

Perbandingan Model Regresi Robust Estimasi M Dan Estimasi *Least Trimmed Squares (LTS)* Pada Jumlah Kasus Tuberkulosis Di Indonesia

Dina Rohmah¹⁾ Yuliana Susanti²⁾, Etik Zukhronah³⁾

Program Studi Statistika FMIPA UNS ^{1),2),3)}

Corresponding author : dinarohmah98@student.uns.ac.id¹

Abstrak. Tuberkulosis merupakan suatu penyakit menular yang disebabkan oleh bakteri *Mycobacterium tuberculosis*. Tuberkulosis sendiri menjadi salah satu penyakit yang menjadi perhatian dunia karena menjadi 10 penyebab kematian tertinggi di dunia pada tahun 2015. Indonesia adalah negara dengan jumlah kasus baru tuberkulosis terbanyak ketiga di dunia dengan angka 8% dari total kasus global. Pada data jumlah kasus tuberkulosis di Indonesia tahun 2019 ditemukan adanya pencilan, maka dari itu, diperlukan adanya metode yang kekar pada pengaruh pencilan yaitu regresi *robust*. Penelitian ini bertujuan untuk memperoleh estimator terbaik antara metode estimasi M dan *Least Trimmed Squares (LTS)* serta mengetahui faktor-faktor yang berpengaruh terhadap jumlah kasus tuberkulosis di Indonesia. Kriteria pemilihan estimasi terbaik dilihat dari nilai *adjusted R²* terbesar dan nilai *Mean Squares Error (MSE)* terkecil. Hasil penelitian menunjukkan bahwa metode estimasi *Least Trimmed Squares (LTS)* merupakan metode yang lebih baik dibandingkan metode estimasi M dan variabel persentase TPM memenuhi syarat kesehatan, jumlah HIV, persentase penduduk, dan jumlah tenaga kesehatan masyarakat berpengaruh signifikan terhadap jumlah kasus tuberkulosis di Indonesia tahun 2019.

Kata kunci: *Regresi robust, Estimasi M, Estimasi LTS, Tuberkulosis.*

PENDAHULUAN

Tuberkulosis merupakan suatu penyakit menular yang disebabkan bakteri *Mycobacterium tuberculosis* (Kementerian Kesehatan RI, 2020). Tuberkulosis sebagai salah satu penyakit yang menjadi perhatian global. Menurut data dari WHO (2019), Indonesia merupakan negara yang menempati posisi ketiga dengan jumlah kasus baru tuberkulosis terbanyak di dunia dengan angka 8 % dari total kasus global. Jumlah kasus tuberkulosis di Indonesia sudah mencapai 543.874 kasus pada tahun 2019, angka tersebut terbilang cukup banyak.

Orang yang mempunyai penyakit HIV/AIDS atau orang yang mempunyai status gizi buruk merupakan orang yang rentan terinfeksi dan terjangkit penyakit tuberkulosis (Infodatin, 2018). Penyakit tuberkulosis juga dapat dipengaruhi oleh persentase penduduk usia produktif, persentase tenaga kesehatan terdidik tuberkulosis, dan persentase tempat umum dan pengelolaan makanan (TPUM) sehat (Lestari dkk., 2014). Selain faktor tersebut, berdasarkan penelitian dari Banapon dkk. (2020) menunjukkan bahwa jumlah tenaga medis juga merupakan variabel yang berpengaruh signifikan pada kasus penderita tuberkulosis di Jawa Barat tahun 2017.

Untuk itu, dalam upaya penanggulangan permasalahan tuberkulosis di Indonesia diperlukan adanya pemodelan regresi untuk memprediksi kasus tuberkulosis serta mengetahui faktor-faktor yang berpengaruh pada kasus tuberkulosis. Analisis regresi digunakan untuk mempelajari ketergantungan antara variabel dependen dengan variabel independen (Gujarati, 2004). Salah satu metode yang digunakan untuk mengestimasi parameter regresi adalah Metode Kuadrat Terkecil (MKT). Penggunaan MKT memerlukan beberapa asumsi yang harus dipenuhi yaitu asumsi normalitas, homoskedastisitas, nonautokorelasi, dan nonmultikolinearitas. Pada kenyataannya, terkadang ditemukan asumsi normalitas tidak terpenuhi yang disebabkan adanya pencilan sehingga penggunaan MKT tidak lagi efisien. Oleh karena itu, untuk menangani adanya pencilan diperlukan metode regresi *robust*.

