Metode Regresi *Robust* Dengan Estimasi *Method of Moment* (Estimasi-MM) Pada Regresi Linier Berganda (Studi Kasus : Data Indeks Harga Konsumen (IHK) Provinsi Kalimantan Timur)

Method of Robust Regression With Method of Moment Estimation (MM-Estimation) in Multiple Linear Regression (Case Study: Costumer Price Index Data (CPI) for The Province of East Borneo)

Hisintus Suban Hurint[1],Ika Purnamasari[2],Memi Nor Hayati[3]
[1]Mahasiswa Program Studi Statistika FMIPA Universitas Mulawarman
[2,3]Dosen Program Studi Statistika FMIPA Universitas Mulawarman
Email: sintus_hurint@yahoo.co.id

Abstract

Method of Ordinary Least Square (OLS) on the regression analysis is a method which is often used to estimate the parameters. In the OLS method, there are several assumptions that must be fulfilled, these assumptions are often not fulfilled when the data contains outlier, so need a method that are robust to the presence of outliers. In this research, studied method of robust regression with MM-estimation. MM-estimation is a combination of estimation methods that have a high breakdown point, namely the Scale estimation(S-estimation) and Least Trimmed Square estimation (LTS estimation) and the method that have higher efficiency point, namely the Maximum Likelihood Type estimation (M-estimation). The first step in the MM-estimation is to find the S-estimator, then set the parameter regression using the M-estimation. The purpose of this study was to determine the effect of price index of foodstuffs (X_1), the price index of education (X_2), and the price index of health (X_3) to the CPI for the province of east borneo, where the CPI data contains outliers, namely observation to 13, 31,and 32. Robust regression equation to MM-estimation obtained is $\hat{Y}$ = *67,413 + 0,051* X_1 + *0,189* X_2 + *0,125* X_3.

Keywords : Method of Moment Estimation (MM-Estimation), outliers, regression analysis, robust regression.

Pendahuluan

Analisis regresi merupakan suatu teknik analisis data dalam statistika yang bertujuan untuk mengetahui hubungan antara beberapa variabel yang terdiri dari satu variabel terikat (respon) dan satu atau lebih variabel bebas (prediktor).Dalam analisis regresi, metode yang sering digunakan untuk mengestimasi model adalah Metode Kuadrat Terkecil (MKT) atau *Ordinary Least Square* (OLS). Metode OLS dikenal sebagai metode estimasi terbaik, namun metode ini sangat peka terhadap adanyapencilan atau *outlier*. Pencilan merupakan suatu keganjilan dan menandakan suatu titik data yang letaknya jauh dari sebaran data yang lain (Draper dan Smith,1992).

Pencilan tidak dapat dibuang atau dihapus begitu saja dari pengamatan. Menurut Draper dan Smith (1992), adakalanya pencilanmemberikan informasi yang tidak bisa diberikan oleh titik data lainnya yang mungkin saja sangat penting dan perlu diselidiki lebih jauh.Pencilandapat diabaikan apabila setelah ditelusuri ternyata merupakan akibat dari kesalahan dalam mencatat pengamatan atau kesalahan ketika menyiapkan peralatan.

Salah satu metode untuk mengatasi pencilanadalah regresi *robust*. Menurut Olive (2005), regresi *robust* merupakan metode regresi yang digunakan untuk menganalisis data yang mengandung pencilan yang berpengaruh pada model.

Pada penelitian ini akan dibahas mengenai satu metode estimasi regresi *robust* yaitu estimasi-MM. Metode estimasi-MM diperkenalkan oleh Yohai pada tahun 1987. Metode ini merupakan gabungan dari metode estimasi LTS, estimasi-S dan estimasi-M (Chen, 2002).

Chandraningtyas (2013) pernah melakukan penelitian tentang regresi *robust* estimasi-MM yang berjudul “Regresi *Robus*t MM-Estimator untuk penanganan pencilan pada regresi linier berganda” dengan menggabungkan metode estimasi LTS dan metode estimasi-M.Oleh karena itu, dalam penelitian ini penulis menggunakan regresi *robust* metode estimasi-MM dengan menggabungkan metode estimasi-S dan metode estimasi-M pada model regresi linier berganda.

Regresi Linier Berganda

Analisis regresi adalah studi yang mempelajari tentang hubungan antara suatu variabel bebas dengan satu variabel terikat dengan tujuan untuk mengestimasi atau meramalkan nilai variabel terikat yang didasarkan pada nilai variabel bebas yang diketahui. Model regresi umum yang mengandung k variabel bebas dapat dituliskan sebagai berikut :

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \varepsilon \qquad (1)$$

Bila pengamatan mengenai $Y, X_1, X_2, \ldots, X_k$ dinyatakan masing-masing dengan $Y_i, X_{i1}, X_{i2}, \ldots, X_{ik}$ dan residual ε_i maka persamaan (1) dapat dituliskan sebagai berikut :

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik} + \varepsilon_i$$

$$i = 1,2,\ldots,n \qquad (2)$$

Dimana :
Y_i : Variabel terikat pada pengamatan ke-i.
β : Parameter model regresi
X_i: Variabel bebas ke-i
ε_i : Residual data ke-i dengan mean nol dan variansi σ^2 atau $\varepsilon_i \sim N(0,\sigma^2)$
(Montgomery dan Peck, 1992).

