

PERBANDINGAN REGRESI *OLS* DAN *ROBUST MM-ESTIMATION* DALAM KASUS DBD DI INDONESIA 2018

Oleh :

Esa Anindika Sari^{1)*}, Hanum Iftitah Rahma²⁾, M Ridwan Firdaus³⁾,
Wahyu Winarto⁴⁾, Yansi Indiyani⁵⁾, Rani Nooraeni^{6)*}

^{1,2,3,4,5,6}Politeknik Statistika STIS, Jakarta

¹*email : esasari16@gmail.com

²email : hanumiftitah10@gmail.com

³email : ridwanfirdaus@gmail.com

⁴email : wahyuwinarto192@gmail.com

⁵email : yansiindiyani@gmail.com

⁶email : raninoor@stis.ac.id

Abstrak

Analisis regresi merupakan salah satu cara untuk membentuk model hubungan antara variabel bebas dengan variabel terikat. Salah satu metode yang umum digunakan adalah metode *Ordinary Least Square (OLS)*. Namun apabila terdapat outlier, maka estimasi koefisien garis regresi dengan metode OLS menjadi tidak tepat. Untuk mengatasi hal tersebut, salah satu yang dapat dilakukan yaitu dengan menggunakan metode regresi robust. Metode estimasi parameter pada regresi robust yang digunakan adalah metode MM-Estimation karena memiliki breakdown point dan efisiensi yang tinggi. Penelitian ini akan menyajikan estimasi parameter dan perbandingan antara metode regresi OLS dengan metode robust MM-Estimation dalam kasus penyakit Demam Berdarah Dengue (DBD) di Indonesia pada tahun 2018. Variabel yang digunakan yaitu jumlah kasus DBD, persentase sanitasi layak, kepadatan penduduk, dan persentase rumah tangga kumuh. Dalam menentukan model terbaik yaitu dengan membandingkan standar error dan R-squared. Berdasarkan hasil dan pembahasan, metode Robust MM-Estimation memiliki standar error yang lebih kecil dibandingkan OLS. Selain itu R-squared Robust MM-Estimation lebih besar dibandingkan R-squared OLS. Pada metode Robust juga semua variabel signifikan, sedangkan pada OLS tidak ada satupun variabel yang signifikan. Sehingga dapat disimpulkan bahwa metode Regresi Robust MM-Estimation lebih baik dibandingkan metode OLS.

Kata kunci: Analisis regresi, OLS, Outlier, Robust, Demam Berdarah

1. PENDAHULUAN

Gujarati (2006) mendefinisikan analisis regresi sebagai kajian terhadap hubungan satu variabel yang disebut sebagai variabel yang diterangkan (*the explained variable*) dengan satu atau dua variabel yang menerangkan (*the explanatory variable*). Variabel pertama disebut juga sebagai variabel terikat dan variabel kedua disebut juga sebagai variabel bebas. Jika variabel bebas lebih dari satu, maka analisis regresi disebut regresi linier berganda. Disebut berganda karena pengaruh beberapa variabel bebas akan dikenakan kepada variabel terikat.

Umumnya, untuk mengestimasi koefisien regresi dalam model regresi, digunakan metode Ordinary Least Squares (OLS) yaitu dengan meminimumkan jumlah kuadrat residual. Penggunaan metode OLS harus memenuhi beberapa asumsi klasik antara lain asumsi normalitas, asumsi homoskedastisitas, asumsi non-multikolinieritas, dan asumsi non-autokorelasi. Jika semua asumsi tersebut terpenuhi maka penduga parameter yang diperoleh bersifat Best Linear Unbiased Estimator (BLUE). Namun, metode ini mempunyai kelemahan ketika terdapat outlier dalam data. Adanya outlier tersebut, mengakibatkan estimasi koefisien regresi yang

diperoleh menjadi kurang tepat. Dalam kasus yang seperti ini, analisis regresi robust merupakan metode yang tepat digunakan.

Regresi robust diperkenalkan oleh Andrews (1972) dan merupakan metode regresi yang digunakan ketika distribusi error tidak normal dan atau adanya beberapa outlier yang berpengaruh pada model (Ryan, 1997). Regresi *robust* dapat diterapkan pada data-data di berbagai bidang di antaranya bidang pendidikan, kesehatan, ekonomi, atau bidang sosial. Salah satu bidang yang memiliki urgensi tinggi dan berpengaruh besar bagi negara yaitu kesehatan. Dalam hal ini yang menjadi fokus penelitian adalah data mengenai masalah penyakit Demam Berdarah *Dengue* (DBD).

