Data Set	WORK.GAJI
Dependent Variable	y
Number of Independent Variables	3
Number of Observations	24
Method	MM Estimation

Number of Observations Read	24
Number of Observations Used	24

Parameter Estimates

Parameter	DF	Estimate
Intercept	1	18.1903
x1	1	1.0284
x2	1	0.3188
x3	1	1.3196

Goodness-of-Fit

Statistic	Value
R-Square	0.7289

Berdasarkan output diatas, terlihat bahwa nilai R^2 (Koefisien determinasi) adalah 0,7289. Nilai tersebut mendekati satu sehingga dapat disimpulkan bahwa variabel-variabel independen pada contoh I memberikan hampir semua informasi yang dibutuhkan untuk memprediksi variasi variabel dependen.

Sehingga didapatkan persamaan modelnya adalah

$$\hat{Y} = 18{,}1903 + 1{,}0284\, X_1 + 0{,}3188\, X_2 + 1{,}3196\, X_3$$

dengan,

Y = Gaji tahunan matematikawan

X_1 = Indeks mutu publikasi

X_2 = Lama pengalaman (dalam tahun)

X_3 = Indeks keberhasilan dalam memperoleh dukungan dana

Makna dari model persamaan diatas adalah sebagai berikut:

- Setiap penambahan satu satuan indeks mutu publikasi (X_1) akan meningkatkan rata-rata gaji tahunan matematikawan sebesar 1,0284 apabila lama pengalaman (X_2), dan indeks keberhasilan (X_3) tetap.
- Setiap penambahan 1 tahun lama pengalaman (X_2) akan meningkatkan rata-rata gaji tahunan matematikawan sebesar 0,3188 apabila indeks mutu publikasi (X_1), dan indeks keberhasilan (X_3) tetap.
- Setiap penambahan satu satuan indeks keberhasilan (X_3) akan meningkatkan rata-rata gaji tahunan matematikawan sebesar 1,3196 apabila indeks mutu publikasi (X_1), dan lama pengalaman (X_2) tetap.
- Jika indeks mutu publikasi (X_1), lama pengalaman (X_2), dan indeks keberhasilan (X_3) sama dengan 0, maka rata-rata gaji tahunan matematikawan sebesar 18,1903.

Selain dengan regresi *robust* MM, cara lain digunakan untuk mengetimasi parameter adalah metode kuadrat terkecil (MKT) dengan data pencilan dihapus.

Tabel 3.4. Hasil Estimasi MKT dengan pencilan dihapus

Parameter Estimates

Variable	DF	Parameter Estimate
Intercept	1	18.58745
x1	1	0.84032
x2	1	0.34892
x3	1	1.31441

Berdasarkan tabel 3.3 dan tabel 3.4, terlihat bahwa hasil estimasi untuk metode MM dan metode kuadrat terkecil dengan data pencilan dihapus hampir sama nilainya. Tetapi menghapus data pencilan bukanlah tindakan yang baik, karena adakalanya data yang mengandung pencilan merupakan data yang berpengaruh terhadap keseluruhan data, selain itu juga dengan mengahapus sebagian data berarti mengubah data asli yang sudah ada yang mungkin dapat memberikan resiko kesalahan yang besar pada hasil estimasi. Dengan begitu, metode estimasi MM merupakan metode yang digunakan untuk data yang mengandung pencilan tanpa menghapus data pencilan tersebut.

Contoh ilustrasi kasus II

Sebagai contoh ilustrasi kasus II adalah data yang berupa 20 sampel potongan kayu pinus yang dipotong melintang dengan ketebalan yang sama. Pada penelitian tersebut akan diteliti berat jenis potongan kayu pinus tersebut. Data terdiri dari X_1, X_2, X_3, X_4, X_5 secara berurutan adalah serat kayu pinus (mm^2), kecepatan tumbuh (mm), kelembaban tanah (%), penyerapan cahaya pada kayu pinus (%), dan kadar air pada kayu (%), dan respon Y adalah berat jenis kayu. Data merupakan data yang dikarang oleh penulis.

Data ditunjukkan pada tabel 3.9. dari data tersebut akan dicari model regresi terbaiknya.

Tabel 3.5. Data Faktor Anatomi Dan Berat Jenis Potongan Kayu Pinus

No	X_1	X_2	X_3	X_4	X_5	Y
1	573	1059	46,5	53,8	84,1	0,534
2	651	1356	52,7	54,5	88,7	0,535
3	606	1273	49,4	52,1	92	0,57
4	630	1151	48,9	50,3	87,9	0,528
5	547	1135	53,1	51,9	91,5	0,548
6	557	1236	54,9	55,2	91,4	0,555
7	489	1231	56,2	45,5	82,4	0,481

No	X_1	X_2	X_3	X_4	X_5	Y
8	685	1564	56,6	44,3	91,3	0,516
9	536	1182	59,2	46,4	85,4	0,475
10	685	1564	63,1	56,4	91,4	0,486
11	664	1588	50,6	48,1	86,7	0,554
12	703	1335	51,9	48,4	81,2	0,519
13	653	1395	62,5	51,9	89,2	0,492
14	586	1114	50,5	56,5	88,9	0,517
15	534	1143	52,1	57	88,9	0,502
16	523	1320	50,5	61,2	91,9	0,508
17	580	1249	54,6	60,8	95,4	0,52
18	448	1028	52,2	53,4	91,8	0,506
19	476	1057	42,9	53,2	92,9	0,595
20	528	1057	42,4	56,6	90	0,568

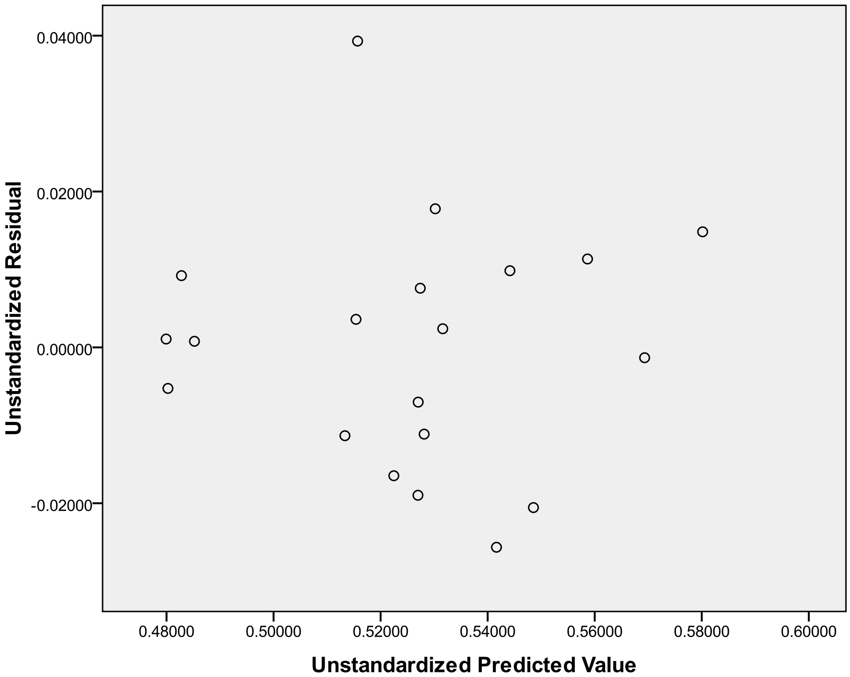
Untuk menganalisis data pada tabel 3.9, langkah pertama yaitu melakukan identifikasi pencilan untuk mengetahui apakah data mengandung pencilan atau tidak.

A. Identifikasi Pencilan

Mengidentifikasi suatu pencilan dapat dilakukan dengan cara sebagai berikut:

a. Diagram Pencar (*Scatter Plot*)

Berdasarkan *output* Minitab didapat plot antara residual (*e*) dengan nilai prediksi Y ($\widehat{Y}$) sebagai berikut:



Gambar 3.2. Scatterplot antara Residual (E) dengan Nilai Prediksi Y ($\widehat{Y}$)

Gambar 3.4 memperlihatkan bahwa ada beberapa data yang terletak jauh dari kumpulan data. Data tersebut yang disebut dengan pencilan. Untuk lebih jelasnya data mana yang teridentifikasi pencilan dapat dilihat pada hasil *DfFITS*.

b. *DfFITS*

Nilai *DfFITS* yang diidentifikasi sebagai pencilan adalah data yang nilai *DfFITS*-nya lebih besar dari $2\sqrt{p/n}$ = 1,09545.

Tabel 3.6. Nilai *DfFITS*

No.	DfFITS	| DfFITS|
1	0,09817	0,09817
2	0,16195	0,16195
3	0,30116	0,30116
4	-0,80356	0,80356
5	0,49159	0,49159
6	0,90684	0,90684

No.	DfFITS	| DfFITS|
7	0,09695	0,09695
8	-2,39149	**2,39149**
9	-0,24917	0,24917
10	0,04253	0,04253
11	0,76534	0,76534
12	0,21751	0,21751
13	0,34570	0,34570
14	-0,29539	0,29539
15	-0,26517	0,26517
16	-1,44988	**1,44988**
17	-0,27626	0,27626
18	-0,66639	0,66639
19	0,77635	0,77635
20	-0,04615	0,04615

Berdasarkan tabel 3.10, data ke 8 dan 16 diindikasikan sebagai pencilan karena mempunyai nilai yang lebih besar dari 1,09545.

Apabila terdapat pencilan metode yang bisa digunakan yaitu regresi *robust* estimasi MM dan metode kuadrat terkecil dengan menghapus pencilan tersebut.