Dalam bentuk matriks persamaan (2) menjadi :

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_1 & X_1 & \cdots & X_{1k} \\ 1 & X_2 & X_2 & \cdots & X_{2k} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} (3)$$

Persamaan (3) dapat ditulis dalam persamaan matriks :

$$\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \qquad (4)$$

dimana :
$\boldsymbol{y}$: Vektor variabel terikat berukuran $n \times 1$.
$\mathbf{X}$: Matriks variabel bebas berukuran $n \times (k+1)$.
$\boldsymbol{\beta}$: Vektor parameter berukuran $(k+1)$.
$\boldsymbol{\varepsilon}$: Vektor residual berukuran $n \times 1$.

Ada sebanyak $k+1$ parameter yang harus ditaksir dari data, termasuk β_0. Taksirannya akan ditulis dalam bentuk umum persamaan :

$$\hat{\boldsymbol{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} + \boldsymbol{\varepsilon} \qquad (5)$$

Salah satu metode yang sering digunakan untuk menaksir koefisien regresi adalah Metode OLS.

Metode *Ordinary Least Square* (OLS)

Metode OLS pertama kali dikemukakan oleh Carl Friedrich Gauss, seorang ahli matematika Jerman.Metode OLS merupakan metode yang paling banyak digunakan untuk mengestimasi parameter-parameter regresi dibandingkan metode-metode lainnya.OLS ini mampu mendapatkan $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_k$ yang *Best Linear Unbiased Estimator* (BLUE) sehingga menyebarnya persamaan regresi sedekat mungkin pada data aktualnya (Sembiring, 1995).

Persamaan metode OLS dalam bentuk matriks adalah:

$$(\mathbf{X}^T\mathbf{X})\hat{\boldsymbol{\beta}} = \mathbf{X}^T\mathbf{y} \text{ atau } \hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \qquad (6)$$

dengan :

$$\boldsymbol{y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \mathbf{X} = \begin{bmatrix} 1 & X_1 & X_1 & \cdots & X_{1k} \\ 1 & X_2 & X_2 & \cdots & X_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_n \end{bmatrix}$$

$$\mathbf{X}^T = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_1 & X_2 & \cdots & X_{n1} \\ X_1 & X_2 & \cdots & X_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ X_{1k} & X_{2k} & \cdots & X_n \end{bmatrix}, \hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_k \end{bmatrix},$$

$$\mathbf{X}^T\mathbf{X} =$$

$$\begin{bmatrix} n & \sum_{i=1}^{n} X_{i1} & \sum_{i=1}^{n} X_{i2} & \cdots & \sum_{i=1}^{n} X_{ik} \\ \sum_{i-1}^{n} X_{i1} & \sum_{i-1}^{n} (X_{i1})^2 & \sum_{i-1}^{n} X_{i1}X_{i2} & \cdots & \sum_{i-1}^{n} X_{i1}X_{ik} \\ \sum_{i-1}^{n} X_{i2} & \sum_{i-1}^{n} X_{i1}X_{i2} & \sum_{i-1}^{n} (X_{i2})^2 & \cdots & \sum_{i-1}^{n} X_{i2}X_{ik} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{i-1}^{n} X_{ik} & \sum_{i-1}^{n} X_{i1}X_{ik} & \sum_{i-1}^{n} X_{i2}X_{ik} & \cdots & \sum_{i-1}^{n} (X_{ik})^2 \end{bmatrix}$$

$$\mathbf{X}^T\boldsymbol{y} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_1 & X_2 & \cdots & X_{n1} \\ X_1 & X_2 & \cdots & X_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ X_{1k} & X_{2k} & \cdots & X_n \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} =$$

$$\begin{bmatrix} \sum_{i-1}^{n} Y_i \\ \sum_{i-1}^{n} X_{1i}Y_i \\ \sum_{i=1}^{n} X_{2i}Y_i \\ \vdots \\ \sum_{i-1}^{n} X_k\, Y_i \end{bmatrix} \qquad (7)$$

Uji Asumsi Model Regresi Linier Berganda

Pendugaan koefisien regresi (parameter) dapat dilakukan dengan berbagai metode diantaranya adalah metode kuadrat terkecil atau OLS.Metode ini paling mudah dan sederhana dibandingkan metode lainnya, sehingga sering digunakan oleh para peneliti.Namun ada beberapa asumsi yang harus dipenuhi dari penggunaan metode OLS tersebut. Asumsi utama yang mendasari pendugaan koefisien regresi dengan menggunakan metode OLS adalah (Draper dan Smith, 1992) :

a. Normalitas residual, regresi linier klasik mengasumsikan bahwa tiap $£_i$ mengikuti distribusi normal, $£_i \sim N(0,\pm^2)$.
b. Non multikolinearitas atau tidak terhadap hubungan linier yang sempurna atau pasti di antara beberapa atau semua variabel yang menjelaskan model regresi.
c. Non autokorelasi antar residual, berarti kov $(\varepsilon_i, \varepsilon_j) = 0,\ i \neq j$

d. Non heteroskedastisitas, var $(\varepsilon_i) = \sigma^2$ yaitu variansi dari sumber residual adalah konstan atau homokedastisitas untuk setiap i, i=1,2…,n.