Penyakit Demam Berdarah *Dengue* (DBD) menjadi salah satu penyakit menular yang sering terjadi di Indonesia. DBD yang ditularkan melalui gigitan nyamuk dari genus *Aedes*, terutama *Aedes aegypti* atau *Aedes albopictus*, dapat muncul sepanjang tahun dan dapat menyerang seluruh kelompok umur.

Menurut World Health Organization (WHO), sebelum tahun 1970 hanya 9 negara yang mengalami epidemi dengue parah. Penyakit ini sekarang endemik di lebih dari 100 negara di wilayah WHO di Afrika, Amerika, Mediterania

Timur, Asia Tenggara, dan Pasifik Barat (WHO, 2018). Berdasarkan data WHO terjadi peningkatan penyakit DBD sebesar 75% di kawasan Asia Pasifik termasuk Indonesia dibandingkan kawasan lain.

2. METODE PENELITIAN

Regresi Linear Berganda

Regresi linear berganda adalah teknik statistika yang digunakan untuk menentukan model hubungan satu variabel respon (Y) dengan melibatkan lebih dari satu variabel prediktor hingga p variabel prediktor dimana banyaknya p kurang dari jumlah observasi (n). Adapun model regresi berganda sebagai berikut :

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip} + \varepsilon_i$$

Di mana Y_i merupakan nilai variabel dependen dalam pengamatan ke- i , $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ adalah parameter yang tidak diketahui nilainya, $X_{i1}, X_{i2}, \dots, X_{ip}$ adalah nilai dari variabel independen dari pengamatan ke- i , dan ε_i adalah error random dan berdistribusi normal dengan rata-rata nol dan varians σ^2 .

OLS (Ordinary Least Square)

Metode estimasi parameter yang digunakan adalah metode OLS (*Ordinary Least Square*), yaitu menduga koefisien regresi (β) dengan meminimumkan kesalahan (error) . Adapun penaksir parameternya adalah sebagai berikut :

$$\hat{\beta} = (X^T X)^{-1} X^T y$$

Dimana $\hat{\beta}$ adalah vektor dari parameter yang diestimasi berukuran $(p + 1) \times 1$, X adalah matriks variabel prediktor berukuran $n \times (p + 1)$ dan y vektor observasi dari variabel respon berukuran $n \times 1$

Uji Multikolinearitas

Menurut Ghazali (2011: 105) uji multikolinearitas digunakan untuk mengetahui ada atau tidaknya korelasi antar variabel independen dalam model regresi. Model regresi yang baik adalah yang tidak mengandung multikolinearitas. Mendeteksi multikolinieritas dapat melihat nilai *tolerance* dan *varian inflation factor* (VIF) sebagai tolak ukur. Apabila nilai *tolerance* $\leq 0,10$ dan nilai VIF ≥ 10 maka dapat disimpulkan bahwa dalam penelitian tersebut terdapat multikolinieritas (Ghozali, 2011: 106).

Uji Heteroskedastisitas

Menurut Ghazali (2011:139) uji heteroskedastisitas digunakan untuk menguji apakah dalam sebuah regresi terjadi ketidaksamaan varian dari residual dari suatu pengamatan ke pengamatan lain. Prasyarat yang harus terpenuhi dalam model regresi adalah tidak adanya gejala heteroskedastisitas. Pada penelitian ini akan dilakukan uji heteroskedastisitas menggunakan uji *glesjer* yaitu mengkorelasikan nilai absolut residual dengan masing-masing variabel. Hasil dari uji *glesjer* menunjukkan tidak ada heteroskedastisitas

apabila dari perhitungan nilai probabilitas signifikansinya diatas tingkat kepercayaan 5% (Ghozali, 2011: 143).

Uji Linieritas

Menurut Ghazali (2016), uji linieritas bertujuan untuk mengetahui apakah spesifikasi model linier yang digunakan sudah benar atau tidak. Uji linearitas dimaksudkan untuk mengetahui ada tidaknya hubungan secara linear antara variabel dependen terhadap setiap variabel independen yang hendak diuji. Jika suatu model tidak memenuhi syarat linearitas maka model regresi linear tidak bisa digunakan. Salah satu uji linieritas adalah Uji Ramsey Reset Test. Aturan untuk keputusan linearitas dapat dengan membandingkan nilai signifikansi dari F-statistic yang dihasilkan dari Ramsey Reset Test (menggunakan bantuan Eviews) dengan nilai alpha yang digunakan. Jika nilai signifikansi dari F-statistic $>$ alpha (0,05) maka nilai tersebut linear.