B. Estimasi dengan regresi *robust* MM

Estimasi MM adalah gabungan dari estimasi dengan *high breakdown* (estimasi S) dan estimasi M. Langkah pertama dalam metode estimasi MM yaitu mencari estimator S, kemudian menetapkan parameter-parameter regresi menggunakan metode estimasi M. Dengan bantuan software SAS 9.1 didapat:

Tabel 3.7. Hasil Estimasi Regresi Robust MM

```
                    The SAS System

                The ROBUSTREG Procedure

                  Model Information

Data Set                                  WORK.CON2
Dependent Variable                                y
Number of Independent Variables                   5
Number of Observations                           20
Method                                MM Estimation


     Number of Observations Read          20
     Number of Observations Used          20

                 Parameter Estimates
                 Parameter DF Estimate

                 Intercept  1   0.4448
                 x1         1   0.0001
                 x2         1   0.0000
                 x3         1  -0.0056
                 x4         1  -0.0024
                 x5         1   0.0047
                   Goodness-of-Fit

               Statistic        Value

               R-Square        0.6638
```

Berdasarkan output diatas, terlihat bahwa nilai R^2 (Koefisien determinasi) adalah 0,6638. Nilai tersebut mendekati satu sehingga dapat disimpulkan bahwa variabel-variabel independen pada contoh II memberikan hampir semua informasi yang dibutuhkan untuk memprediksi variasi variabel dependen.

Sehingga didapatkan persamaan modelnya adalah

$$\widehat{Y} = 0,4448 + 0,0001\,X_1 - 0,0056\,X_3 - 0,0024\,X_4 + 0,0047\,X_5$$

dengan,

Y = Berat jenis kayu pinus

X_1 = Serat kayu pinus (mm^2)

X_2 = Kecepatan tumbuh (mm)

X_3 = Kelembaban tanah (%)

X_4 = Penyerapan cahaya pada kayu pinus (%)

X_5 = Kadar air pada kayu (%)

Makna dari model persamaan diatas adalah sebagai berikut:

- Setiap penambahan 1 mm^2 serat kayu pinus (X_1) akan meningkatkan rata-rata berat jenis potongan kayu pinus sebesar 0,0001, jika kecepatan tumbuh (X_2), kelembaban tanah (X_3), penyerapan cahaya pada kayu pinus (X_4), dan kadar air pada kayu (X_5) tetap.
- Setiap penambahan 1 mm pertumbuhan pinus (X_2) tidak membuat rata-rata berat jenis potongan kayu pinus berubah karena nilai estimasinya nol.
- Setiap penambahan 1 persen kelembaban tanah (X_3) akan menurunkan rata-rata berat jenis potongan kayu pinus sebesar 0,0056 apabila serat kayu pinus (X_1), kecepatan tumbuh (X_2), penyerapan cahaya pada kayu pinus (X_4), dan kadar air pada kayu (X_5) tetap.
- Setiap penambahan 1 persen penyerapan cahaya pada kayu pinus (X_4) akan menurunkan rata-rata berat jenis potongan kayu pinus sebesar 0,0024 apabila serat kayu pinus (X_1), kecepatan tumbuh (X_2), kelembaban tanah (X_3), dan kadar air pada kayu (X_5) tetap.
- Setiap penambahan 1 persen kadar air pada kayu (X_5) akan meningkatkan rata-rata berat jenis potongan kayu pinus sebesar 0,0047 apabila serat kayu pinus (X_1), kecepatan tumbuh (X_2), kelembaban tanah (X_3), dan penyerapan cahaya pada kayu pinus (X_4) tetap.

- Jika serat kayu pinus (X_1), kecepatan tumbuh (X_2), kelembaban tanah (X_3), penyerapan cahaya pada kayu pinus (X_4), dan kadar air pada kayu (X_5) sama dengan 0, maka berat jenis potongan kayu pinus adalah 0,4448. Dalam hal ini berarti tidak bermakna.

Selain dengan regresi robust MM, cara lain digunakan untuk mengetimasi parameter adalah metode kuadrat terkecil (MKT) dengan data pencilan dihapus.

Tabel 3.8. Hasil Estimasi MKT dengan pencilan dihapus

Parameter Estimates

Variable	DF	Parameter Estimate
Intercept	1	0.38038
x1	1	0.00007251
x2	1	0.00004751
x3	1	-0.00579
x4	1	-0.00260
x5	1	0.00548

Berdasarkan tabel 3.7 dan tabel 3.8, terlihat bahwa hasil estimasi untuk metode MM dan metode kuadrat terkecil dengan data pencilan dihapus tidak jauh berbeda. Tetapi menghapus data pencilan bukanlah tindakan yang baik, karena adakalanya data yang mengandung pencilan merupakan data yang berpengaruh terhadap keseluruhan data, selain itu juga dengan mengahapus sebagian data berarti mengubah data asli yang sudah ada yang mungkin dapat memberikan resiko kesalahan yang besar pada hasil estimasi. Dengan demikian metode yang digunakan untuk mendapatkan hasil estimasi yang baik dan bersifat *robust* pada data yang mengandung pencilan adalah metode estimasi MM.

BAB IV
PENUTUP

A. Kesimpulan

Dari hasil pembahasan dapat diambil kesimpulan sebagai berikut:

1. Prosedur estimasi parameter pada model regresi linier ganda dengan regresi *robust* estimasi MM adalah sebagai berikut:

   a. Menghitung estimasi awal koefisien $\hat{\beta}_j^{(1)}$ dan residual $e_i^{(1)}$ dari regresi robust dengan *high breakdown point* (estimasi S) dengan bobot huber / bisquare (dilihat sebagai bentuk estimasi M).

   b. Residual $e_i^{(1)}$ pada langkah pertama digunakan untuk menghitung skala estimasi $\hat{\sigma}_s$ dan dihitung pula pembobot awal $w_i^{(1)}$.

   c. Residual $e_i^{(1)}$ dengan skala estimasi $\hat{\sigma}_s$ pada langkah kedua digunakan dalam iterasi awal sebagai penaksir WLS untuk menghitung koefisien regresi.

   $\sum_{i=1}^{n} w_i^{(1)} \left(\frac{e_i^{(1)}}{\hat{\sigma}_s} \right) x_i = 0$, $w_i^{(1)}$ merupakan pembobot Huber/bisquare.

   d. Menghitung bobot baru $w_i^{(2)}$ dengan skala estimasi dari iterasi awal WLS.

   e. Mengulang langkah 2, 3, 4 (dengan skala estimasi tetap konstan) sampai mendapatkan $\sum_{i=1}^{n} | e_i^{(m)} |$ konvergen (selisih $\hat{\beta}_j^{(m+1)}$ dan $\hat{\beta}_j^{(m)}$ mendekati 0, dengan m banyaknya iterasi).

2. Dalam penulisan ini data yang digunakan adalah data regresi linier berganda dan data yang teridentifikasi pencilan. Contoh kasus pertama mengenai

hubungan antara gaji tahunan matematikawan dengan indeks mutu publikasi, lama pengalaman, dan indeks keberhasilan dalam memperoleh dukungan dana dan pada kasus kedua mengenai hubungan antara berat jenis kayu pinus dengan serat kayu pinus, kecepatan tumbuh, kelembaban tanah, penyerapan cahaya pada kayu pinus, dan kadar air pada kayu. Hasil pada kedua contoh menunjukkan bahwa regresi *robust* estimasi MM menghasilkan persamaan regresi yang tidak jauh berbeda dengan MKT dengan data pencilan yang dihapus. Tetapi MKT tidak baik digunakan dalam kasus yang mengandung pencilan karena menghapus data tidaklah baik, dengan menghapus sebagian data berarti akan mengubah data asli. Jika menggunakan regresi robust estimasi MM, data pencilan tidak dihapus sehingga dapat mengestimasi dengan tetap menggunakan data asli. Dengan begitu regresi *robust* estimasi MM adalah alternatif yang tepat untuk data yang mengandung pencilan.

B. Saran

1. Dalam penulisan skripsi ini metode regresi robust yang digunakan adalah estimasi MM. Oleh karena itu bagi yang berminat untuk membahas regresi robust dapat menggunakan estimasi lain seperti estimasi S, LTS dan LMS.
2. Dalam penulisan ini untuk mendapatkan hasil estimasi dibantu dengan software SAS 9.1, tetapi software lain juga bisa digunakan seperti software S-PLUS.

DAFTAR PUSTAKA

Bain, J. L. & Engelhardt, M. (1992). *Introduction To Probability And Mathematical Statistics*. 2nd. ed. California: Duxbury Press.

Chen, C. (2002). "Robust Regression and Outlier Detection with the ROBUSTREG Procedure". SUGI Paper 265-27. North Carolina: SAS Institute. www.sas.com.

Copt, S. & Heritier, S. (2006). "Robust MM-Estimation and Inference in Mixed Linear Models". NHMRC Clinical Trials Centre, University of Sidney. http://www.unige.ch/ses/metri/.

Draper, R. N. & Smith, H. (1998). *Applied Regression Analysis*. New York: John Wiley & Sons Inc.

Fox, J. (2002). "Robust Regression. Appendix to An R and S-Plus Companion to Applied Regression". January, 2002.

Imam Ghozali. (2006). *Aplikasi Analisis Multivariat dengan Program SPSS*. Semarang: Badan Penerbit Universitas Diponegoro.

Greene, W. H. (2000). *Econometrics Analysis*. 4th. ed. New Jersey: Prentice Hall.

Gujarati, D. N. (2004). *Basic Econometrics*. 4th. ed. New York: McGraw-Hill.