Keempat asumsi tersebut sering disebut dengan asumsi klasik dari metode OLS.Jika asumsi-asumsi tersebut tidak terpenuhi, maka kesimpulan yang diperoleh akan menjadi bias dan model yang diperoleh tidak dapat dijadikan sebagai model prediksi. Sehingga untuk memastikan bahwa kesimpulan yang diperoleh benar dan parameternya yang diduga bersifat *BLUE* maka perlu dilakukan pengujian terhadap asumsi tersebut.

Pencilan (*Outlier*)

Menurut Ferguson (1961), pencilan adalah suatu data yang menyimpang dari sekumpulan data yang lain, Menurut Barnett (1981), pencilan adalah data yang tidak mengikuti sebagian besar pola dan terletak jauh dari pusat data. Menurut Sembiring (1995), pencilan adalah data yang jauh dari pusat data dan adakalanya pencilan memberikan informasi yang tidak bisa diberikan oleh titik data lainnya.

Secara umum data pencilan adalah data yang berbeda dari data-data lainnya pada suatu hasil, kemungkinan nilainya terlalu besar (lebih besar dari nilai pengamatan pada umumnya) atau terlalu kecil. Pencilan merupakan suatu keganjilan dan menandakan suatu titik yang khas dibanding pengamatan lain. Oleh karena itu pencilanperlu diperiksa secara seksama, kemungkinan saja alasan keganjilan dapat diketahui dan berpengaruh besar terhadap koefisien regresi (Draper dan Smith, 1992).

Pendeteksian Pencilan Dengan Diagram Pencar (*Scatter Plot*)

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

Metode diagram pencar (*scatter plot*) ini dilakukan dengan cara memplot data dengan pengamatan ke-i (i = 1,2, …, n). Selain itu, jika sudah didapatkan model regresi maka dapat dilakukan dengan cara memplot data nilai residual (ε) dengan data variabel terikat (Y). Jika terdapat satu atau beberapa data yang terletak jauh dari pola kumpulan data keseluruhan maka hal ini mengindikasikan adanya pencilan (Soemartini, 2007).

Pendeteksian Pencilan dengan Uji *Difference in Fit Standardized* (DfFITS)

Uji DfFITS digunakan untuk mendeteksi pencilan yang berpengaruh terhadap nilai variabel terikat (Y) , hipotesis yang digunakan adalah:

H_0 :Pencilan ke – i tidak berpengaruh, i=1,2,3...n
H_1 : Pencilan ke – i berpengaruh

Menurut Montgomery dan Peck (1992), rumus DfFITS didefinisikan sebagai berikut:

$$\text{DfFITS}= \varepsilon_i \left[\frac{n-k-1}{J\;(1-h_{i})-\varepsilon_i^2}\right]^{\frac{1}{2}}\left[\frac{h_i}{1-h_i}\right]^{\frac{1}{2}} \qquad (8)$$

dimana :
ε_i= Nilai residual ke-i, i = 1, 2, 3, ...n
JKG = Jumlah Kuadrat Residual
h_{ii}= Nilai *Leverage*
n= Banyaknya pengamatan

Leverage adalah pengamatan dengan nilai ekstrim pada variabel bebas atau ukuran jauhnya variabel bebas menyimpang dari rata-ratanya.Nilai h_{ii} didapatkan dengan rumus :

$$\mathbf{H} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T \quad (9)$$

Dimana h_{ii} merupakan elemen-eleman diagonal dari matriks hat **H** (matriks topi). **H** adalah matriks n x n.

$$\mathbf{H} = \begin{bmatrix} h_1 & h_1 & \cdots & h_{1n} \\ h_2 & h_2 & \cdots & h_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n1} & h_{n2} & \cdots & h_n \end{bmatrix},$$

$$\mathbf{X} = \begin{bmatrix} 1 & X_1 & X_1 & \cdots & X_{1k} \\ 1 & X_2 & X_2 & \cdots & X_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_n \end{bmatrix}$$

$$\mathbf{X}^T\mathbf{X} = \begin{bmatrix} n & \sum_{i=1}^{n} X_{i1} & \sum_{i=1}^{n} X_{i2} & \cdots & \sum_{i=1}^{n} X_{i} \\ \sum_{i-1}^{n} X_{i1} & \sum_{i-1}^{n} (X_{i1})^2 & \sum_{i-1}^{n} X_{i1}X_{i2} & \cdots & \sum_{i-1}^{n} X_{i1}X_{i} \\ \sum_{i-1}^{n} X_{i2} & \sum_{i-1}^{n} X_{i1}X_{i2} & \sum_{i-1}^{n} (X_{i2})^2 & \cdots & \sum_{i-1}^{n} X_{i2}X_{i} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{i-1}^{n} X_{i} & \sum_{i-1}^{n} X_{i1}X_{i} & \sum_{i-1}^{n} X_{i2}X_{i} & \cdots & \sum_{i-1}^{n} (X_{i})^2 \end{bmatrix}$$

$$\mathbf{X}^T = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_1 & X_2 & \cdots & X_{n1} \\ X_1 & X_2 & \cdots & X_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ X_{1k} & X_{2k} & \cdots & X_n \end{bmatrix} \qquad (10)$$

Menurut Montgomery dan Peck (1992), nilai kritis untuk DfFITS adalah $2\sqrt{\frac{k+1}{n}}$,dimana k adalah banyaknya parameter sedangkan n adalah banyaknya data penelitian.