Uji Normalitas

Uji normalitas berfungsi untuk menguji apakah dalam sebuah model regresi, variabel pengganggu memiliki distribusi normal (Ghozali, 2011:160). Dalam penelitian ini akan digunakan uji *One Sample Kolmogorov-Smirnov* dengan menggunakan taraf signifikansi 0,05. Data dinyatakan berdistribusi normal jika signifikansi lebih besar dari 0,05.

Outlier

Menurut B Sembiring (1995) pencilan adalah data yang dari pusat data dan ada kalanya pencilan memberikan informasi yang tidak bisa diberikan oleh titik data lainnya. Keberadaan data outlier akan mengganggu proses analisis data dan harus dihindari dari beberapa hal. Dalam kaitannya dalam analisis regresi, outlier dapat menyebabkan hal-hal berikut (Soemartini, 2007: 7): (1) Residual yang besar dari model yang terbentuk $E(\varepsilon_i) \neq 0$ (2) Varians pada data tersebut menjadi lebih besar (3) Taksiran interval memiliki rentang yang lebar. Untuk mengidentifikasi apakah terdapat data outlier atau tidak, dapat menggunakan beberapa metode salah satunya yang akan dibahas adalah metode boxplot. Dengan metode boxplot suatu data dikatakan ada outlier jika nilai data pengamatan lebih kecil dari $Q1 - (1,5 * IQR)$ atau lebih besar dari $Q3 - (1,5 * IQR)$. Selain itu juga bisa menggunakan metode DfFITS (*Difference fitted value FITS*). *Difference fitted value FITS* merupakan metode yang menampilkan nilai perubahan dalam harga yang diprediksi bilamana kasus tertentu dikeluarkan dan yang sudah distandarkan. Apabila nilai DfFits $> 2 * \sqrt{(k+1/n)}$ maka terdapat outlier atau pencilan, dengan k banyaknya variabel independen dan n banyaknya observasi. Selanjutnya dilakukan analisis regresi menggunakan metode robust untuk data yang mengandung outlier agar hasil regresi yang dihasilkan lebih tepat.

Regresi Robust

Regresi Robust pertama kali diperkenalkan oleh Andrews (1972). Menurut Hurint (2006) Regresi Robust merupakan metode regresi yang digunakan ketika distribusi dari nilai residual tidak normal dan adanya beberapa pencilan berpengaruh pada model. Metode ini merupakan alat penting untuk menganalisa data yang dipengaruhi oleh pencilan sehingga dihasilkan model yang *robust* atau tahan terhadap pencilan.

Dalam regresi *robust*, dikenal beberapa jenis estimasi parameter, yaitu estimasi *Least Trimmed Square* (LTS-estimation), estimasi *Maximum Likelihood Type* (M-estimation), estimasi *Scale* (S-estimation), dan estimasi *Method of Moment* (MM-estimation).

Robust M-Estimation

Salah satu estimasi parameter regresi robust ialah estimasi-M. Menurut Yuliana (2008), estimasi-M berasal dari metode estimasi parameter Maximum Likelihood dan estimasi robust. Pada estimasi-M, diperlukan fungsi pembobot untuk memperoleh skala yang konstan atau invariant dari estimasi ini. Pada penelitian ini, digunakan fungsi pembobot Bisquare, yaitu:

$$w_i \begin{cases} \left[1 - \left[\frac{u_i}{4,689} \right]^2 \right]^2 & ; |u_i| \leq c \\ 0 & ; |u_i| > c \end{cases}$$