Regresi *robust* merupakan metode regresi untuk menganalisis data yang dipengaruhi oleh adanya pencilan, sehingga menghasilkan model yang kekar

terhadap pengaruh pencilan (Draper and Smith, 1998). Pada regresi *robust* terdapat beberapa metode estimasi untuk mengestimasi parameter regresi, salah satunya adalah estimasi M yang diperkenalkan oleh Huber (1973) dan estimasi *Least Trimmed Squares (LTS)* yang diperkenalkan oleh Rousseeuw (1984). Metode estimasi M mempunyai prinsip dasar meminimumkan fungsi residu ρ (Yuliana *et al.*, 2014). Sedangkan metode estimasi *LTS* mempunyai prinsip dasar meminimumkan jumlah kuadrat residu terpangkas (Williems and Aelst, 2004).

Orhan *et al.* (2001) pernah melakukan penelitian yang bertujuan membandingkan teknik regresi *robust* dengan *high breakdown* (estimasi *LTS*) dengan Metode Kuadrat Terkecil (MKT). Berdasarkan indikator dari nilai *standard error*, R^2 dan statistik F dapat disimpulkan bahwa metode estimasi *LTS* lebih baik dibandingkan metode MKT.

Berdasarkan uraian sebelumnya, maka ingin dilakukan penelitian dengan membandingkan model regresi *robust* estimasi M dan estimasi *Least Trimmed Squares (LTS)* serta mengetahui faktor-faktor yang berpengaruh terhadap jumlah kasus tuberkulosis di Indonesia.

METODE PENELITIAN

Regresi Linier Berganda

Model regresi linier berganda merupakan analisis untuk mengetahui hubungan yang mungkin terjadi antara variabel dependen y dengan k variabel independen, x_1, x_2, \dots, x_k . Menurut Montgomery and Peck (2006), model regresi linier berganda mempunyai persamaan umum sebagai berikut

$$y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + \varepsilon_i \quad (1)$$

dengan y_i adalah variabel dependen pada pengamatan ke- i , $\beta_0, \beta_1, \dots, \beta_k$ adalah parameter regresi, $x_{i1}, x_{i2}, \dots, x_{ik}$ adalah variabel independen, dan ε_i adalah residu ke- i yang diasumsikan $\varepsilon_i \sim NID(0, \sigma^2)$

Metode Kuadrat Terkecil (MKT)

Menurut Gujarati (2004), tujuan dari MKT ini adalah untuk meminimumkan jumlah kuadrat dari residu. Estimator parameter dari persamaan (1) dengan menggunakan MKT adalah

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

dengan

$$Y = [y_1 \ y_2 \ \dots \ y_n], X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ \vdots & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix}$$

$$\hat{\beta} = [\hat{\beta}_0 \ \hat{\beta}_1 \ \dots \ \hat{\beta}_k]$$

dengan Y adalah vektor dari variabel dependen dengan ukuran $n \times 1$, X adalah matriks dari variabel independen dengan ukuran $n \times m$, dengan $m = k + 1$, $\hat{\beta}$ adalah vektor estimator parameter model dengan ukuran $m \times 1$.