Adapun kriteria pengujian yang melandasi keputusan adalah:

$$\text{DfFITS} = \begin{cases} \leq 2\sqrt{\dfrac{k+1}{n}}, H_0 \text{ diterima} \\ > 2\sqrt{\dfrac{k+1}{n}}, H_0 \text{ ditolak} \end{cases} \quad (11)$$

Jika nilai DfFITS lebih besar dari nilai kritis maka menolak H_0, jadi pencilan ke-*i* berpengaruh dari himpunan data sehingga menyebabkan penduga dari Y_i berubah (Soemartini, 2007).

Regresi *Robust*

Regresi *robust* diperkenalkan oleh Andrews (1972) dan merupakan metode regresi yang digunakan ketika distribusi dari nilai residual tidak normal dan adanya beberapa pencilan yang berpengaruh pada model.Metode ini merupakan alat penting untuk menganalisa data yang dipengaruhi oleh pencilan sehingga dihasilkan model yang *robust* terhadap pencilan.Suatu estimasi yang resistan adalah relatif tidak terpengaruh oleh perubahan besar pada sebagian kecil data atau perubahan kecil pada sebagian besar data.

Beberapa estimasi parameter dalam regresi *robust* adalah estimasi *Maximum Likelihood Type* (M), estimasi *Least Trimmed Square* (LTS), estimasi *Least Median Square* (LMS), estimasi *Scale* (S), estimasi *Method of Moment* (MM).

Regresi *Robust* Estimasi-M

Salah satu caramencari koefisien parameter pada regresi *robust* yang paling penting dan paling luas digunakan adalah estimasi-M. Pada prinsipnya estimasi-M merupakan estimasi yang meminimumkan suatu fungsi objektif ρ.

$$\min_{\beta} \sum_{i=1}^{n} \rho(\varepsilon_i) = \min_{\beta} \sum_{i=1}^{n} \rho\left(Y_i - \sum_{j=0}^{k} X_i\ \beta_j\right)$$

Fungsi ρ merupakan representasi pembobot dari residual. Untuk memperoleh suatu skala invariant dari estimasi ini, biasanya dilakukan dengan menyelesaikan persamaan:

$$\min_{\beta} \sum_{i-1}^{n} \rho\left(\frac{\varepsilon_i}{\hat{\sigma}}\right) \quad (12)$$

Pada metode estimasi-M terdapat 2 jenis pembobot yang dapat digunakan, antara lain :

1. Fungsi pembobot yang disarankan oleh Huber memakai fungsi objektif

$$\rho(u_i) = \begin{cases} \dfrac{1}{2}u_i^2 & |u_i| \leq c \\ |u_i|c - \dfrac{1}{2}c^2 & |u_i| > c \end{cases}$$

dan, fungsi pembobot

$$w(u_i) = \begin{cases} 1 & |u_i| \leq c \\ \frac{c}{|u_i|} & |u_i| > c \end{cases} \quad (13)$$

Dengan nilai c = 1,345

2. Fungsi pembobot yang disarankan oleh Tukey Bisquare memakai fungsi objektif

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

dan, fungsi pembobot

$$w(u_i) = \begin{cases} \left[1 - \left(\frac{u_i}{c}\right)^2\right]^2 & |u_i| \leq c \\ 0 & |u_i| > c \end{cases} \quad (14)$$

dengan nilai c = 4,685

Regresi *Robust* Estimasi-S

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

$$\hat{\beta}_s = \arg m \quad \hat{\sigma}(\varepsilon_1, \varepsilon_2, \ldots \varepsilon_n) \quad (15)$$

dengan menentukan nilai estimator skala *robust* $(\hat{\sigma})$ yang minimum dan memenuhi persamaan

$$m \quad \sum_{i-1}^{n} \rho\left(\frac{Y_i - \sum_{i=1}^{k} X_i\ \beta_j}{\hat{\beta}_s}\right) \quad (16)$$

dengan,

$$\hat{\sigma} = \sqrt{\frac{n\sum_{i=1}^{n}(\varepsilon_i{}^2) - (\sum_{i=1}^{n}\varepsilon_i)^2}{n(n-1)}} \quad (17)$$

definisi dari fungsi pembobot adalah

$$w(u_i) = \frac{\psi(u_i)}{u_i} = \begin{cases} \left[1 - \left(\frac{u_i}{c}\right)^2\right]^2 & |u_i| \leq c \\ 0 & |u_i| > c \end{cases} \quad (18)$$

dimana, $u_i = \frac{\varepsilon_i}{\hat{\sigma}}$ dan c = 1,547 (19)

(Chen, 2002)

Regresi *Robust* Estimasi-MM

Metode estimasi-MM dikenalkan oleh Yohai pada tahun 1987. Estimasi ini merupakan gabungan dari metode estimasi yang mempunyai nilai *breakdown* tinggi (metode LTS atau S) dan metode estimasi-M.Langkah pertama dalam estimasi ini adalah mencari estimator-S, kemudian menetapkan parameter-parameter regresi menggunakan estimasi-M. Bentuk dari metode estimasi-MM adalah

$$\hat{\beta}_M = \arg\min \sum_{i=1}^{n} \rho\left(\frac{\varepsilon_i}{\hat{\sigma}}\right)$$

$$= \arg\min \sum_{i=1}^{n} \rho\left(\frac{y_i - \sum_{j=0}^{k} x_i\ \beta_j}{\hat{\sigma}}\right) \quad (20)$$