Huber, P. J. (1973). "Robust Regression: Asymptotics, conjectures and Monte Carlo", *Ann. Stat*., Vol. 1, No.5, 799-821.

Khattree, R. & Naik, N. D. (1999). *Applied Multivariate Statistics With SAS Software*. 2nd. ed. North Carolina: SAS Institute Inc.

Maronna, A. R., Martin, D. R., & Yohai, J. V. (2006). *Robust Statistics Theory And Methods*. San Francisco: John Wiley & Sons Inc.

Montgomery, C. D., & Peck, A. E. (1982). *Introduction To Linear Regression Analysis.* New York: John Wiley & Sons Inc.

Netter, J., W. Wasserman, & M. H. Kutner. (1997). *Applied Linear Statistical Models* (Bambang Sumantri. Terjemahan). Illinois: Homewood. Buku asli diterbitkan tahun1990.

O'Kelly, M. (2006). "A Tour Around PROC ROBUSTREG". *Paper ST01*. Dublin: Quintiles Ireland Ltd.

Rousseeuw, P. J. “Least Median of Squares Regression”. (1984). *Journal of American Statistical Association*, Vol. 79, No. 388, 871-880.

Rousseeuw, P. J and Yohai, V. (1984). “Robust Regression by Means of S Estimator”, in *Robust and Nonlinier Time Series Analysis*, edited by J. Franke, W, Hardle, and R.D. Martin, Lecture Notes in Statistics 26, Springer Verlag, New York, 256-272.

SAS STAT User’s Guide, Version 9.1. (2004). North Carolina: SAS Institute.

Soemartini. (2007). “Outlier (Pencilan)”. Bandung: UNPAD.

Yaffe, A. R. (2002). “Robust Regression Analysis: Some Popular Statistical Package Options”. Academic Computing Services, Information Technology Services.

Yohai, V. J. (1987), “High Breakdown Point and High Efficiency Robust Estimates for Regression”, *Annals of Statistics*, Vol. 15, No. 20, 642-656.

http://statistikaterapan.wordpress.com.

http://www.stats.ox.ac.uk.

http://statisticsanalyst.wordpress.com.

LAMPIRAN

Lampiran 1

Prosedur manual mencari estimator MM:

A. Prosedur manual pada contoh kasus I

Prosedur estimasi parameter pada model regresi linier ganda dengan regresi *robust* estimasi MM secara manual:

1. Menghitung estimator awal dan residual $e_i^{(1)}$ dari metode estimasi S.

   Dengan bantuan software SAS 9,1 didapat:

The ROBUSTREG Procedure

Model Information

Data Set	WORK.GAJI
Dependent Variable	y
Number of Independent Variables	3
Number of Observations	24
Method	S Estimation

Number of Observations Read	24
Number of Observations Used	24

Parameter Estimates

Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	18.3309	2.3718	13.6822	22.9796	59.73	<.0001
x1	1	0.9958	0.3974	0.2170	1.7746	6.28	0.0122
x2	1	0.3181	0.0457	0.2284	0.4077	48.37	<.0001
x3	1	1.3334	0.3548	0.6379	2.0289	14.12	0.0002
Scale	0	1.8275					

Berdasarkan output diatas didapatkan nilai parameter iterasi 1:

$$\widehat{\beta}_0^{(1)} = 18,3309$$
$$\widehat{\beta}_1^{(1)} = 0,9958$$
$$\widehat{\beta}_2^{(1)} = 0,3181$$
$$\widehat{\beta}_3^{(1)} = 1,3334$$

Estimator dari metode S tersebut kemudian digunakan untuk mencari nilai residual $e_i^{(1)}$:

Dengan $e_i^{(1)} = Y_i - \widehat{Y}_i^{(1)}$

Y	X1	X2	X3	$\widehat{Y}^{(1)}$	$\lvert e_i^{(1)} \rvert$
33,2	3,5	9	6,1	32,81284	0,38716
40,3	5,3	20	6,4	38,5044	1,7956
38,7	5,1	18	7,4	39,00244	0,30244
46,8	5,8	33	6,7	43,53762	3,26238
41,4	4,2	31	7,5	42,37486	0,97486
37,5	6	13	5,9	36,30806	1,19194
39	6,8	25	6	41,05524	2,05524
40,7	5,5	30	4	38,6844	2,0156
30,1	3,1	5	5,8	30,7421	0,6421
52,9	7,2	47	8,3	51,51858	1,38142
38,2	4,5	25	5	37,4315	0,7685
31,8	4,9	11	6,4	35,24318	3,44318
43,3	8	23	7,6	43,74744	0,44744
44,1	6,5	35	7	45,2709	1,1709
42,8	6,6	39	5	43,97608	1,17608
33,6	3,7	21	4,4	34,56242	0,96242
34,2	6,2	7	5,5	34,06526	0,13474
48	7	40	7	47,3593	0,6407
38	4	35	6	41,448	3,448
35,9	4,5	23	3,5	34,7952	1,1048
40,4	5,9	33	4,9	41,23708	0,83708
36,8	5,6	27	4,3	38,2297	1,4297
45,2	4,8	34	8	44,59334	0,60666
35,1	3,9	15	5	33,65302	1,44698

2. Residual $e_i^{(1)}$ pada langkah pertama digunakan untuk menghitung pembobot awal $w_i^{(1)}$ (dengan bobot Tukey bisquare).

Berdasarkan output SAS pada langkah pertama diatas terlihat nilai *scale* (penduga dari $\hat{\sigma}_S$) adalah 1,8275.

$e_i^{(1)} / \hat{\sigma}_S$	psi ($e_i^{(1)} / \hat{\sigma}_S$)	$w_i^{(1)}$
0,211852	0,115451	0,54496
0,982544	0,49139	0,50012
-0,16549	-0,09033	0,545831
1,78516	0,713772	0,399836
-0,53344	-0,28438	0,5331
0,652224	0,343194	0,526191
-1,12462	-0,54651	0,485951
1,102927	0,538476	0,488224
-0,35135	-0,1901	0,541058
0,755907	0,392374	0,519077
0,42052	0,226414	0,538414
-1,88409	-0,72446	0,384514
-0,24484	-0,13324	0,544211
-0,64071	-0,3376	0,526919
-0,64355	-0,33898	0,526741
-0,52663	-0,28093	0,533455
0,073729	0,040324	0,546925
0,350588	0,189698	0,541084
-1,88673	-0,72469	0,384099
0,604542	0,319878	0,529125
-0,45805	-0,24587	0,536785
-0,78233	-0,40454	0,517105
0,331962	0,179829	0,541715
0,791781	0,408863	0,516384

3. Residual $e_i^{(1)}$ dengan $\hat{\sigma}_s$ pada langkah kedua digunakan dalam iterasi awal sebagai penaksir WLS untuk menghitung koefisien regresi,

   $w_i^{(1)}$ menggunakan pembobot Tukey bisquare.

   Nilai $\mathbf{w_i^{(1)}}$ dijadikan matriks diagonal nxn dengan w_i merupakan elemen diagonalnya, Kemudian dimasukkan kepersamaan dibawah ini untuk mendapatkan nilai $\tilde{\boldsymbol{\beta}}_j^{(2)}$.

$$\underset{p \times 1}{\widetilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(2)}} = \underset{p \times p}{\left(\boldsymbol{X}'\boldsymbol{W}^{(1)}\boldsymbol{X}\right)^{-1}} \; \underset{p \times 1}{\boldsymbol{X}'\boldsymbol{W}^{(1)}\boldsymbol{y}}$$

Sehingga didapat nilai parameter pada iterasi kedua:

$$\widetilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(2)} = \begin{bmatrix} 18{,}0604 \\ 1{,}0564 \\ 0{,}3200 \\ 1{,}3077 \end{bmatrix}$$

4. Menghitung bobot baru $w_i^{(2)}$ dengan skala parameter dari iterasi awal WLS,

Mencari nilai $e_i^{(2)} = Y_i - \widehat{Y}_i^{(2)}$

Y	$\widehat{Y}_i^{(2)}$	$e_i^{(2)}$	$\mathbf{e_i^{(2)}}/\ \widehat{\sigma}_S$	psi ($\mathbf{e_i^{(2)}}/\ \widehat{\sigma}_S$)	$\mathbf{w_i^{(2)}}$
33,2	32,61522	0,584783	0,31999	0,173468	0,542102
40,3	38,42932	1,87068	1,023628	0,507923	0,496198
38,7	38,88572	-0,18572	-0,10162	-0,05556	0,546681
46,8	43,51008	3,289924	1,800232	0,715658	0,397537
41,4	42,2259	-0,8259	-0,45193	-0,24271	0,537059
37,5	36,27486	1,225145	0,670394	0,351968	0,525017
39	41,09098	-2,09098	-1,14418	-0,55363	0,483868
40,7	38,70225	1,997749	1,09316	0,534812	0,489235
30,1	30,52026	-0,42026	-0,22996	-0,12523	0,544562
52,9	51,56168	1,338315	0,73232	0,38138	0,520783
38,2	37,35344	0,84656	0,463234	0,248547	0,536549
31,8	35,12659	-3,32659	-1,8203	-0,71803	0,394455
43,3	43,81104	-0,51104	-0,27964	-0,15193	0,543304
44,1	45,28194	-1,18194	-0,64675	-0,34054	0,526539
42,8	44,05221	-1,25221	-0,6852	-0,35907	0,524036
33,6	34,44359	-0,84359	-0,46161	-0,24771	0,536623
34,2	34,04296	0,157044	0,085934	0,046991	0,546827
48	47,41024	0,589756	0,322712	0,174915	0,542015
38	41,3331	-3,3331	-1,82386	-0,71843	0,393906
35,9	34,75183	1,148172	0,628275	0,331535	0,527691
40,4	41,26183	-0,86183	-0,47159	-0,25285	0,536163
36,8	38,24016	-1,44016	-0,78805	-0,40716	0,516669
45,2	44,47368	0,726318	0,397438	0,214358	0,539348
35,1	33,51941	1,580595	0,864894	0,441558	0,510534