Estimasi M

Prinsip dasar dari estimasi M adalah untuk meminimumkan fungsi residu ρ (Yuliana *et.al.*, 2014).

$$\hat{\beta}_M = \rho \left(y_i - \sum_{j=0}^k x_{ij} \beta_j \right)$$

Estimasi M didefinisikan sebagai fungsi dari ε , $\rho(\varepsilon)$ dan disebut dengan fungsi *huber*.

b_0, b_1, \dots, b_p masing-masing adalah penaksir dari $\beta_0, \beta_1, \dots, \beta_p$, dipilih sedemikian rupa sehingga $\sum p(e)$ minimum dan fungsi Huber didefinisikan sebagai berikut (Birkes and Dodge, 1993)

$$\rho(e) = \begin{cases} e^2 & \text{jika } -k \leq e \leq k \\ 2k|e| - k^2 & \text{jika } e < -k \text{ atau } k < e \end{cases}$$

Berdasarkan saran dari Huber diambil nilai $k = 1,5\hat{\sigma}$, dimana $\hat{\sigma}$ adalah estimasi standar deviasi σ dari populasi kesalahan acak. Untuk mengestimasi $\hat{\sigma}$ dapat menggunakan $\hat{\sigma} = 1.483 \text{ MAD}$, dengan MAD (*Median of Absolute Deviation*) merupakan median dari nilai mutlak residu $|\hat{e}_j|$. Untuk mendapatkan nilai estimasi parameter dari estimasi M , diperlukan algoritma perhitungan, langkah-langkahnya sebagai berikut (Birkes and Dodge, 1993):

1. Misalkan b^0 adalah estimasi dari α dan β , seperti pada metode kuadrat terkecil, menghitung $e_i^0 = y_i - \hat{y}_i$ dengan $\hat{y}_i = b_0^0 + b_1^0 x_{i1} + \dots + b_p^0 x_{ip}$ serta menghitung $\hat{\sigma} = 1,483 \text{ MAD}$,
2. Memotong e_i^0 untuk mendapatkan nilai e_i^* dimana $e_i^* = 1,5\hat{\sigma}$ jika $e_i^0 >$

$1,5\hat{\sigma}$ dan $e_i^* = -1,5\hat{\sigma}$ jika $e_i^0 < 1,5\hat{\sigma}$

3. Menghitung $y_i^* = \hat{y}_i^0 + e_i^*$ dan kemudian mencari nilai b_0 menggunakan MKT,
4. Mengulangi tahap 1 sampai 3 hingga diperoleh nilai b_0 yang sama dengan iterasi sebelumnya (konvergen).

Estimasi *Least Trimmed Squares (LTS)*

Estimasi *LTS* adalah metode regresi robust dengan *breakdown point* tinggi (Chen, 2002). Rousseeuw and Leroy (1987) mendefinisikan penduga *Least Trimmed Squared (LTS)* (β) sebagai berikut

$$\hat{\beta}_{LTS} = \min \hat{\beta} \sum_{i=1}^h e_i^2$$

dengan meminimumkan jumlah kuadrat residu sebanyak h residu.

dengan:

$h = \frac{n}{2} + \frac{p+1}{2} = \frac{n+p+1}{2}$, dengan $h \leq n$, merupakan konstanta pemangkasan (*trimming*)

$e_1^2 \leq e_2^2 \leq \dots \leq e_n^2$: residu kuadrat yang diurutkan dari terkecil ke terbesar

n : banyaknya data pengamatan

p : banyaknya parameter

Algoritma estimasi *LTS* yang lebih dikenal dengan algoritma *FAST-LTS* dan *C-steps* adalah:

1. Menghitung estimasi parameter $\hat{\beta}_{awal}$ menggunakan MKT,
2. Menghitung nilai residu (e_i) dengan $e_i = y_i - \hat{y}_i$ yang bersesuaian dengan $\hat{\beta}_{awal}$, kemudian menghitung jumlah $h_0 = \frac{n+p+1}{2}$ observasi dengan nilai e_i^2 terkecil,
3. Menghitung $\sum_{i=1}^{h_0} e_i^2$
4. Menghitung estimasi parameter $\hat{\beta}_{baru}$ dari h himpunan bagian data,
5. Menentukan n kuadrat residu $e_i^2 = (y_i - \hat{y}_i)^2$ yang bersesuaian dengan $\hat{\beta}_{baru}$, kemudian menghitung sejumlah h observasi dengan nilai e_i^2 terkecil,
6. Menghitung $\sum_{i=1}^{h_{baru}} e_i^2$
7. Melakukan *C-steps* dengan mengulangi langkah 4 sampai 6 hingga

mendapatkan fungsi objektif ($\sum_{i=1}^h e_i^2$) yang konvergen.