Fungsi pembobot pada regresi *robust* estimasi-MM yang sering digunakan adalah fungsi pembobot Tukey Bisquare. Kriteria fungsi pembobot Tukey Bisquare adalah sebagai berikut:

$$w(u_i) = \begin{cases} \left[1 - \left(\frac{u_i}{c}\right)^2\right]^2 & |u_i| \leq c \\ 0 & |u_i| > c \end{cases} \quad (21)$$

dimana, nilai $u_i = \frac{\varepsilon_i}{\hat{\sigma}}$ dan nilai c = 4,685 (22)

Menurut Yohai (1987), pilihan estimasi-MM untuk $\hat{\sigma}$ adalah skala $\hat{\sigma}$ dari estimasi-S pada iterasi ke-n yang bersifat tetap.

Menurut Montgomery dan Peck (1992), estimasi parameter pada regresi *robust* estimasi-MM dilakukan dengan estimasi *Iteratively Reweighted Least Square* (IRLS). Iterasi ini membutuhkan proses iterasi dimana nilai w_i akan berubah nilainya di setiap iterasi. Persamaan dalam bentuk matriks dapat ditulis sebagai berikut:

$$\mathbf{X}^{\mathrm{T}}\mathbf{W}^{l}\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}^{\mathrm{T}}\mathbf{W}^{l}\boldsymbol{y} \quad (23)$$

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

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^{\mathrm{T}}\mathbf{W}^{l}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{W}^{l}\boldsymbol{y} \quad l = 1, 2, 3 \ldots. \quad (24)$$

Dimana :

$$\boldsymbol{y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \mathbf{X} = \begin{bmatrix} 1 & X_1 & X_1 & \cdots & X_{1k} \\ 1 & X_2 & X_2 & \cdots & X_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_n \end{bmatrix}$$

$$\mathbf{X}^{\mathrm{T}} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_1 & X_2 & \cdots & X_{n1} \\ X_1 & X_2 & \cdots & X_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ X_{1k} & X_{2k} & \cdots & X_n \end{bmatrix}$$

$$\mathbf{W} = \begin{bmatrix} w_{1,1} & 0 & \cdots & 0 \\ 0 & w_{1,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & w_{1,n} \end{bmatrix}$$

Iterasi akan berhenti sampai didapatkan nilai $\hat{\beta}_j$ yang konvergen yaitu selisih nilai $\hat{\beta}_j^{\,l+1}$ dan $\hat{\beta}_j^{\,l}$ mendekati 0

Indeks Harga Konsumen (IHK)

IHK merupakan salah satu indikator ekonomi penting yang dapat memberikan informasi mengenai perkembangan harga barang dan jasa yang dibayar oleh konsumen atau masyarakat, khususnya masyarakat perkotaan.Kegunaan utama dari IHK adalah untuk menilai daya beli uang.Pada saat harga naik, nilai sebenarnya dari uang atau daya beli menurun sehingga hanya dapat membeli kuantitas yang lebih sedikit dari barang atau jasa.

IHK dihitung dengan rumus *Laspeyres* termodifikasi. Dalam perhitungan rata-rata harga komoditas, ukuran yang digunakan adalah rata-rata aritmatik, tetapi untuk beberapa komoditas seperti beras, minyak goreng, bensin, dan sebagainya digunakan rata-rata geometri :

$$L = \frac{\sum P_T Q_0}{\sum P_0 Q_T} \times 100 \quad (23)$$

dimana,
P_t = harga komoditi pada periode tertentu
P_0 = harga komoditi pada periode dasar
Q_0 = kuantitas komoditi pada periode dasar
(BPS Provinsi Kaltim, 2014)

Metodologi Penelitian

Data yang digunakan dalam penelitian ini adalah data bulanan yang diambil dari Badan Pusat Statistik (BPS) Provinsi Kalimantan Timur dari bulan Januari 2011 sampai dengan Desember 2014. Variabel dalam penelitian ini terdiri dari variable terikat yaitu data IHK Provinsi Kalimantan Timur dan variabel respon yaitu data indeks bahan makanan, pendidikan dan kesehatan.

Adapun teknik analisis data dalam penelitian ini adalah :

1. Analisis deskriptif
2. Estimasi Parameter Model dengan metode OLS.
3. Pengujian Signifikansi Parameter
   a. Pengujian signifikansi parameter secara simultan (Uji F)
   b. Pengujian signifikansi parameter secara parsial (Uji t)
4. Pengujian asumsi regresi linier berganda
5. Pendeteksian Pencilan
   Pendeteksian pencilan dilakukan dengan cara sebagai berikut :
   a. Diagram pencar (*Scatter Plot)*
   b. Uji DfFits
6. Metode regresi *robust* estimasi-MM
   tahapan-tahapan regresi *robust* dengan estimasi-MM adalah :
   1. Menghitung estimator awal koefisien $\hat{\beta}_j^{(1)}$ dan residual $\varepsilon_i^{(1)}$ dari regresi *robust* estimasi-S. Langkah-langkah perhitungannya adalah sebagai berikut :
      a. Menghitung parameter regresi awal dengan metode OLS sehingga didapatkan nilai $\hat{\beta}_0$ dan mendapatkan nilai $\varepsilon_i = Y_i - \hat{Y}_i$ untuk mencari nilai pembobot awal.