Sehingga diperoleh nilai parameter pada iterasi ketiga:

$$\underset{p\times 1}{\tilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(3)}} = \underset{p\times p}{\left(\boldsymbol{X}'\boldsymbol{W}^{(2)}\boldsymbol{X}\right)^{-1}} \; \underset{p\times 1}{\boldsymbol{X}'\boldsymbol{W}^{(2)}\boldsymbol{y}}$$

$$\tilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(3)} = \begin{bmatrix} 18{,}0359 \\ 1{,}0627 \\ 0{,}31999 \\ 1{,}3051 \end{bmatrix}$$

5. Estimasi pada iterasi keempat

Mencari nilai $e_i^{(3)} = Y_i - \hat{Y}_i^{(3)}$

Y	$\hat{Y}_i^{(3)}$	$e_i^{(3)}$	**$e_i^{(3)}$/ $\hat{\sigma}_S$**	**psi ($e_i^{(3)}$/ $\hat{\sigma}_S$)**	**$w_i^{(3)}$**
33,2	32,59617	0,603831	0,330413	0,179007	0,541766
40,3	38,42049	1,879508	1,028458	0,509836	0,495728
38,7	38,87304	-0,17304	-0,09468	-0,05177	0,546749
46,8	43,50333	3,296671	1,803924	0,716106	0,396971
41,4	42,20711	-0,80711	-0,44165	-0,23739	0,537514
37,5	36,27184	1,228157	0,672042	0,352761	0,524908
39	41,09247	-2,09247	-1,14499	-0,55392	0,483781
40,7	38,70083	1,999171	1,093938	0,535105	0,489155
30,1	30,49958	-0,39958	-0,21865	-0,11912	0,544815
52,9	51,55917	1,340828	0,733695	0,382024	0,520685
38,2	37,34324	0,856765	0,468818	0,251423	0,536292
31,8	35,11544	-3,31544	-1,81419	-0,71732	0,395395
43,3	43,81581	-0,51581	-0,28225	-0,15333	0,543231
44,1	45,27872	-1,17872	-0,64499	-0,33968	0,52665
42,8	44,05483	-1,25483	-0,68664	-0,35976	0,52394
33,6	34,43006	-0,83006	-0,4542	-0,24389	0,536958
34,2	34,04236	0,15764	0,08626	0,047169	0,546825
48	47,41005	0,589948	0,322817	0,174971	0,542012
38	41,31695	-3,31695	-1,81502	-0,71742	0,395267
35,9	34,74563	1,154373	0,631668	0,333194	0,527482
40,4	41,26046	-0,86046	-0,47084	-0,25246	0,536198
36,8	38,23863	-1,43863	-0,78721	-0,40678	0,516734
45,2	44,45725	0,742754	0,406432	0,219063	0,53899
35,1	33,50565	1,594353	0,872423	0,444852	0,509904

Sehingga diperoleh nilai parameter pada iterasi keempat:

$$\underset{p \times 1}{\tilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(4)}} = \underset{p \times p}{\left(\boldsymbol{X}'\boldsymbol{W}^{(3)}\boldsymbol{X}\right)^{-1}} \underset{p \times 1}{\boldsymbol{X}'\boldsymbol{W}^{(3)}\boldsymbol{y}}$$

$$\tilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(4)} = \begin{bmatrix} 18{,}0332 \\ 1{,}0634 \\ 0{,}31996 \\ 1{,}3048 \end{bmatrix}$$

6. Estimasi pada iterasi kelima

Mencari nilai $e_i^{(4)} = Y_i - \hat{Y}_i^{(4)}$

Y	$\hat{Y}_i^{(4)}$	$e_i^{(4)}$	**$e_i^{(4)}/\ \hat{\sigma}_S$**	**psi ($e_i^{(4)}/\ \hat{\sigma}_S$)**	**$w_i^{(4)}$**
33,2	32,59423	0,605771	0,331475	0,17957	0,541731
40,3	38,41942	1,880578	1,029044	0,510067	0,495671
38,7	38,87164	-0,17164	-0,09392	-0,05135	0,546756
46,8	43,50208	3,297917	1,804606	0,716188	0,396867
41,4	42,20453	-0,80453	-0,44024	-0,23666	0,537575
37,5	36,27168	1,228321	0,672132	0,352804	0,524903
39	41,09244	-2,09244	-1,14497	-0,55392	0,483783
40,7	38,70015	1,999852	1,09431	0,535245	0,489116
30,1	30,49756	-0,39756	-0,21755	-0,11853	0,544838
52,9	51,55806	1,341941	0,734304	0,382309	0,520641
38,2	37,34174	0,858264	0,469638	0,251845	0,536254
31,8	35,1144	-3,3144	-1,81362	-0,71726	0,395482
43,3	43,81635	-0,51635	-0,28254	-0,15348	0,543222
44,1	45,27785	-1,17785	-0,64452	-0,33945	0,52668
42,8	44,0544	-1,2544	-0,6864	-0,35964	0,523956
33,6	34,42825	-0,82825	-0,45322	-0,24338	0,537002
34,2	34,04267	0,157334	0,086093	0,047078	0,546826
48	47,40937	0,590627	0,323188	0,175168	0,542
38	41,31446	-3,31446	-1,81366	-0,71726	0,395477
35,9	34,74458	1,155421	0,632241	0,333474	0,527447
40,4	41,25974	-0,85974	-0,47045	-0,25226	0,536216
36,8	38,23805	-1,43805	-0,7869	-0,40663	0,516757
45,2	44,45489	0,745113	0,407723	0,219737	0,538938
35,1	33,50406	1,595936	0,873289	0,44523	0,509831

Sehingga diperoleh nilai parameter pada iterasi kelima:

$$\underset{p\times 1}{\tilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\mathbf{5})}} = \underset{p\times p}{\left(\boldsymbol{X}'\boldsymbol{W}^{(\mathbf{4})}\boldsymbol{X}\right)^{-\mathbf{1}}} \ \underset{p\times 1}{\boldsymbol{X}'\boldsymbol{W}^{(\mathbf{4})}\boldsymbol{y}}$$

$$\tilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\mathbf{5})} = \begin{bmatrix} 18{,}0328 \\ 1{,}0635 \\ 0{,}31995 \\ 1{,}3048 \end{bmatrix}$$

7. Estimasi pada iterasi keenam

Mencari nilai $e_i^{(5)} = Y_i - \widehat{Y}_i^{(5)}$

Y	$\widehat{Y}_i^{(5)}$	$e_i^{(5)}$	**$e_i^{(5)}$/ $\widehat{\sigma}_S$**	**psi ($e_i^{(5)}$/ $\widehat{\sigma}_S$)**	**$w_i^{(5)}$**
33,2	32,59402	0,605983	0,331591	0,179632	0,541727
40,3	38,41929	1,880713	1,029118	0,510096	0,495664
38,7	38,87148	-0,17148	-0,09383	-0,0513	0,546757
46,8	43,50187	3,298132	1,804723	0,716202	0,396849
41,4	42,20415	-0,80415	-0,44003	-0,23655	0,537584
37,5	36,2717	1,228305	0,672123	0,352799	0,524903
39	41,09242	-2,09242	-1,14497	-0,55392	0,483784
40,7	38,69999	2,00001	1,094397	0,535277	0,489107
30,1	30,49736	-0,39736	-0,21743	-0,11847	0,544841
52,9	51,55782	1,342178	0,734434	0,38237	0,520632
38,2	37,3415	0,858501	0,469768	0,251912	0,536248
31,8	35,11431	-3,31431	-1,81357	-0,71725	0,39549
43,3	43,81644	-0,51644	-0,28259	-0,15351	0,543221
44,1	45,27769	-1,17769	-0,64442	-0,33941	0,526685
42,8	44,05424	-1,25424	-0,68632	-0,3596	0,523962
33,6	34,42798	-0,82798	-0,45307	-0,2433	0,537009
34,2	34,04277	0,15723	0,086036	0,047047	0,546827
48	47,40921	0,590789	0,323277	0,175215	0,541997
38	41,31405	-3,31405	-1,81344	-0,71723	0,395511
35,9	34,74439	1,155606	0,632343	0,333523	0,52744
40,4	41,25958	-0,85958	-0,47036	-0,25222	0,53622
36,8	38,23793	-1,43793	-0,78683	-0,4066	0,516763
45,2	44,45453	0,745471	0,407918	0,21984	0,53893

35,1	33,50386	1,59614	0,873401	0,445279	0,509822

Sehingga diperoleh nilai parameter pada iterasi keenam:

$$\underset{p\times 1}{\tilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\boldsymbol{6})}} = \underset{p\times p}{\left(\boldsymbol{X}'\boldsymbol{W}^{(\boldsymbol{5})}\boldsymbol{X}\right)^{-\boldsymbol{1}}} \underset{p\times 1}{\boldsymbol{X}'\boldsymbol{W}^{(\boldsymbol{5})}\boldsymbol{y}}$$

$$\tilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\boldsymbol{6})} = \begin{bmatrix} 18{,}0327 \\ 1{,}0636 \\ 0{,}31995 \\ 1{,}3048 \end{bmatrix}$$