Data Penelitian

Penelitian ini menggunakan data jumlah kasus tuberkulosis (Y), persentase Tempat Pengelolaan Makanan (TPM) memenuhi syarat (X_1), jumlah HIV (X_2), persentase penduduk (X_3), dan jumlah tenaga kesehatan masyarakat (X_4) di Indonesia tahun 2019 yang diperoleh dari publikasi Kementerian Kesehatan RI dan Badan Pusat Statistika RI.

Alur Penelitian

1. Memodelkan kasus tuberkulosis dengan menggunakan regresi MKT.
2. Pengujian asumsi klasik pada model regresi MKT kasus tuberkulosis.
3. Melakukan pendeteksian pencilan pada model regresi MKT kasus tuberkulosis.
4. Memodelkan kasus tuberkulosis menggunakan regresi *robust* estimasi M dan estimasi LTS .
5. Membandingkan hasil yang diperoleh pada regresi *robust* estimasi M dan estimasi LTS berdasarkan nilai *adjusted R²* dan nilai *Mean Squares Error (MSE)*.
6. Menentukan model regresi *robust* terbaik berdasarkan nilai *adjusted R²* yang terbesar dan nilai *Mean Squares Error (MSE)* yang terkecil.
7. Menginterpretasikan hasil yang diperoleh serta membuat kesimpulan.

HASIL DAN PEMBAHASAN

Hasil Penelitian

Langkah awal dari penelitian yaitu memodelkan data menggunakan Metode Kuadrat Terkecil (MKT). Model regresi yang diperoleh dengan menggunakan MKT pada data kasus tuberkulosis adalah

$$Y = 2649,64 - 61,2329X_1 + 0,69596X_2 + 5361,86X_3 - 1,34792X_4$$

Setelah diperoleh model regresi menggunakan MKT, selanjutnya melakukan pengujian asumsi klasik yang meliputi uji normalitas, nonmultikolinearitas, nonautokorelasi, dan homoskedastisitas. Hasil pengujian asumsi klasik terhadap model regresi menggunakan MKT menyimpulkan bahwa asumsi homoskedastisitas, nonmultikolinearitas, dan nonautokorelasi terpenuhi namun

asumsi normalitas tidak terpenuhi. Oleh karena asumsi normalitas tidak terpenuhi serta ditemukan adanya empat pencilan yaitu data ke 11, 12, 13, dan 15, dimana keempat data tersebut mempunyai nilai $|DFFITs| > 2\sqrt{\frac{p}{n}} = 0,76696$. Oleh karena itu, pemodelan regresi dilakukan menggunakan regresi *robust*.

Berdasarkan algoritma estimasi M , diperoleh model regresi menggunakan estimasi M dengan 23 iterasi adalah sebagai berikut

$$Y = 1636,24 - 49,50705X_1 + 1,13071X_2 + 4930,18X_3 - 0,13568X_4 \quad (1)$$

Dari model regresi tersebut dilakukan uji signifikansi untuk mengetahui apakah variabel persentase Tempat Pengelolaan Makanan (TPM) memenuhi syarat (X_1), jumlah HIV (X_2), persentase penduduk (X_3), dan jumlah tenaga kesehatan masyarakat (X_4) berpengaruh signifikan terhadap jumlah kasus tuberkulosis (Y). Uji signifikansi estimasi M menggunakan *reduced models*.

Hipotesis pada uji signifikansi estimasi M adalah $H_0 : \beta_j = 0, j = q + 1, q + 2, \dots, p$ (parameter regresi ke- j tidak berpengaruh signifikan dalam model) dan $H_1 : \beta_j \neq 0, j = q + 1, q + 2, \dots, p$ (salah satu parameter regresi ke- j berpengaruh signifikan dalam model). Daerah kritisnya, H_0 ditolak jika $F_M > F_{(\alpha; p-q; n-p-1)}$, dengan $F_M = \frac{STR_{reduce} - STR_{full}}{(p-q)\hat{\lambda}}$, dimana q adalah jumlah variabel independen yang tereduksi, p adalah jumlah parameter dalam model, dan n adalah jumlah data pengamatan. Tabel 1 merupakan hasil pengujian signifikansi estimasi M dengan lima belas kemungkinan *reduced models*.