b. Mencari nilai $\hat{\sigma}$ dengan menggunakan persamaan (17).
c. Menghitung nilai u_i dengan menggunakan persamaan (19).
d. Mencari nilai pembobot $w(u_i)$ sebagai nilai pembobot awal dengan dengan menggunakan kriteria fungsi pembobot Tukey Bisquare pada persamaan (18).
e. Mencari nilai parameter regresi *robust* dengan metode estimasi-S menggunakan persamaan (24).
f. Mengulang langkah a sampai e hingga didapatkan nilai $\hat{\beta}_j$ yang konvergen yaitu selisih nilai $\hat{\beta}_j^{\,i+1}$ dan $\hat{\beta}_j^{\,i}$ sampai mendekati 0.

2. Menghitung nilai residual $\varepsilon_i = Y_i - \hat{Y}_i$ dari persamaan regresi *robust* estimasi-S yang didapat pada langkah 1.
3. Menghitung nilai u_i dengan menggunakan nilai $\hat{\sigma}$ pada estimasi-S
4. Menghitung nilai pembobot $w(u_i)$ sebagai nilai pembobot awal dengan menggunakan persamaan (21) dan menggunakan kriteria fungsi pembobot Tukey Bisquare pada persamaan (14)
5. Mencari nilai parameter regresi *robust* estimasi-MM menggunakan persamaan (24).
6. Mengulang langkah 2 sampai 5 hingga didapatkan nilai $\hat{\beta}_j$ yang konvergen yaitu selisih nilai $\hat{\beta}_j^{\,i+1}$ dan $\hat{\beta}_j^{\,i}$ sampai mendekati 0.

Hasil dan Pembahasan

Tabel 4.1 Analisis Deskriptif

Variabel	maks	min	Rata-rata	variansi
IHK	161,17	113,16	137,58	225,69
Bahan makanan	299,95	123,92	174,78	1817,63
Pendidikan	288,04	124,22	209,96	2757,05
Kesehatan	216,67	116,44	174,27	771,86

Berdasarkan Tabel 4.1 didapatkan hasil bahwa rata-rata harga indeks untuk IHK sebesar 137,58, rata-rata harga indeks bahan makanan, pendidikan dan kesehatan yang dikonsumsi oleh konsumen masing-masing sebesar 174,78, 209,96 dan 174,27. Standar deviasi untuk harga indeks IHK sebesar 15,02. Harga indeks Bahan makanan memiliki standar deviasi sebesar 42,63, harga indeks pendidikanmemiliki standar deviasi sebesar 52,51 dan harga indeks kesehatan memiliki standar deviasi sebesar 27,78. IHK, bahan makanan, pendidikan, dan kesehatanmemiliki variansi masing-masing sebesar 225,69,1.817,63, 2.757,05 dan 771,86.

Estimasi Parameter Model dengan Metode OLS

Estimasi parameter model regresi dengan metode OLS,dilakukan dengan bantuan *software* SPSS 21, berdasarkan hasil analisis diperoleh estimasi model regresi untuk data IHK Provinsi Kalimantan Timur Tahun 2011-2014 adalah sebagai berikut :

$$\hat{Y} = 67{,}182 + 0{,}041\,X_1 + 0{,}190\,X_2 + 0{,}134\,X_3$$

dengan X_1adalah bahan makanan, X_2 adalah pendidikan dan X_3adalah kesehatan. Sebelum di analisis lebih lanjut maka dilakukan pengujian secara simultan (uji F), pengujian variabel bebas secara parsial (uji t), dan koefisien determinasi (R^2), dan dilanjutkan dengan pengujian asumsi.

Pengujian Signifikansi Parameter Secara Simultan (Uji F)

Uji F menunjukkan apakah semua variabel bebas yang dimasukkan dalam model mempenyai pengaruh secara bersama-sama terhadap variabel terikat. Dengan menggunakan bantuan software SPSS 21, maka diperoleh nilai F_{hit} = 489,655 $>F_{(0,0\ ;3;4\ )}$= 2,82, maka dapat disimpulkan bahwa paling sedikit ada satu variabel yang berpengaruh antara bahan makanan, pendidikan dan kesehatan terhadap IHK.

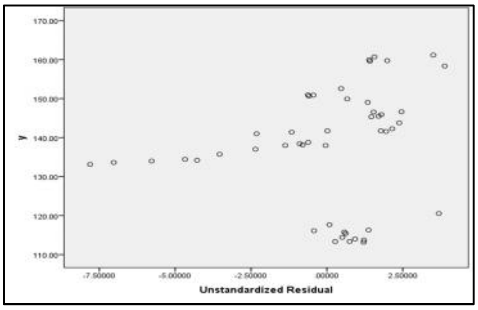
Pengujian Signifikansi Parameter Secara Parsial (Uji t)

Pengujian parameter secara parsial (uji t) bertujuan untuk melihat ada tidaknya pengaruh secara individual dari semua variabel bebas (bahan makanan, pendidikan, dan kesehatan. Dengan menggunakan *software* SPSS 21 diperoleh nilai *p-value* untuk parameter β_1 = 0,003, β_2 = 0,000, β_3 = 0,001 < α = 0,05, makadisimpulkan bahwa variabel bahan makanan, pendidikan dan kesehatanberpengaruh terhadap IHK.