8. Estimasi pada iterasi ketujuh

Mencari nilai $e_i^{(6)} = Y_i - \widehat{Y}_i^{(6)}$

Y	$\widehat{Y}_i^{(6)}$	$e_i^{(6)}$	$\mathbf{e_i^{(6)}}/\ \widehat{\sigma}_S$	**psi** $(\mathbf{e_i^{(6)}}/\ \widehat{\sigma}_S)$	$\mathbf{w_i^{(6)}}$
33,2	32,59399	0,606008	0,331605	0,179639	0,541727
40,3	38,41927	1,880731	1,029128	0,5101	0,495663
38,7	38,87146	-0,17146	-0,09382	-0,0513	0,546757
46,8	43,50183	3,298168	1,804743	0,716205	0,396846
41,4	42,20409	-0,80409	-0,44	-0,23654	0,537585
37,5	36,2717	1,228297	0,672119	0,352797	0,524903
39	41,09242	-2,09242	-1,14496	-0,55392	0,483784
40,7	38,69996	2,000041	1,094414	0,535284	0,489105
30,1	30,49733	-0,39733	-0,21742	-0,11846	0,544841
52,9	51,55778	1,342221	0,734458	0,382381	0,52063
38,2	37,34146	0,85854	0,469789	0,251923	0,536247
31,8	35,1143	-3,3143	-1,81357	-0,71725	0,395491
43,3	43,81646	-0,51646	-0,2826	-0,15352	0,543221
44,1	45,27766	-1,17766	-0,64441	-0,3394	0,526686
42,8	44,05421	-1,25421	-0,6863	-0,3596	0,523963
33,6	34,42794	-0,82794	-0,45305	-0,24329	0,53701
34,2	34,04279	0,157206	0,086023	0,047039	0,546827
48	47,40918	0,59082	0,323294	0,175224	0,541997
38	41,31398	-3,31398	-1,8134	-0,71723	0,395517
35,9	34,74436	1,155639	0,632361	0,333532	0,527439
40,4	41,25955	-0,85955	-0,47034	-0,25221	0,536221
36,8	38,2379	-1,4379	-0,78681	-0,4066	0,516764

45,2	44,45447	0,745527	0,407949	0,219856	0,538929
35,1	33,50383	1,596169	0,873416	0,445286	0,50982

Sehingga diperoleh nilai parameter pada iterasi ketujuh:

$$\underset{p\times 1}{\tilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\mathbf{7})}} = \underset{p\times p}{\left(\boldsymbol{X}'\boldsymbol{W}^{(\mathbf{6})}\boldsymbol{X}\right)^{-\mathbf{1}}} \ \underset{p\times 1}{\boldsymbol{X}'\boldsymbol{W}^{(\mathbf{6})}\boldsymbol{y}}$$

$$\tilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\mathbf{7})} = \begin{bmatrix} 18,0327 \\ 1,0636 \\ 0,31995 \\ 1,3048 \end{bmatrix}$$

9. Selisih nilai pada iterasi keenam dan ketujuh konvergen (selisih $\tilde{\beta}_j = 0$) sehingga nilai parameter regresi dari estimasi MM adalah:

$$\tilde{\boldsymbol{\beta}}_{\boldsymbol{j}} = \begin{bmatrix} 18,0327 \\ 1,0636 \\ 0,31995 \\ 1,3048 \end{bmatrix}$$

Diperoleh model regresi dari estimasi MM:

$\hat{Y} = 18,0327 + 1,0636\ X_1 + 0,31995\ X_2 + 1,3048\ X_3$

B. Prosedur manual pada contoh kasus II

Prosedur estimasi parameter pada model regresi linier ganda dengan regresi *robust* estimasi MM secara manual:

1. Menghitung estimator awal dan residual $e_i^{(1)}$ dari metode estimasi S.

   Dengan bantuan software SAS 9,1 didapat:

The ROBUSTREG Procedure

Model Information

Data Set	WORK.CON2S
Dependent Variable	y
Number of Independent Variables	5
Number of Observations	20
Method	S Estimation

Number of Observations Read	20
Number of Observations Used	20

Parameter Estimates

Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	0.3825	0.0822	0.2215	0.5436	21.67	<.0001
x1	1	0.0002	0.0001	0.0001	0.0003	7.83	0.0051
x2	1	0.0000	0.0000	-0.0001	0.0001	0.05	0.8228
x3	1	-0.0059	0.0006	-0.0071	-0.0047	93.69	<.0001
x4	1	-0.0039	0.0009	-0.0058	-0.0021	18.33	<.0001
x5	1	0.0062	0.0011	0.0040	0.0084	31.21	<.0001
Scale	0	0.0173					

Berdasarkan output diatas didapatkan nilai parameter iterasi 1:

$\widehat{\beta}_0^{(1)} = 0{,}3825$ $\qquad$ $\widehat{\beta}_3^{(1)} = -0{,}0059$

$\widehat{\beta}_1^{(1)} = 0{,}0002$ $\qquad$ $\widehat{\beta}_4^{(1)} = -0{,}0039$

$\widehat{\beta}_2^{(1)} = 0$ $\qquad$ $\widehat{\beta}_5^{(1)} = 0{,}0062$

Estimator dari metode S tersebut kemudian digunakan untuk mencari nilai residual $e_i^{(1)}$:

Dengan $e_i^{(1)} = Y_i - \widehat{Y}_i^{(1)}$

Y	X1	X2	X3	X4	X5	$\widehat{Y}^{(1)}$	$\lvert e_i^{(1)} \rvert$
0,534	573	1059	46,5	53,8	84,1	0,53435	0,00035
0,535	651	1356	52,7	54,5	88,7	0,53916	0,00416
0,57	606	1273	49,4	52,1	92	0,57945	0,00945
0,528	630	1151	48,9	50,3	87,9	0,5688	0,0408
0,548	547	1135	53,1	51,9	91,5	0,5435	0,0045
0,555	557	1236	54,9	55,2	91,4	0,52139	0,03361
0,481	489	1231	56,2	45,5	82,4	0,48215	0,00115
0,516	685	1564	56,6	44,3	91,3	0,57885	0,06285
0,475	536	1182	59,2	46,4	85,4	0,48894	0,01394
0,486	685	1564	63,1	56,4	91,4	0,49393	0,00793
0,554	664	1588	50,6	48,1	86,7	0,56671	0,01271
0,519	703	1335	51,9	48,4	81,2	0,53157	0,01257
0,492	653	1395	62,5	51,9	89,2	0,49498	0,00298
0,517	586	1114	50,5	56,5	88,9	0,53258	0,01558
0,502	534	1143	52,1	57	88,9	0,51079	0,00879
0,508	523	1320	50,5	61,2	91,9	0,52025	0,01225
0,52	580	1249	54,6	60,8	95,4	0,53072	0,01072
0,506	448	1028	52,2	53,4	91,8	0,52502	0,01902
0,595	476	1057	42,9	53,2	92,9	0,59309	0,00191
0,568	528	1057	42,4	56,6	90	0,5752	0,0072

2. Residual $e_i^{(1)}$ pada langkah pertama digunakan untuk menghitung pembobot awal $w_i^{(1)}$ (dengan bobot Tukey bisquare).

   Berdasarkan output SAS pada langkah pertama diatas terlihat nilai *scale* (penduga dari $\widehat{\sigma}_S$) adalah 0,0173.

$\mathbf{e_i^{(1)}} / \widehat{\sigma}_S$	**psi** $(\mathbf{e_i^{(1)}} / \widehat{\sigma}_S)$	$\mathbf{w_i^{(1)}}$
-0,02023	-1,16939	57,80131
-0,24046	-13,8264	57,49932
-0,54624	-30,7221	56,24257
-2,35838	-75,9875	32,2202
0,260116	14,94303	57,44765
1,942775	76,99804	39,63303

$e_i^{(1)}/\ \hat{\sigma}_S$	psi ($e_i^{(1)}/\ \hat{\sigma}_S$)	$w_i^{(1)}$
-0,06647	-3,84088	57,7802
-3,63295	-33,3796	9,188029
-0,80578	-43,8621	54,43427
-0,45838	-25,9912	56,70209
-0,73468	-40,4042	54,99551
-0,72659	-40,0033	55,05628
-0,17225	-9,93	57,64729
-0,90058	-48,2806	53,61063
-0,50809	-28,6827	56,45174
-0,70809	-39,0816	55,19278
-0,61965	-34,5759	55,79878
-1,09942	-56,7438	51,61237
0,110405	6,374684	57,73929
-0,41618	-23,6787	56,89477

3. Residual $e_i^{(1)}$ dengan $\hat{\sigma}_s$ pada langkah kedua digunakan dalam iterasi awal sebagai penaksir WLS untuk menghitung koefisien regresi,

   $w_i^{(1)}$ menggunakan pembobot Tukey bisquare.