Di mana $u_i = \frac{e_i}{\hat{\sigma}_i}$ dan $c = 4,685$ untuk pembobot fungsi bisquare. Kemudian mencari estimasi parameter β_M dengan Weighted Least Square dan pembobot W , dengan rumus:

$$\hat{\beta}_M = (X'WX)^{-1}X'WY$$

Di mana W adalah matriks diagonal berdimensi 4×4 dengan baris a dan kolom b yang berisi $W_{ab} = w_i ; a = b = i$. Sehingga, dibentuk model persamaan:

$$Y_M = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

Robust S-Estimation

Menurut Rousseeuw (1984), estimasi-S dihitung berdasarkan skala residual dari estimasi-M. Kekurangan dari estimasi-S ialah tidak memperhitungkan distribusi dari data karena hanya menggunakan median sebagai penimbang. Metode ini menggunakan standar deviasi dari residual untuk mengatasi kekurangan dari median. Menurut Salibian dan Yohai (2006), estimasi-S didefinisikan sebagai $\hat{\beta}_S = \min_{\beta} \hat{\sigma}_S(e_1, e_2, \dots, e_n)$ dengan meminimumkan skala estimasi robust $\hat{\sigma}_S$. Sama seperti estimasi-M, estimasi-S juga menggunakan metode iterasi untuk mengestimasi parameter β_S . Nilai dari $\hat{\sigma}_S$ didapat dari:

$$\hat{\sigma}_i \begin{cases} \frac{\text{median}|e_i - \text{median } e_i|}{0,6745} & ; \text{Iterasi} = 1 \\ \sqrt{\frac{1}{nK} \sum_{i=1}^n w_i e_i^2} & ; \text{Iterasi} > 1 \end{cases}$$

Di mana e_i merupakan residual ke- i dari estimasi OLS. Selanjutnya, untuk mengestimasi parameter β_S , digunakan metode estimasi Weighted Least Square (WLS). Fungsi pembobot yang digunakan untuk estimasi-S:

$$w_i \begin{cases} \left[1 - \left(\frac{u_i}{1,547} \right)^2 \right]^2 & ; |u_i| \leq c \\ 0 & ; |u_i| > c \\ \frac{\rho(u)}{u^2} & ; \text{Iterasi} > 1 \end{cases}$$

Di mana $u_i = \frac{e_i}{\hat{\sigma}_i}$ dan $c = 1,547$ untuk *breakdown value* = 0,5. Setelah didapat pembobot untuk masing masing koefisien β_{OLS} , selanjutnya mencari estimasi parameter β_S dengan metode estimasi WLS dan pembobot W , yang didapat dari:

$$\hat{\beta}_S = (X'WX)^{-1}X'WY$$

Di mana W adalah matriks diagonal berdimensi 4×4 dengan baris a dan kolom b yang berisi $W_{ab} = w_i ; a = b = i$. Sehingga, dibentuk model persamaan regresi :

$$Y_S = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

Robust MM-Estimation

Menurut Chen (2002), MM-estimation pertama kali dikenalkan oleh Yohai (1987) untuk regresi linier. Menurut Yohai (1987), MM-estimation merupakan gabungan dari metode estimasi yang mempunyai nilai *breakdown* tinggi (metode LTS-estimation atau S-estimation) dan metode M-estimation memiliki tiga tahap. Tahap pertama yaitu menghitung estimasi parameter awal dengan metode yang memiliki *breakdown value* tinggi seperti S-estimation dan LTS-estimation. Tahap kedua, menghitung residual dan skala estimasi *robust* dengan menggunakan M-estimator. Ketiga, menghitung estimasi parameter akhir dengan M-estimation. MM-estimation didefinisikan sebagai berikut.

$$\hat{\beta} = \min \sum \rho \left(\frac{e_i}{\hat{\sigma}} \right) = \min \sum_{i=1}^n \rho \left(\frac{y_i - \sum_{j=1}^n x_{ij} \beta_j}{\hat{\sigma}_S} \right)$$

Demam Berdarah Dengue (DBD) dan Faktor-Faktor Penyebabnya

Demam Berdarah Dengue (DBD) adalah penyakit yang disebabkan oleh virus dengue, virus yang ditularkan dari nyamuk *Aedes sp*. Penyakit ini sering terjadi di negara-negara tropis seperti Indonesia. DBD memiliki gejala serupa dengan Demam Dengue, namun DBD memiliki gejala lain berupa sakit/nyeri pada ulu hati terus menerus,

pendarahan pada hidung, mulut, gusi, atau memar pada kulit (Pusdatin Kemenkes, 2018).