<i>Reduced Models</i>	F_M dan $F_{(0,05; 4-q; 29)}$	Kesimpulan
$Y_i = \beta_0 + \varepsilon_i$	$F_M = 107,230$ $F_{(0,05;4;29)} = 2,70$	Salah satu variabel X_1, X_2, X_3, X_4 signifikan
$Y_i = \beta_0 + \beta_1 X_{i1} + \varepsilon_i$	$F_M = 174,093$ $F_{(0,05;3;29)} = 2,93$	Salah satu variabel X_2, X_3, X_4 signifikan
$Y_i = \beta_0 + \beta_2 X_{i2} + \varepsilon_i$	$F_M = 60,8532$ $F_{(0,05;3;29)} = 2,93$	Salah satu variabel X_1, X_3, X_4 signifikan
$Y_i = \beta_0 + \beta_3 X_{i3} + \varepsilon_i$	$F_M = 2,97639$ $F_{(0,05;3;29)} = 2,93$	Salah satu variabel X_1, X_2, X_4 signifikan

$Y_i = \beta_0 + \beta_4 X_{i4} + \varepsilon_i$	$F_M = 126,361$ $F_{(0,05;3;29)} = 2,93$	Salah satu variabel X_1, X_2, X_3 signifikan
$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \varepsilon_i$	$F_M = 80,2006$ $F_{(0,05;2;29)} = 3,33$	Salah satu variabel X_3, X_4 signifikan
$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_3 X_{i3} + \varepsilon_i$	$F_M = 3,63797$ $F_{(0,05;2;29)} = 3,33$	Salah satu variabel X_2, X_4 signifikan
$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_4 X_{i4} + \varepsilon_i$	$F_M = 191,477$ $F_{(0,05;2;29)} = 3,33$	Salah satu variabel X_2, X_3 signifikan
$Y_i = \beta_0 + \beta_2 X_{i2} + \beta_3 X_{i3} + \varepsilon_i$	$F_M = 1,57835$ $F_{(0,05;2;29)} = 3,33$	Salah satu variabel X_1, X_4 tidak signifikan
$Y_i = \beta_0 + \beta_2 X_{i2} + \beta_4 X_{i4} + \varepsilon_i$	$F_M = 71,7643$ $F_{(0,05;2;29)} = 3,33$	Salah satu variabel X_1, X_3 signifikan
$Y_i = \beta_0 + \beta_3 X_{i3} + \beta_4 X_{i4} + \varepsilon_i$	$F_M = 4,53589$ $F_{(0,05;2;29)} = 3,33$	Salah satu variabel X_1, X_2 signifikan
$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \varepsilon_i$	$F_M = 0,43633$ $F_{(0,05;1;29)} = 4,17$	variabel X_4 tidak signifikan
$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_4 X_{i4} + \varepsilon_i$	$F_M = 142,315$ $F_{(0,05;1;29)} = 4,17$	variabel X_3 signifikan
$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_3 X_{i3} + \beta_4 X_{i4} + \varepsilon_i$	$F_M = 6,77698$ $F_{(0,05;1;29)} = 4,17$	variabel X_2 signifikan
$Y_i = \beta_0 + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \varepsilon_i$	$F_M = 2,88253$ $F_{(0,05;1;29)} = 4,17$	variabel X_1 tidak signifikan

Tabel 1. Hasil Pengujian Signifikansi Estimasi M

Tabel 1 menunjukkan bahwa variabel jumlah HIV (X_2) dan persentase penduduk (X_3) berpengaruh signifikan terhadap jumlah kasus tuberkulosis.