   Nilai $\mathbf{w_i^{(1)}}$ dijadikan matriks diagonal nxn dengan w_i merupakan elemen diagonalnya, Kemudian dimasukkan kepersamaan dibawah ini untuk mendapatkan nilai $\tilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\mathbf{2})}$.

$$\underset{p\times 1}{\tilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\mathbf{2})}} = \underset{p\times p}{\left(\boldsymbol{X}'\boldsymbol{W}^{(\mathbf{1})}\boldsymbol{X}\right)^{-\mathbf{1}}} \ \underset{p\times 1}{\boldsymbol{X}'\boldsymbol{W}^{(\mathbf{1})}\boldsymbol{y}}$$

   Sehingga didapat nilai parameter pada iterasi kedua:

$$\tilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\mathbf{2})} = \begin{bmatrix} 0{,}3875 \\ 0{,}0001 \\ 0{,}0000 \\ -0{,}0056 \\ -0{,}0033 \\ 0{,}0058 \end{bmatrix}$$

4. Menghitung bobot baru $w_i^{(2)}$ dengan skala parameter dari iterasi awal WLS.

Mencari nilai $e_i^{(2)} = Y_i - \widehat{Y}_i^{(2)}$

Y	$\widehat{Y}_i^{(2)}$	$e_i^{(2)}$	$\mathbf{e_i^{(2)}} / \widehat{\sigma}_S$	**psi** $(\mathbf{e_i^{(2)}} / \widehat{\sigma}_S)$	$\mathbf{w_i^{(2)}}$
0,534	0,52669173	0,007308	0,42244307	24,02321672	56,86734725
0,535	0,52938082	0,005619	0,32480781	18,59496519	57,24913279
0,57	0,56784875	0,002151	0,12434993	7,177733074	57,72205352
0,528	0,55577868	-0,02778	-1,6057038	-72,2906874	45,02118523
0,548	0,53535741	0,012643	0,73078562	40,21136937	55,02485003
0,555	0,51576975	0,03923	2,26764477	76,85490127	33,89194913
0,481	0,47825579	0,002744	0,15862497	9,148063058	57,67101613
0,516	0,56243899	-0,04644	-2,6843348	-70,009601	26,0808003
0,475	0,48256426	-0,00756	-0,4372407	-24,8356712	56,80090803
0,486	0,48651041	-0,00051	-0,0295037	-1,70528129	57,79888353
0,554	0,55388202	0,000118	0,0068198	0,394206608	57,80322324
0,519	0,51819775	0,000802	0,046373	2,679995172	57,79214228
0,492	0,48637669	0,005623	0,32504679	18,60838241	57,24831872
0,517	0,52535172	-0,00835	-0,4827584	-27,3156636	56,5824771
0,502	0,50713684	-0,00514	-0,2969269	-17,0257979	57,34003123
0,508	0,51889612	-0,0109	-0,6298334	-35,1024871	55,73297573
0,52	0,52560208	-0,0056	-0,3238197	-18,5394861	57,25249224
0,506	0,52191116	-0,01591	-0,9197199	-49,1443417	53,43402842
0,595	0,58545915	0,009541	0,55149394	31,00092308	56,21262729
0,568	0,56793955	6,04E-05	0,00349404	0,201967411	57,80340391

Sehingga diperoleh nilai parameter pada iterasi ketiga:

$$\underset{p \times 1}{\tilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\mathbf{3})}} = \underset{p \times p}{\left(\boldsymbol{X}'\boldsymbol{W}^{(\mathbf{2})}\boldsymbol{X}\right)^{-\mathbf{1}}} \; \underset{p \times 1}{\boldsymbol{X}'\boldsymbol{W}^{(\mathbf{2})}\boldsymbol{y}}$$

$$\tilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\mathbf{3})} = \begin{bmatrix} 0,4244 \\ 0,0001 \\ 0,0000 \\ -0,0055 \\ -0,0026 \\ 0,0050 \end{bmatrix}$$

5. Estimasi pada iterasi keempat

Mencari nilai $e_i^{(3)} = Y_i - \widehat{Y}_i^{(3)}$

Y	$\widehat{Y}_i^{(3)}$	$e_i^{(3)}$	$\mathbf{e_i^{(3)}}/\ \widehat{\sigma}_S$	**psi** ($\mathbf{e_i^{(3)}}/\ \widehat{\sigma}_S$)	$\mathbf{w_i^{(3)}}$
0,534	0,529052	0,004948	0,285983	16,40784	57,3735
0,535	0,527986	0,007014	0,405431	23,08565	56,94095
0,57	0,562634	0,007366	0,425798	24,20766	56,85248
0,528	0,551711	-0,02371	-1,37059	-66,2444	48,33262
0,548	0,531673	0,016327	0,94376	50,21503	53,20741
0,555	0,514691	0,040309	2,329989	76,29731	32,74578
0,481	0,478862	0,002138	0,123574	7,133059	57,72307
0,516	0,551559	-0,03556	-2,05542	-77,4752	37,69311
0,475	0,480547	-0,00555	-0,32062	-18,3596	57,26331
0,486	0,484837	0,001163	0,06725	3,885666	57,77965
0,554	0,549468	0,004532	0,261984	15,04904	57,44253
0,519	0,517114	0,001886	0,109018	6,294803	57,74089
0,492	0,483485	0,008515	0,492171	27,8247	56,53467
0,517	0,52599	-0,00899	-0,51966	-29,3036	56,38987
0,502	0,509447	-0,00745	-0,43049	-24,4654	56,8315
0,508	0,522348	-0,01435	-0,82938	-44,9834	54,23719
0,52	0,524911	-0,00491	-0,28385	-16,2874	57,37987
0,506	0,520932	-0,01493	-0,86312	-46,562	53,94628
0,595	0,582198	0,012802	0,740012	40,66744	54,95514
0,568	0,568292	-0,00029	-0,01685	-0,97397	57,80197

Sehingga diperoleh nilai parameter pada iterasi keempat:

$$\underset{p\times 1}{\widetilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\mathbf{4})}} = \underset{p\times p}{\left(\boldsymbol{X}'\boldsymbol{W}^{(\mathbf{3})}\boldsymbol{X}\right)^{-\mathbf{1}}} \ \underset{p\times 1}{\boldsymbol{X}'\boldsymbol{W}^{(\mathbf{3})}\boldsymbol{y}}$$

$$\widetilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\mathbf{4})} = \begin{bmatrix} 0{,}4409 \\ 0{,}0001 \\ 0{,}0000 \\ -0{,}0055 \\ -0{,}0022 \\ 0{,}0046 \end{bmatrix}$$

6. Estimasi pada iterasi kelima

Mencari nilai $e_i^{(4)} = Y_i - \widehat{Y}_i^{(4)}$

Y	$\widehat{Y}_i^{(4)}$	$e_i^{(4)}$	$\mathbf{e_i^{(4)}}/\ \widehat{\sigma}_S$	**psi** ($\mathbf{e_i^{(4)}}/\ \widehat{\sigma}_S$)	$\mathbf{w_i^{(4)}}$
0,531547	0,530426	0,003574	0,206601	11,89583	57,57887
0,527399	0,527538	0,007462	0,431348	24,51246	56,82764
0,558743	0,560408	0,009592	0,554439	31,15709	56,19571

Y	$\widehat{Y}_i^{(4)}$	$e_i^{(4)}$	**e$_i^{(4)}$**/ $\widehat{\sigma}_S$	**psi (e$_i^{(4)}$**/ $\widehat{\sigma}_S$ **)**	**w$_i^{(4)}$**
0,548659	0,550152	-0,02215	-1,28044	-63,3697	49,49057
0,53011	0,530182	0,017818	1,029967	53,91988	52,35107
0,515499	0,514344	0,040656	2,350071	76,08181	32,37426
0,479845	0,479099	0,001901	0,109889	6,345	57,73988
0,54197	0,546589	-0,03059	-1,76813	-75,1631	42,50991
0,48014	0,479735	-0,00474	-0,2737	-15,7131	57,40958
0,485039	0,484182	0,001818	0,105084	6,068095	57,74532
0,544462	0,547392	0,006608	0,381948	21,7854	57,03765
0,515576	0,516873	0,002127	0,122969	7,098227	57,72385
0,482653	0,48234	0,00966	0,558366	31,36509	56,17302
0,527968	0,526585	-0,00959	-0,55405	-31,1366	56,19794
0,513075	0,510695	-0,00869	-0,50257	-28,3857	56,48079
0,526704	0,523956	-0,01596	-0,92229	-49,2596	53,41005
0,526728	0,524809	-0,00481	-0,278	-15,9565	57,39712
0,522209	0,520604	-0,0146	-0,84417	-45,6791	54,11097
0,580175	0,580814	0,014186	0,819973	44,538	54,31639
0,569276	0,568674	-0,00067	-0,03895	-2,25101	57,79548

Sehingga diperoleh nilai parameter pada iterasi kelima:

$$\underset{p\times 1}{\tilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\mathbf{5})}} = \underset{p\times p}{\left(\boldsymbol{X}'\boldsymbol{W}^{(\mathbf{4})}\boldsymbol{X}\right)^{-\mathbf{1}}} \quad \underset{p\times 1}{\boldsymbol{X}'\boldsymbol{W}^{(\mathbf{4})}\boldsymbol{y}}$$

$$\tilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\mathbf{5})} = \begin{bmatrix} 0{,}4467 \\ 0{,}0001 \\ 0{,}0000 \\ -0{,}0055 \\ -0{,}0021 \\ 0{,}0045 \end{bmatrix}$$

7. Estimasi pada iterasi keenam

Mencari nilai $e_i^{(5)} = Y_i - \widehat{Y}_i^{(5)}$

Y	$\widehat{Y}_i^{(5)}$	$e_i^{(5)}$	**e$_i^{(5)}$**/ $\widehat{\sigma}_S$	**psi (e$_i^{(5)}$**/ $\widehat{\sigma}_S$ **)**	**w$_i^{(5)}$**
0,531547	0,530905	0,003095	0,17892	10,31203	57,63498
0,527399	0,527391	0,007609	0,43982	24,97696	56,7891
0,558743	0,559614	0,010386	0,600322	33,57055	55,92089
0,548659	0,549581	-0,02158	-1,24745	-62,2452	49,89779
0,53011	0,52964	0,01836	1,061254	55,21027	52,02363
0,515499	0,514227	0,040773	2,356847	76,00517	32,24867