Pengujian Asumsi Regresi Linier Berganda

Hasil dari pengujian ini menunjukkan bahwa terdapat satu asumsi yang tidak terpenuhi yang menyebabkan masalah heteroskedastisitas dalam model regresi.Asumsi tersebut tidak terpenuhi karena mungkin saja terdapat pencilan pada data, sehingga perlu dilakukan pendeteksian terhadap pencilan.

Pendeteksian Pencilan



Gambar 4.1 Diagram pencar

Gambar 4.1 memperlihatkan bahwa ada data yang terletak jauh dari kumpulan data, yaitu data yang bertitik hitam tebal. Data tersebut yang disebut dengan pencilan. Untuk mengetahui lebih jelasnya data mana yang teridentifikasi pencilan dapat dilihat pada hasil uji *DfFITS*

Dari hasil uji DfFITS dapat diketahui bahwa data pengamatan ke-13, 31, dan 32 dengan nilai pencilan sebesar 0,5861, 0,5919, 0,84931memiliki nilai yang lebih besar dari kriteria uji sebesar 0,57. Maka dapat disimpulkan bahwa data pengamatan ke-13, 31, dan 32 merupakan data pencilan yang berpengaruh terhadap model regresi. Apabila terdapat pencilan, maka analisis akan dilanjutkan dengan metode regresi *robust* estimasi-MM.

Metode Regresi *Robust* Estimasi-MM

Estimasi awal untuk mengestimasi parameter β dengan regresi *robust* estimasi-MM adalah menggunakan regresi *robust* estimasi-S.Iterasi dihentikan jika telah didapatkan nilai $\hat{\beta}_j$ yang konvergen yaitu selisih nilai $\hat{\beta}_j^{i+1}$dan $\hat{\beta}_j^{i}$mendekati 0. Dengan bantuan *software* SPSS 21 didapatkan hasil iterasi yang konvergen, yaitu iterasi ke-12 dan ke-13 mendekati atau sama dengan 0.

Tabel 3. Hasil parameter regresi *robust* estimasi-S

$\hat{\beta}_j$	Iterasi 12	Iterasi 13	Selisih
$\hat{\beta}_0$	68,793	68,793	0
$\hat{\beta}_1$	0,074	0,074	0
$\hat{\beta}_2$	0,202	0,202	0
$\hat{\beta}_3$	0,084	0,084	0

Jadi, persamaan regresi *robust* estimasi-S adalah

$\hat{Y}$ = 68,793 + 0,074 X_1 + 0,202 X_2 + 0,084 X_3

Langkah selanjutnya, menggunakan regresi *robust* estimasi-S untuk mendapatkan penyelesaian dari regresi *robust* estimasi-MM.Iterasi dihentikan jika telah didapatkan nilai $\hat{\beta}_j$ yang konvergen yaitu selisih nilai $\hat{\beta}_j^{i+1}$dan $\hat{\beta}_j^{i}$mendekati 0. Dengan bantuan *software* SPSS 21 didapatkan hasil iterasi yang konvergen, yaitu iterasi ke-5 dan ke-6 mendekati atau sama dengan 0.

Tabel 4. Hasil parameter regresi *robust* estimasi-MM

$\hat{\beta}_j$	Iterasi 5	Iterasi 6	Selisih
$\hat{\beta}_0$	67,413	67,413	0
$\hat{\beta}_1$	0,051	0,051	0
$\hat{\beta}_2$	0,189	0,189	0
$\hat{\beta}_3$	0,125	0,125	0

Jadi, persamaan regresi *robust* estimasi-MM adalah

$\hat{Y}$ = 67,413 + 0,051 X_1 + 0,189 X_2 + 0,125 X_3

Interpretasi untuk persamaan diatasadalah jika indeks harga bahan makanan (X_1), indeks harga pendidikan (X_2), dan indeks harga kesehatan (X_3) sama dengan 0, maka harga rata-rata IHK Provinsi Kalimantan Timur sebesar 67,413. Setiap penambahan satuan harga indeksbahan makanan (X_1) akan meningkatkan harga rata-rata IHK Provinsi Kalimantan Timur sebesar 0,051, apabila indeks harga pendidikan (X_2), dan indeks harga kesehatan (X_3) tetap. Setiap penambahan satuan harga indeks pendidikan (X_2)akan meningkatkan IHK Provinsi Kalimantan Timur sebesar 0,189 apabila indeks harga bahan makanan (X_1), dan indeks harga kesehatan (X_3) tetap. Setiap penambahan satuan harga indekskesehatan (X_3) akan meningkatkan harga rata-rata IHK Provinsi Kalimantan Timur sebesar 0,125 apabila indeks harga bahan makanan (X_1), dan indeks harga pendidikan (X_2) tetap.