Faktor-faktor yang menyebabkan penyakit Demam Berdarah (DBD) antara lain yaitu sanitasi, kepadatan penduduk, dan rumah tangga kumuh. Faktor pertama yaitu sanitasi, Ehler dan Steel mengemukakan bahwa sanitasi adalah usaha-usaha pengawasan yang ditujukan terhadap faktor lingkungan yang dapat menjadi mata rantai penularan penyakit. Sanitasi adalah upaya kesehatan dengan cara memelihara dan melindungi kebersihan lingkungan dari subyeknya. Misalnya menyediakan air yang bersih untuk keperluan mencuci tangan, menyediakan tempat sampah untuk mawadahi sampah agar tidak dibuang sembarangan (Depkes RI, 2004). Sanitasi menjadi salah satu faktor yang mampu meningkatkan jumlah kasus DBD. Sanitasi yang kurang baik membuat nyamuk *Aedes sp.* berkembang biak dengan cepat. Faktor kedua yaitu kepadatan penduduk, kepadatan penduduk adalah jumlah penduduk di suatu wilayah dibandingkan dengan luas wilayah tersebut. Kepadatan penduduk mengakibatkan kepadatan pemukiman sehingga penularan penyakit DBD lebih cepat. Antonius (2005) mengatakan bahwa daerah yang terjangkit demam berdarah dengue pada umumnya adalah kelurahan yang padat penduduknya dan jarak antar rumah yang saling berdekatan memudahkan penularan penyakit. Faktor selanjutnya yaitu rumah tangga kumuh, menurut Badan Pusat Statistik (BPS), rumah tangga kumuh adalah rumah tangga yang tidak memiliki akses air minum (air minum layak), akses sanitasi (sanitasi layak), *sufficient living area*, dan *durability of housing*. Berbagai penelitian penyakit menular membuktikan bahwa kondisi perumahan yang berdesak-desakan dan kumuh mempunyai kemungkinan lebih besar terserang penyakit (Mukono, 2006).

Berdasarkan latar belakang dan kajian literatur di atas maka dapat dirumuskan masalahnya sebagai berikut yaitu bagaimanakah model regresi dan perbandingannya antara model regresi dengan metode *Ordinary Least Square* (OLS) dan metode *Robust MM-Estimation* dalam kasus penyakit Demam Berdarah Dengue (DBD) di Indonesia tahun 2018. Dari perumusan masalah maka tujuan penulisan yaitu untuk mendapatkan model regresi sekaligus melakukan perbandingan antara model regresi dengan metode *Ordinary Least Square* (OLS) dan metode *Robust MM-Estimation* dalam kasus penyakit Demam Berdarah Dengue (DBD) di Indonesia tahun 2018.

Sumber Data dan Variabel

Sumber data yang digunakan pada penelitian ini yaitu data pada Profil Kesehatan Indonesia Tahun 2018 yang diterbitkan oleh Kementerian Kesehatan RI. Variabel-variabel yang digunakan pada penelitian ini adalah sebagai berikut:

Tabel 2.1 Variabel Penelitian

Jenis Variabel	Nama
Dependen	Jumlah Kasus Demam Berdarah
Independen	Persentase Sanitasi Layak (X_1)
	Kepadatan Penduduk (X_2)
	Persentase Rumah Tangga Kumuh (X_3)

Spesifikasi Model

Persamaan regresi linier berganda :

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

Di mana:

Y = Variabel dependen (Demam Berdarah)

β_0 = Intercept

β_i = Koefisien dari variabel ke- i ; $i \neq 0$

X_i = Variabel independen ke- i

Atau bisa dituliskan dalam bentuk matriks:

$$Y = X\beta + \varepsilon$$

$$Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix} X = \begin{bmatrix} 1 & X_{11} & X_{21} & X_{31} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & X_{1n} & X_{2n} & X_{3n} \end{bmatrix} \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}$$

$$\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

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

Dengan rumus residual estimasi OLS:

$$\varepsilon = Y - X\beta$$

Estimasi-S dimulai dengan mencari standar deviasi:

$$\hat{\sigma}_i = \begin{cases} \frac{\text{median}|e_i - \text{median } e_i|}{0.6745} & ; \text{Iterasi} = 1 \\ \sqrt{\frac{1}{nK} \sum_{i=1}^n w_i e_i^2} & ; \text{Iterasi} > 1 \end{cases}$$