Y	$\widehat{Y}_i^{(5)}$	$e_i^{(5)}$	$e_i^{(5)} / \widehat{\sigma}_S$	psi ($e_i^{(5)} / \widehat{\sigma}_S$)	$w_i^{(5)}$
0,479845	0,479192	0,001808	0,104526	6,035947	57,74594
0,54197	0,544833	-0,02883	-1,66665	-73,4975	44,09884
0,48014	0,479443	-0,00444	-0,2568	-14,7551	57,45664
0,485039	0,483985	0,002015	0,116481	6,724665	57,73203
0,544462	0,546683	0,007317	0,42294	24,05056	56,86515
0,515576	0,516792	0,002208	0,127632	7,366638	57,7177
0,482653	0,481947	0,010053	0,581077	32,5628	56,03874
0,527968	0,526791	-0,00979	-0,56596	-31,7664	56,12872
0,513075	0,511142	-0,00914	-0,52845	-29,7739	56,34196
0,526704	0,524556	-0,01656	-0,95697	-50,7964	53,08064
0,526728	0,524783	-0,00478	-0,27648	-15,8705	57,40155
0,522209	0,520476	-0,01448	-0,83679	-45,3323	54,17427
0,580175	0,580307	0,014693	0,849293	45,91856	54,0668
0,569276	0,5688	-0,0008	-0,04622	-2,6714	57,79221

Sehingga diperoleh nilai parameter pada iterasi keenam:

$$\underset{p\times 1}{\tilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\boldsymbol{6})}} = \underset{p\times p}{\left(\boldsymbol{X}'\boldsymbol{W}^{(\boldsymbol{5})}\boldsymbol{X}\right)^{-\boldsymbol{1}}} \; \underset{p\times 1}{\boldsymbol{X}'\boldsymbol{W}^{(\boldsymbol{5})}\boldsymbol{y}}$$

$$\tilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\boldsymbol{6})} = \begin{bmatrix} 0,4486 \\ 0,0001 \\ 0,0000 \\ -0,0055 \\ -0,0021 \\ 0,0045 \end{bmatrix}$$

8. Estimasi pada iterasi ketujuh

Mencari nilai $e_i^{(6)} = Y_i - \widehat{Y}_i^{(6)}$

Y	$\widehat{Y}_i^{(6)}$	$e_i^{(6)}$	$e_i^{(6)} / \widehat{\sigma}_S$	psi ($e_i^{(6)} / \widehat{\sigma}_S$)	$w_i^{(6)}$
0,531547	0,531051	0,002949	0,170485	9,828527	57,65048
0,527399	0,527346	0,007654	0,442437	25,12024	56,77705
0,558743	0,559359	0,010641	0,615105	34,33997	55,82785
0,548659	0,54939	-0,02139	-1,23642	-61,8606	50,03199
0,53011	0,529461	0,018539	1,071597	55,63033	51,91345
0,515499	0,514189	0,040811	2,359033	75,98002	32,20812
0,479845	0,479224	0,001776	0,102635	5,926969	57,748
0,54197	0,544277	-0,02828	-1,63453	-72,8805	44,5881
0,48014	0,479346	-0,00435	-0,25123	-14,4385	57,47151
0,485039	0,483931	0,002069	0,119603	6,904438	57,72815

Y	$\widehat{Y}_i^{(6)}$	$e_i^{(6)}$	$\mathbf{e_i^{(6)}}/\ \widehat{\sigma}_S$	**psi** ($\mathbf{e_i^{(6)}}/\ \widehat{\sigma}_S$)	$\mathbf{w_i^{(6)}}$
0,544462	0,546468	0,007532	0,435392	24,73431	56,80933
0,515576	0,516765	0,002235	0,129208	7,457333	57,71557
0,482653	0,481823	0,010177	0,588246	32,939	55,99527
0,527968	0,526851	-0,00985	-0,5694	-31,948	56,10844
0,513075	0,511283	-0,00928	-0,53661	-30,2092	56,2968
0,526704	0,524755	-0,01676	-0,96851	-51,3006	52,9685
0,526728	0,524774	-0,00477	-0,27597	-15,8414	57,40304
0,522209	0,52043	-0,01443	-0,83411	-45,2062	54,1971
0,580175	0,58014	0,01486	0,858971	46,36951	53,98262
0,569276	0,568834	-0,00083	-0,04822	-2,78647	57,79122

Sehingga diperoleh nilai parameter pada iterasi ketujuh:

$$\underset{p\times 1}{\widetilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(7)}} = \underset{p\times p}{\left(\boldsymbol{X}'\boldsymbol{W}^{(6)}\boldsymbol{X}\right)^{-\mathbf{1}}}\ \underset{p\times 1}{\boldsymbol{X}'\boldsymbol{W}^{(6)}\boldsymbol{y}}$$

$$\widetilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(7)} = \begin{bmatrix} 0,4491 \\ 0,0001 \\ 0,0000 \\ -0,0055 \\ -0,0021 \\ 0,0045 \end{bmatrix}$$

9. Estimasi pada iterasi kedelapan

Mencari nilai $e_i^{(7)} = Y_i - \widehat{Y}_i^{(7)}$

Y	$\widehat{Y}_i^{(7)}$	$e_i^{(7)}$	$\mathbf{e_i^{(7)}}/\ \widehat{\sigma}_S$	**psi** ($\mathbf{e_i^{(7)}}/\ \widehat{\sigma}_S$)	$\mathbf{w_i^{(7)}}$
0,531547	0,531094	0,002906	0,167972	9,684422	57,65496
0,527399	0,527332	0,007668	0,443228	25,16357	56,77339
0,558743	0,55928	0,01072	0,619678	34,57716	55,79862
0,548659	0,54933	-0,02133	-1,23293	-61,7379	50,07429
0,53011	0,529405	0,018595	1,074856	55,76196	51,87855
0,515499	0,514177	0,040823	2,359718	75,9721	32,19542
0,479845	0,479235	0,001765	0,102019	5,891444	57,74866
0,54197	0,544107	-0,02811	-1,6247	-72,6833	44,73636
0,48014	0,479316	-0,00432	-0,24947	-14,3385	57,47614
0,485039	0,483916	0,002084	0,120474	6,954621	57,72705
0,544462	0,546404	0,007596	0,439087	24,93681	56,79246
0,515576	0,516756	0,002244	0,129712	7,486306	57,71488
0,482653	0,481785	0,010215	0,590449	33,05441	55,98181

Y	$\widehat{Y}_i^{(7)}$	$e_i^{(7)}$	$\mathbf{e_i^{(7)}}/\ \widehat{\sigma}_S$	**psi** $(\mathbf{e_i^{(7)}}/\ \widehat{\sigma}_S)$	$\mathbf{w_i^{(7)}}$
0,527968	0,526868	-0,00987	-0,57038	-31,9996	56,10265
0,513075	0,511326	-0,00933	-0,53909	-30,3414	56,28293
0,526704	0,524818	-0,01682	-0,97215	-51,4589	52,93287
0,526728	0,524771	-0,00477	-0,27579	-15,8314	57,40355
0,522209	0,520415	-0,01441	-0,83321	-45,1641	54,20469
0,580175	0,580087	0,014913	0,862005	46,51038	53,95604
0,569276	0,568844	-0,00084	-0,04877	-2,81822	57,79094

Sehingga diperoleh nilai parameter pada iterasi kedelapan:

$$\underset{p\times 1}{\tilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\mathbf{8})}} = \underset{p\times p}{\left(\boldsymbol{X}'\boldsymbol{W}^{(7)}\boldsymbol{X}\right)^{-\mathbf{1}}}\ \underset{p\times 1}{\boldsymbol{X}'\boldsymbol{W}^{(7)}\boldsymbol{y}}$$

$$\tilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\mathbf{8})} = \begin{bmatrix} 0,4493 \\ 0,0001 \\ 0,0000 \\ -0,0055 \\ -0,0021 \\ 0,0044 \end{bmatrix}$$

10. Estimasi pada iterasi kesembilan

Mencari nilai $e_i^{(8)} = Y_i - \widehat{Y}_i^{(8)}$

Y	$\widehat{Y}_i^{(8)}$	$e_i^{(8)}$	$\mathbf{e_i^{(8)}}/\ \widehat{\sigma}_S$	**psi** $(\mathbf{e_i^{(8)}}/\ \widehat{\sigma}_S)$	$\mathbf{w_i^{(8)}}$
0,531547	0,531107	0,002893	0,167221	9,641338	57,65628
0,527399	0,527328	0,007672	0,443467	25,17664	56,77228
0,558743	0,559255	0,010745	0,621072	34,64939	55,78967
0,548659	0,549311	-0,02131	-1,23185	-61,6999	50,08733
0,53011	0,529388	0,018612	1,075859	55,80242	51,86779
0,515499	0,514173	0,040827	2,359929	75,96966	32,19151
0,479845	0,479238	0,001762	0,101826	5,880343	57,74887
0,54197	0,544056	-0,02806	-1,62172	-72,6227	44,78117
0,48014	0,479306	-0,00431	-0,24893	-14,3076	57,47756
0,485039	0,483911	0,002089	0,120727	6,969165	57,72673
0,544462	0,546385	0,007615	0,440187	24,9971	56,78741
0,515576	0,516753	0,002247	0,129868	7,495276	57,71467
0,482653	0,481774	0,010226	0,59112	33,08955	55,9777
0,527968	0,526872	-0,00987	-0,57066	-32,0146	56,10097
0,513075	0,511339	-0,00934	-0,53984	-30,3813	56,27873
0,526704	0,524838	-0,01684	-0,97327	-51,5075	52,92187