Selanjutnya model regresi menggunakan estimasi LTS adalah

$$Y = 3527 - 77,75516X_1 + 1,38269X_2 + 4686,44X_3 - 0,80935X_4$$

Dari model regresi tersebut akan dilakukan uji signifikansi yang terdiri dari dua uji yaitu uji simultan dan uji parsial. Hipotesis pada uji simultan adalah $H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$ (variabel persentase TPM memenuhi syarat kesehatan, jumlah HIV, persentase penduduk, dan jumlah tenaga kesehatan masyarakat tidak berpengaruh signifikan terhadap jumlah kasus tuberkulosis) dan $H_1 : \beta_j \neq 0$ untuk $j = 0, 1, \dots, k$ (paling tidak terdapat satu variabel persentase TPM memenuhi syarat kesehatan, jumlah HIV, persentase penduduk, dan jumlah Tenaga kesehatan masyarakat berpengaruh signifikan terhadap jumlah kasus tuberkulosis). Daerah kritisnya adalah H_0 ditolak jika jika $F_{hitung} > F_{\alpha;k;n-k-1} = F_{0,05;4;29} = 2,70$. Diperoleh

$F_{hitung} = 37270,37$ sehingga H_0 ditolak ditolak berarti bahwa paling tidak terdapat satu dari persentase TPM memenuhi syarat kesehatan, jumlah kasus HIV, persentase penduduk, dan jumlah tenaga kesehatan berpengaruh signifikan pada jumlah kasus tuberkulosis. Selanjutnya dilakukan pengujian parsial dengan $H_0 : \beta_j = 0, j = 1,2,3,4$ (variabel independen ke- j tidak berpengaruh signifikan terhadap model) dan $H_1 : \beta_j \neq 0, j = 1,2,3,4$ (variabel independen ke- j berpengaruh signifikan terhadap model). Daerah kritisnya adalah H_0 ditolak jika $|t_{hitung}| > t_{\frac{\alpha}{2}; n-k-1} = t_{\frac{0,05}{2}; 34-4} = t_{0,025; 29} = 2,045$. Hasil pengujian parsial dapat dilihat pada Tabel 2.

Variabel Independen	$ t_{hitung} $	Kesimpulan
X_1	15,45	Signifikan
X_2	26,76	Signifikan
X_3	114,31	Signifikan
X_4	12,99	Signifikan

Tabel 2. Hasil Uji Parsial Model Regresi Estimasi LTS

Tabel 2 menunjukkan seluruh variabel independen berpengaruh signifikan pada jumlah kasus tuberkulosis.

Setelah mendapatkan 2 model regresi dari model regresi estimasi M dan estimasi LTS , selanjutnya membandingkan kedua model tersebut berdasarkan nilai *adjusted* R^2 dan *Mean Squares Error* (MSE). Berdasarkan hasil estimasi parameter metode estimasi M dan Estimasi *Least Trimmed Squares* (LTS), nilai *adjusted* R^2 dan *Mean Squares Error* (MSE) dari kedua metode dapat dilihat pada Tabel 3.

Estimasi	Nilai <i>adjusted</i> R^2	Nilai <i>Mean Squares Error</i> (MSE)
M	99,33%	3448713
LTS	100,00%	4000

Tabel 3. Nilai *adjusted* R^2 dan *Mean Squares Error* (MSE)

Tabel 3 menunjukkan nilai *adjusted* R^2 dari metode estimasi LTS lebih besar dibandingkan metode estimasi M dan nilai *Mean Squares Error* (MSE) dari metode estimasi LTS lebih kecil dibandingkan metode estimasi M . Maka dari itu, estimator

dengan metode estimasi *LTS* lebih baik dibandingkan metode estimasi *M* dalam mengestimasi parameter pada data jumlah kasus tuberkulosis di Indonesia tahun 2019.