Di mana e_i merupakan residual ke- i dari estimasi OLS. Fungsi pembobot estimasi-S:

$$w_i = \begin{cases} \left[1 - \left(\frac{|u_i|}{1.547} \right)^2 \right]^2 & ; |u_i| \leq 1.547 \\ 0 & ; |u_i| > 1.547 \end{cases} ; \text{Iterasi} = 1$$

$$\frac{\rho(u)}{u^2} ; \text{Iterasi} > 1$$

Di mana $u_i = \frac{e_i}{\hat{\sigma}_i}$

Estimasi parameter β dengan Weighted Least Square dan pembobot W , dengan rumus:

$$\hat{\beta} = (X'WX)^{-1}X'WY$$

Di mana W adalah matriks diagonal berdimensi 4×4 dengan baris a dan kolom b yang berisi $W_{ab} = w_i$; $a = b = i$. Sehingga, dibentuk model persamaan:

$$Y_5 = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

Di mana koefisien estimasi parameter β -nya merupakan hasil dari estimasi-S. Kemudian, residual yang didapat dari model regresi robust estimasi-S akan dilanjutkan ke tahap estimasi-M, dengan fungsi pembobot Bisquare:

$$w_i \begin{cases} \left[1 - \left[\frac{u_i}{4,689} \right]^2 \right]^2 & ; |u_i| \leq 4,685 \\ 0 & ; |u_i| > 4,685 \end{cases}$$

Di mana $u_i = \frac{\varepsilon_i}{\hat{\sigma}_i}$

Estimasi parameter β dengan Weighted Least Square dan pembobot W , dengan rumus:

$$\hat{\beta} = (X'WX)^{-1}X'WY$$

Di mana W adalah matriks diagonal berdimensi 4×4 dengan baris a dan kolom b yang berisi $W_{ab} = w_i ; a = b = i$. Sehingga, dibentuk model persamaan:

$$Y_{MM} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

3. HASIL DAN PEMBAHASAN

Pemodelan DBD di Regresi Berganda

a. Uji Multikolinieritas

Tabel 3.1 Uji Multikolinieritas

Variabel	VIF
Persentase Sanitasi Layak (X_1)	1,910174
Kepadatan Penduduk (X_2)	1,197800
Persentase Rumah Tangga Kumuh (X_3)	1,682764

Sumber : Olah Eviews versi 10

Berdasarkan Tabel 3.1 terlihat bahwa nilai VIF < 10 , maka dapat dinyatakan bahwa tidak terdapat masalah multikolinieritas dalam model.

b. Uji Heteroskedastisitas

Tabel 3.2 Uji Heteroskedastisitas

F-statistic	0,278877	Prob. F(3,30)	0,8402
Obs*R-squared	0,922456	Prob. Chi-square(3)	0,8200
Scaled explained SS	1,263837	Prob. Chi-square(3)	0,7377

Sumber : Olah Eviews versi 10

Berdasarkan hasil uji Glejser, nilai p-value yang ditunjukkan dengan nilai Prob. Chi-Square(3) pada Obs*R-Squared yaitu sebesar 0,8200. Oleh karena p-value $0,8200 > 0,05$ maka model regresi bersifat homoskedastisitas.

c. Uji Linieritas

Tabel 3.3 Uji Linieritas

	Value	df	Probability
t-statistic	0,075404	29	0,9404
F-statistic	0,005686	(1,29)	0,9404
Likelihood ratio	0,006665	1	0,9349

Sumber : Olah Eviews versi 10

Uji Linieritas menggunakan Ramsey Reset Test, nilai p value yang ditunjukkan pada kolom probability baris F-statistic sebesar 0,9404 dimana $> 0,05$ sehingga dapat disimpulkan bahwa asumsi linieritas terpenuhi.

d. Uji Normalitas

Tabel 3.4 Uji Normalitas

	OLS
p-value	0,0000

Sumber : Olah Eviews versi 10

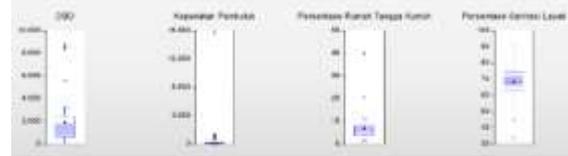
Sumber : Olah Eviews versi 10

Berdasarkan Uji Kolmogorov-Smirnov, diperoleh nilai p-value sebesar 0,0000. Karena nilai p-value = $0,0000 < 0,05$ dapat disimpulkan bahwa asumsi normalitas tidak terpenuhi.