Y	$\widehat{Y}_i^{(8)}$	$e_i^{(8)}$	$\mathbf{e_i^{(8)}}/\ \widehat{\sigma}_S$	psi ($\mathbf{e_i^{(8)}}/\ \widehat{\sigma}_S$)	$\mathbf{w_i^{(8)}}$
0,526728	0,52477	-0,00477	-0,27573	-15,8282	57,40371
0,522209	0,52041	-0,01441	-0,83293	-45,1508	54,2071
0,580175	0,580071	0,014929	0,862935	46,55353	53,94788
0,569276	0,568846	-0,00085	-0,04892	-2,82734	57,79086

Sehingga diperoleh nilai parameter pada iterasi kesembilan:

$$\underset{p\times 1}{\tilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\mathbf{9})}} = \underset{p\times p}{\left(\boldsymbol{X}'\boldsymbol{W}^{(\mathbf{8})}\boldsymbol{X}\right)^{-\mathbf{1}}}\ \underset{p\times 1}{\boldsymbol{X}'\boldsymbol{W}^{(\mathbf{8})}\boldsymbol{y}}$$

$$\tilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\mathbf{9})} = \begin{bmatrix} 0,4494 \\ 0,0001 \\ 0,0000 \\ -0,0055 \\ -0,0021 \\ 0,0044 \end{bmatrix}$$

11. Estimasi pada iterasi kesepuluh

Mencari nilai $e_i^{(9)} = Y_i - \widehat{Y}_i^{(9)}$

Y	$\widehat{Y}_i^{(9)}$	$e_i^{(9)}$	$\mathbf{e_i^{(9)}}/\ \widehat{\sigma}_S$	psi ($\mathbf{e_i^{(9)}}/\ \widehat{\sigma}_S$)	$\mathbf{w_i^{(9)}}$
0,531547	0,531111	0,002889	0,166995	9,628389	57,65668
0,527399	0,527327	0,007673	0,443539	25,18059	56,77195
0,558743	0,559248	0,010752	0,621495	34,67129	55,78695
0,548659	0,549305	-0,02131	-1,23152	-61,6883	50,0913
0,53011	0,529382	0,018618	1,076165	55,81475	51,8645
0,515499	0,514172	0,040828	2,359993	75,96891	32,19031
0,479845	0,479239	0,001761	0,101767	5,87694	57,74893
0,54197	0,54404	-0,02804	-1,62082	-72,6043	44,79472
0,48014	0,479304	-0,0043	-0,24876	-14,2982	57,478
0,485039	0,48391	0,00209	0,120802	6,973482	57,72663
0,544462	0,546379	0,007621	0,440518	25,01518	56,78589
0,515576	0,516752	0,002248	0,129915	7,498014	57,71461
0,482653	0,48177	0,01023	0,591324	33,10021	55,97645
0,527968	0,526874	-0,00987	-0,57075	-32,0191	56,10047
0,513075	0,511343	-0,00934	-0,54006	-30,3933	56,27746
0,526704	0,524844	-0,01684	-0,97362	-51,5224	52,91852
0,526728	0,52477	-0,00477	-0,27572	-15,8271	57,40377
0,522209	0,520408	-0,01441	-0,83284	-45,1466	54,20785
0,580175	0,580066	0,014934	0,863218	46,56665	53,94539
0,569276	0,568847	-0,00085	-0,04897	-2,83003	57,79084

Sehingga diperoleh nilai parameter pada iterasi kesepuluh:

$$\underset{p\times 1}{\widetilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\mathbf{10})}} = \underset{p\times p}{\left(\boldsymbol{X}'\boldsymbol{W}^{(\mathbf{9})}\boldsymbol{X}\right)^{-\mathbf{1}}} \; \underset{p\times 1}{\boldsymbol{X}'\boldsymbol{W}^{(\mathbf{9})}\boldsymbol{y}}$$

$$\widetilde{\boldsymbol{\beta}}_{\boldsymbol{j}}^{(\mathbf{10})} = \begin{bmatrix} 0{,}4494 \\ 0{,}0001 \\ 0{,}0000 \\ -0{,}0055 \\ -0{,}0021 \\ 0{,}0044 \end{bmatrix}$$

12. Selisih nilai pada iterasi kesembilan dan kesepuluh konvergen (selisih $\tilde{\beta}_j =$ 0) sehingga nilai parameter regresi dari estimasi MM adalah:

$$\widetilde{\boldsymbol{\beta}}_{\boldsymbol{j}} = \begin{bmatrix} 0{,}4494 \\ 0{,}0001 \\ 0{,}0000 \\ -0{,}0055 \\ -0{,}0021 \\ 0{,}0044 \end{bmatrix}$$

Diperoleh model regresi dari estimasi MM:

$\hat{Y} = 0{,}4494 + 0{,}0001\, X_1 - 0{,}0055\, X_3 - 0{,}0021\, X_4 + 0{,}0044\, X_5$

Lampiran 2

Sintaks SAS 9.1:

A. Sintaks SAS 9.1 untuk contoh kasus I

1. Estimasi dengan regresi robust MM

```
data gaji;
input x1 x2 x3 y;
cards;
 3.5        9      6.1   33.2
 5.3        20     6.4   40.3
 5.1        18     7.4   38.7
 5.8        33     6.7   46.8
 4.2        31     7.5   41.4
 6         13      5.9   37.5
 6.8        25     6     39
 5.5        30     4     40.7
 3.1        5      5.8   30.1
 7.2        47     8.3   52.9
 4.5        25     5     38.2
 4.9        11     6.4   31.8
 8          23     7.6   43.3
 6.5        35     7     44.1
 6.6        39     5     42.8
 3.7        21     4.4   33.6
 6.2        7      5.5   34.2
 7         40      7     48
 4         35      6     38
 4.5        23     3.5   35.9
 5.9        33     4.9   40.4
 5.6        27     4.3   36.8
 4.8        34     8     45.2
 3.9        15     5     35.1
;
proc robustreg method=mm data=gaji;
model y=x1 x2 x3;
run;
```

2. Estimasi dengan metode kuadrat terkecil data outlier dihapus

```
data gaji;
input x1 x2 x3 y;
cards;
 3.5        9      6.1   33.2
 5.3        20     6.4   40.3
 5.1        18     7.4   38.7
 5.8        33     6.7   46.8
 4.2        31     7.5   41.4
 6         13      5.9   37.5
 6.8        25     6     39
 5.5        30     4     40.7
 3.1        5      5.8   30.1
 7.2        47     8.3   52.9
 4.5        25     5     38.2
 4.9        11     6.4   31.8
 8         23      7.6   43.3
 6.5        35     7     44.1
 6.6        39     5     42.8
 3.7        21     4.4   33.6
 6.2        7      5.5   34.2
 7          40     7     48
 4.5        23     3.5   35.9
 5.9        33     4.9   40.4
 5.6        27     4.3   36.8
 4.8        34     8     45.2
 3.9        15     5     35.1
;
proc reg data=gaji;
model y=x1 x2 x3;
run;
```

B. Sintaks SAS 9.1 untuk contoh kasus II

1. Estimasi dengan regresi robust MM

```
data con2;
input x1 x2 x3 x4 x5 y;
cards;
   573    1059  46.5  53.8  84.1  0.534
   651    1356  52.7  54.5  88.7  0.535
   606    1273  49.4  52.1  92.0  0.570
   630    1151  48.9  50.3  87.9  0.528
   547    1135  53.1  51.9  91.5  0.548
   557    1236  54.9  55.2  91.4  0.555
   489    1231  56.2  45.5  82.4  0.481
   685    1564  56.6  44.3  91.3  0.516
   536    1182  59.2  46.4  85.4  0.475
   685    1564  63.1  56.4  91.4  0.486
   664    1588  50.6  48.1  86.7  0.554
   703    1335  51.9  48.4  81.2  0.519
   653    1395  62.5  51.9  89.2  0.492
   586    1114  50.5  56.5  88.9  0.517
   534    1143  52.1  57.0  88.9  0.502
   523    1320  50.5  61.2  91.9  0.508
   580    1249  54.6  60.8  95.4  0.520
   448    1028  52.2  53.4  91.8  0.506
   476    1057  42.9  53.2  92.9  0.595
   528    1057  42.4  56.6  90.0  0.568
;
proc robustreg data = con2 method=MM ;
model y=x1 x2 x3 x4 x5;
run;
```

2. Estimasi dengan metode kuadrat terkecil data outlier dihapus

```
data con2;
input x1 x2 x3 x4 x5 y;
cards;
   573     1059   46.5   53.8   84.1   0.534
   651     1356   52.7   54.5   88.7   0.535
   606     1273   49.4   52.1   92.0   0.570
   630     1151   48.9   50.3   87.9   0.528
   547     1135   53.1   51.9   91.5   0.548
   557     1236   54.9   55.2   91.4   0.555
   489     1231   56.2   45.5   82.4   0.481
   685     1564   56.6   44.3   91.3   0.516
   536     1182   59.2   46.4   85.4   0.475
   685     1564   63.1   56.4   91.4   0.486
   664     1588   50.6   48.1   86.7   0.554
   703     1335   51.9   48.4   81.2   0.519
   653     1395   62.5   51.9   89.2   0.492
   586     1114   50.5   56.5   88.9   0.517
   534     1143   52.1   57.0   88.9   0.502
   523     1320   50.5   61.2   91.9   0.508
   580     1249   54.6   60.8   95.4   0.520
   448     1028   52.2   53.4   91.8   0.506
   476     1057   42.9   53.2   92.9   0.595
   528     1057   42.4   56.6   90.0   0.568
;
proc reg data=con2;
model y=x1 x2 x3 x4 x5;
run;
```