Diskusi Hasil Penelitian

Berdasarkan hasil penelitian yang diperoleh, metode estimasi *Least Trimmed Squares (LTS)* merupakan metode estimasi regresi *robust* yang lebih baik dibandingkan metode estimasi *M* pada data jumlah kasus tuberkulosis di Indonesia tahun 2019. Model regresi estimasi *LTS* adalah

$$Y = 3527 - 77,75516X_1 + 1,38269X_2 + 4686,44X_3 - 0,80935X_4$$

Peningkatan satu persen persentase TPM memenuhi syarat kesehatan dan satu tenaga kesehatan masyarakat akan menurunkan masing-masing 77,75561 dan 0,80935 jumlah kasus tuberkulosis. Sedangkan peningkatan satu penderita HIV dan satu persen persentase penduduk akan meningkatkan masing-masing 1,38269 dan 4686,44 jumlah kasus tuberkulosis. Variabel persentase TPM memenuhi syarat kesehatan, jumlah HIV, persentase penduduk, dan jumlah tenaga kesehatan masyarakat berpengaruh terhadap jumlah kasus tuberkulosis di Indonesia tahun 2019.

SIMPULAN

Berdasarkan hasil dan pembahasan sebelumnya, dapat diambil kesimpulan bahwa metode estimasi *LTS* merupakan metode yang lebih baik dibandingkan metode estimasi *M*. Hal ini dilihat dari nilai *adjusted R²* dari metode estimasi *LTS* lebih besar dibandingkan metode estimasi *M* dan nilai *Mean Squares Error (MSE)* dari metode estimasi *LTS* lebih kecil dibandingkan metode estimasi *M*. Model regresi yang lebih baik adalah model regresi menggunakan metode estimasi *LTS*, yaitu

$$Y = 3527 - 77,75516X_1 + 1,38269X_2 + 4686,44X_3 - 0,80935X_4$$

DAFTAR PUSTAKA

- Banapon, A., Putra, M. L. P., dan Widodo, E. (2020). "Penerapan Regresi Binomial Negatif untuk Mengatasi Pelanggaran Overdispersi pada Regresi Poisson (Studi Kasus Penderita Tuberkulosis di Provinsi Jawa Barat Tahun 2017)". *Biostatistics: Jurnal Statistika Teori dan Aplikasi: Biomedics, Industry & Business And Social*. 39-52.
- Birkes, D. and Dodge, Y. (1993). *Alternative Methods of Regression*. New York: John Wiley & Sons, Inc.
- BPS. (2020). *Statistik Indonesia 2020*. Jakarta: Badan Pusat Statistik (BPS) RI.
- Chen, C. (2002). "Robust Regression and Outlier Detection with the ROBUSTREG Procedure". *SUGI Proceedings*. 265-27.
- Draper, N. R., and Smith, H. (1998). *Applied Regression Analysis* (Third ed.). Canada: John Wiley & Sons, Inc.
- Gujarati, D. N. (2004). *Basic Econometrics* (Fourth ed.). United State (USA): The McGraw-Hill.
- Kementerian Kesehatan RI. (2020). *Profil Kesehatan Indonesia 2019*. Jakarta: Kementerian Kesehatan RI.
- Lestari, R. D., Wulandari, S. P., dan Purhadi. (2014). "Pemodelan Faktor-Faktor yang Mempengaruhi Jumlah Kasus Penyakit Tuberkulosis di Jawa Timur dengan Pendekatan Generalized Poisson Regression dan Geographically Weighted Poisson Regression". *Jurnal Sains dan Seni POMITS*. 3, (2).
- Pusat Data Informasi Kementerian Kesehatan RI. (2018). *InfoDatin: Tuberkulosis*. Jakarta: Kementerian Kesehatan RI.
- Rousseeuw, P. J., and Leroy, A. M. (1987). *Robust Regression And Outlier Detection*. Canada: John Wiley & Sons, Inc.
- Susanti, Y., Pratiwi, H., Handajani, S. S., and Liana, T. (2014). "M Estimation, S Estimation, and MM Estimation In Robust Regression". *International Journal of Pure and Applied Mathematics*. 349-360.
- WHO. (2019). *Global Tuberculosis Report 2019*. New York : World Health Organization.
- Williems, G., and Aelst, S. V. (2004). "Fast and Robust bootstrap for LTS". In *Computational Statistics dan Data Analysis*.
- Zaman, A., Rousseeuw, P. J., and Orhan, M. (2001). "Econometric applications of high-breakdown robust regression techniques". *Economic Letters*. 1-8