Pendeteksian Outlier

a. Metode Boxplot

Berikut merupakan boxplot dari masing-masing variabel :



Gambar 1. Boxplot Variabel DBD, Sanitasi Layak, Kepadatan Penduduk, dan Persentase Rumah Kumuh

Berdasarkan keempat tampilan boxplot di atas, dapat terlihat bahwa di setiap variabel ada pencilan atau outlier dalam data.

b. Metode DfFits

Deteksi outlier selanjutnya adalah melihat nilai DfFITS. Batas nilai penentuan jika DfFITS $> 0,6860$ merupakan data outlier. Berdasarkan hasil pengolahan data metode DfFITS mendeteksi outlier pada data ke 11,12,15, dan 34.

Perbandingan Estimasi Parameter OLS dan Robust

Langkah selanjutnya adalah melakukan analisis regresi untuk mendapatkan nilai estimasi parameter dari data tersebut menggunakan metode OLS dan Robust MM-Estimation. Selanjutnya dapat dibandingkan metode mana yang lebih baik dengan melihat R-squared dan standar errornya.

Tabel 3.5 Parameter Model

	OLS	ROBUST
Intercept	6257,528	4752,249
Persentase Sanitasi Layak (X_1)	-53,16182	-43,00716
Kepadatan Penduduk (X_2)	0,193817	0,171930
Persentase Rumah Tangga Kumuh (X_3)	-118,5858	-82,25501

Sumber : Olah Eviews versi 10

Berdasarkan tabel 3.5 terbentuk model OLS sebagai berikut :

$$\hat{Y} = 6257,528 - 53,16182 X_1 + 0,193817 X_2 - 118,5858 X_3$$

Sedangkan model robust yang terbentuk adalah :

$$\hat{Y}_{MM} = 4752,249 - 43,00716 X_1 + 0,171930 X_2 - 82,25501 X_3$$

Tabel 3.6 Uji Signifikansi Parameter (P-Value)

	OLS	ROBUST
Intercept	0,0382	0,0003
Persentase Sanitasi Layak (X_1)	0,1696	0,0123
Kepadatan Penduduk (X_2)	0,1889	0,0087
Persentase Rumah Tangga Kumuh (X_3)	0,0862	0,0068

Sumber : Olah Eviews versi 10

Berdasarkan tabel 3.6, semua variabel di metode OLS tidak signifikan. Sedangkan di metode robust semua variabel signifikan.

Tabel 3.7 Pemilihan Model Terbaik

	OLS	ROBUST
Standard error	37,77518	17,17253
Persentase Sanitasi Layak (X_1)		

	Kepadatan Penduduk (X_2)	0,144168	0,065538
	Persentase Rumah Tangga Kumuh (X_3)	66,84169	30,38610
R-squared		0,120022	0,168821

Sumber : Olah Eviews versi 10

Pemilihan model terbaik dari model OLS dan model robust dilihat dari standard error dan R-squarednya. Jika dilihat pada tabel 3.7 dapat disimpulkan bahwa metode Robust MM-Estimation memiliki standar error yang lebih kecil dibandingkan OLS. Selain itu R-squared metode Robust MM-Estimation lebih besar dibandingkan OLS yaitu sebesar 0,1688 yang artinya proporsi keragaman variabel Y dapat dijelaskan oleh variabel X sebesar 16,88%, sedangkan sisanya dijelaskan oleh variabel lain yang tidak dibahas dalam penelitian ini. Sehingga dapat disimpulkan bahwa model yang dipilih adalah model Robust MM-Estimation.

$$\hat{Y}_{MM} = 4752,249 - 43,00716 X_1 + 0,171930 X_2 - 82,25501 X_3$$

Interpretasi Model Robust MM-Estimation :

- $\hat{\beta}_1 = -43,00716$ Jika Persentase Sanitasi Layak bertambah 1 persen, maka jumlah kasus demam berdarah berkurang sebesar 43,00716 dengan asumsi variabel kepadatan penduduk dan persentase rumah tangga kumuh sama.
- $\hat{\beta}_2 = 0,171930$ Jika kepadatan penduduk bertambah 1 satuan, maka jumlah kasus demam berdarah bertambah sebesar 0,171930 dengan asumsi variabel persentase sanitasi layak dan persentase rumah tangga kumuh sama.
- $\hat{\beta}_3 = -82,25501$ Jika Persentase Rumah Tangga Kumuh bertambah 1 persen, maka jumlah kasus demam berdarah berkurang sebesar 82,25501 dengan asumsi variabel kepadatan penduduk dan persentase sanitasi layak sama.