Kesimpulan

Berdasarkan hasil dari penelitian yang telah dibahas, maka dapat ditarik kesimpulan sebagai berikut :

1. Model regresi linier berganda pada data Indeks Harga Konsumen (IHK) Provinsi Kalimantan Timur dengan metode OLS adalah: $\hat{Y}$ = 67,182 + 0,041 X_1 + 0,190 X_2 + 0,134 X_3

   dengan $\hat{Y}$ adalah IHK Provinsi Kalimantan Timur, X_1 adalah indeks harga bahan makanan, X_2 adalah indeks harga pendidikan, X_3 adalah indeks harga kesehatan.
2. Berdasarkan identifikasi pencilan yang telah dilakukan, terdapat 3 pencilan pada data IHK Provinsi Kalimantan Timur, yakni pada pengamatan ke-13, 31, dan 32.
3. Model regresi *robust* pada data IHK Provinsi Kalimantan Timur dengan metode estimasi-MM adalah : $\hat{Y}$ = 67,413 + 0,051 X_1 + 0,189 X_2 + 0,125 X_3

   dengan $\hat{Y}$ adalah IHK Provinsi Kalimantan Timur, X_1 adalah indeks harga bahan makanan, X_2 adalah indeks harga pendidikan, X_3 adalah indeks harga kesehatan.

Daftar Pustaka

BPS (Badan Pusat Statistik) Provinsi Kalimantan Timur.2014. *Indeks Harga Konsumen (IHK) Provinsi Kalimantan Timur*. Katalog BPS : 1403.5107

Candraningtyas, Sherly. 2013. *Regresi Robust MM-Estimator untuk Penanganan Pencilan pada Regresi Linier Berganda*; Jurnal : Universitas Diponegoro.

Chen, C. 2002. *Robust Regression and outlier detection with the Robustreg Procedure*.Paper 265-27. North Carolina: SAS Institute.

Draper, N. R., H. Smith. 1992. *Applied Regression Analysis,* 3^{th}*edition.* Bandung: Alfabeta.

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

Sembiring, R.K. 1995. *Analisis Regresi*. Bandung: Penerbit ITB.

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

Yohai, V. J. 1987. *High Breakdown-Point And High Efficiency Robust estimates For Regression*. The Annals of Statistics.


ADA SATU HAL
YANG PERLU
KALIAN PAHAMI
TENTANG
BAGAIMANA
VIRUS CORONA
MENYEBAR


Virus akan menyebar ketika tetesan kecil ini
Masuk ke dalam
Mata
Hidung
Mulut
ORANG SAKIT
ORANG SEHAT
weimankow

Jadi jika kalian melihat seseorang yang sedang batuk/bersin/sakit, kalian bisa memilih untuk:

1. JAGA JARAK.
2 meter sampai 0,5 meter akan menjagamu tetap aman dari percikan yang besar.

Atau,

2 BERI MEREKA MASKER.
Jadi mereka dapat bersin/batuk di maskernya dan melindungi orang-orang yang ada di dekatnya.

Secara umum, ide yang bagus untuk menghindari kerumunan orang. Karena kalian tidak tahu siapa yang sebenarnya sakit.

Bagaimanapun, terkadang air liur orang sakit dapat menempel pada...

Dan jika kalian tidak sengaja menyentuh benda-benda tadi, lalu menyentuh wajah kalian sendiri,

Kalian semua mungkin dapat jatuh sakit.


Virus dapat bertahan hingga 24 jam di sebuah benda.
Sabun
Dan satu-satunya cara yang efektif untuk menghilangkannya adalah mencucinya dengan sabun.

Itulah sebabnya,
ada baiknya untuk
mengikuti

5

tindakan pencegahan
berikut ini.

1 JANGAN MENYENTUH WAJAH KALIAN (atau wajah siapapun juga)

Jika memang harus, cuci bersih dulu tangan kalian dengan sabun.

Bagaimana cuci tangan yang bersih?

2 Buang segera masker kalian jika terasa sudah tidak bersih, jangan gunakan masker lebih dari sehari!

Bakteri dapat berkembang di BAGIAN DALAM masker jika digunakan terlalu lama.

Jangan menyentuh bagian luar masker.

Jika terpaksa menyentuhnya, segera cuci tangan dengan sabun setelahnya.

3 Jangan berbagi MAKANAN, ALAT MAKAN, GELAS, dan HANDUK.

4 Buka & tutup pintu dengan siku atau bahu, jika memungkinkan.

Terakhir,

5 Selalu cuci tangan kalian dengan sabun...

- Sebelum makan

- Setelah dari tempat umum


JAGALAH KESEHATAN!
sabun
@weimankow

Referensi utama untuk informasi medis:

1. "How to Avoid the Coronavirus? Wash Your Hands", by Elizabeth Rosenthal, in The New York Times, Opinion. (Jan 28, 2020)

2. "The Wuhan Virus: How to stay Safe", by Laurie Garrett, in Foreign Policy, Report. (Jan 25, 2020)

3. "This animation shows how far your sneeze can actually travel." by Chia-Yi Hou and Andrea Schmitz, in Business Insider (Jan 21, 2020)

Baca artikel di atas untuk tips yang lebih banyak!

Sumber : HTTPS://M.FACEBOOK.COM/
STORY.PHP?
STORY_FBID=474262819924443&ID=108902046460524