4. KESIMPULAN

Berdasarkan hasil dan pembahasan, maka kesimpulan dari penelitian ini adalah Pemodelan Kasus DBD dengan metode *Robust MM-Estimation* lebih baik daripada metode *Ordinary Least Square* (OLS). Hal ini ditunjukkan dengan R-Square *Robust MM-Estimation* yang lebih besar yaitu sebesar 0,168821 dan semua faktor signifikan terhadap persebaran penyakit Demam Berdarah Dangu (BDB).

Saran bagi pemerintah yaitu diharapkan mampu membuat dan memantau kebijakan yang berkaitan dengan akses sanitasi layak, perumahan kumuh serta mengurangi kepadatan penduduk sehingga penderita Demam Berdarah Dongue (DBD) bisa ditekan. Kemudian bagi masyarakat diharapkan untuk tetap waspada terhadap penyakit DBD serta membuat sanitasi layak di rumah masing – masing. Selain itu, masyarakat juga diharapkan untuk menjaga kondisi rumahnya agar

tetap bersih terutama pada daerah yang tingkat kepadatan penduduknya sangat tinggi, karena hal ini mampu menekan jumlah penderita DBD. Penelitian ini diharapkan menjadi suatu aspek pengembangan diri untuk melakukan penelitian serta dapat digunakan sebagai dasar untuk melakukan penelitian selanjutnya dengan variabel lain yang lebih mendukung.

5. REFERENSI

- Antonius, W.K. 2003. *Kebijakan Pemberantasan Wabah Penyakit Menular, Kasus Kejadian Luar Biasa Demam Berdarah Dengue (KLB DBD)*. Diakses dari <http://www.theindonesiainstitute.com>
- Anwar. 1997. *Sanitasi Makanan Dan Minuman Pada Institusi Pendidikan Tenaga Sanitasi*. Pusat Pendidikan Tenaga Sanitasi, Pusat pendidikan Tenaga Kesehatan Depkes RI. Jakarta, h. 6.
- Chen, C. 2002. *Robust Regression and Outlier Detection with the ROBUSTREG Procedure*. Paper 265-27. North Carolina: SAS Institute.
- Ghozali, Imam. 2011. *Aplikasi Analisis Multivariate dengan Program SPSS*. Semarang: Badan Penerbit Universitas Diponegoro.
- Hurint, Hisintus S., dkk. 2016. *Metode Regresi Robust Dengan Estimasi Method of Moment (Estimasi-MM) Pada Regresi Linier Berganda (Studi Kasus: Data Indeks Harga Konsumen (IHK))*. Jurnal Eksponensial Vol. 7 No. 2 Samarinda: Universitas Mulawarman.
- Kementrian Kesehatan RI. 2018. Pusat Data dan Informasi Kementrian Kesehatan RI. Jakarta: Kemenkes RI. Diakses pada tanggal 15 Februari 2020, dari <https://pusdatin.kemkes.go.id/>
- M. Salibian dan V.J. Yohai. 2006. *A fast Algorithm for S-Regression Estimates*. Journal of Computational and Graphical Statistics. **15**, No. 20, 414-427, doi: 10.1198/106186006X113629.
- Mukono, H.J. 2006. *Prinsip Dasar Kesehatan Lingkungan: Edisi Kedua*. Surabaya.
- P. J. Rousseeuw and V. J. Yohai. 1984 *Robust Regression by Mean of S-Estimators, Robust and Nonlinear Time Series Analysis*. New York. 256-274, doi: 10.1007/978-1-4615-7821-5-15.
- Sembiring, R. K. 1995. *Analisis Regresi*. Bandung: Penerbit ITB.
- Susanti, Yuliana dkk. 2014. *M Estimation, and MM Estimation in Robust Regression*. International journal of Pure and Applied Mathematics. Vol. 91, No. 3, 349-360.
- Soemartini. 2007. *Outlier (Pencilan)*. Bandung. UNPAD.

- Yohai, V. J. 1987. *High Breakdown-Point And High Efficiency ROBust estimates For Regression*. The Annuals of Statistics. Vol. 15, No. 20, 642-656.
- Yuliana dan Y. Susanti. 2008. *Estimasi M dan sifat-sifatnya pada Regresi Linear Robust*. Jurnal Math-Info. 1, No. 11, 8-16.