

ESTIMASI PARAMETER METODE *WEIGHTED LEAST SQUARE* DALAM MENGATASI MASALAH HETEROSKEDASTISITAS

Hidayatun Nisa, Dadan Kusnandar, Shantika Martha

INTISARI

Regresi linear berganda merupakan salah satu metode statistika yang digunakan untuk menganalisis pengaruh antara dua atau lebih variabel independen terhadap variabel dependen. Salah satu asumsi yang harus terpenuhi pada analisis regresi linear berganda adalah tidak terjadinya heteroskedastisitas atau varians error harus tetap (konstan) di dalam model regresi. Metode Weighted Least Square (WLS) merupakan bentuk dari pengembangan penduga least square yang digunakan untuk mengatasi masalah heteroskedastisitas. Tujuan dari penelitian ini adalah menduga parameter metode WLS untuk mengatasi masalah heteroskedastisitas. Data yang digunakan dalam penelitian ini adalah data sosial ekonomi dari 33 provinsi di Indonesia. Data yang dianalisis adalah data tingkat pengangguran terbuka (Y), pertumbuhan ekonomi (X_1), jumlah penduduk (X_2), tingkat partisipasi angkatan kerja (X_3), dan kebutuhan hidup minimum (X_4). Hasil analisis menunjukkan bahwa metode WLS dapat mengatasi masalah heteroskedastisitas. Model regresi yang diperoleh menggunakan metode WLS hanya melibatkan tiga variabel yaitu: $\hat{Y} = 30,616 + 5,640 \times 10^{-5} X_2 - 0,439 X_3 + 2,365 \times 10^{-6} X_4$ dengan nilai adjusted R^2 sebesar 95,49% yang berarti bahwa besarnya pengaruh variabel jumlah penduduk, tingkat partisipasi angkatan kerja dan kebutuhan hidup minimum terhadap tingkat pengangguran terbuka sebesar 95,49%.

Kata Kunci: *regresi linear berganda, transformasi, weighted least square*

PENDAHULUAN

Analisis regresi merupakan suatu metode analisis statistika yang digunakan untuk menganalisis pengaruh antara variabel independen terhadap variabel dependen, dimana variabel independen adalah variabel yang mempengaruhi dan variabel dependen adalah variabel yang dipengaruhi. Analisis regresi menghasilkan suatu model yang menghubungkan variabel dependen dengan variabel independen. Terdapat beberapa metode yang dapat digunakan untuk menduga parameter, salah satunya yaitu metode kuadrat terkecil (*ordinary least square*). Prinsip metode kuadrat terkecil adalah meminimumkan jumlah kuadrat *error*. Metode kuadrat terkecil menghendaki beberapa asumsi klasik yang harus terpenuhi yakni *error* berdistribusi normal, homoskedastisitas, tidak terjadi multikolinearitas dan tidak terjadi autokorelasi [1].

Salah satu metode yang dapat digunakan untuk mengatasi masalah heteroskedastisitas adalah metode *Weighted Least Square* (WLS). Metode WLS digunakan jika efisiensi estimator dianggap lebih penting daripada sifat *unbiased* dan konsisten jika dalam kondisi heteroskedastisitas. Metode WLS sama halnya seperti metode OLS yang meminimumkan jumlah kuadrat *error*, hanya saja pada metode WLS dilakukan pembobotan yang tepat kemudian baru menggunakan metode OLS terhadap data yang telah diboboti.

Data yang digunakan merupakan data sekunder yang diperoleh dari BPS Republik Indonesia tahun 2014 meliputi pertumbuhan ekonomi (X_1), jumlah penduduk (X_2), tingkat partisipasi angkatan kerja (X_3), kebutuhan hidup minimum (X_4) dan variabel dependen adalah tingkat pengangguran terbuka (Y). Setelah data diperoleh tahapan selanjutnya melakukan analisis regresi linear berganda untuk menentukan model regresi dengan menggunakan metode kuadrat terkecil. Selanjutnya melakukan standarisasi data karena variabel-variabel yang diteliti terdapat perbedaan ukuran satuan pada data asli. Kemudian mendeteksi adanya heteroskedastisitas dengan melihat nilai Breusch Pagan (BP). Apabila

ditemukan adanya heteroskedastisitas pada data, maka selanjutnya menentukan pembobot yang sesuai untuk mengatasi masalah heteroskedastisitas pada data.

Selanjutnya menduga parameter menggunakan metode WLS terhadap data yang telah ditransformasi. Kemudian kembalikan bentuk baku data ke bentuk awal. Kemudian dilanjutkan dengan uji terhadap model secara simultan dengan uji F dan uji individu untuk koefisien regresi dengan uji t . Selanjutnya memastikan kembali bahwa sudah tidak terjadi heteroskedastisitas dengan melihat nilai BP. Kemudian langkah terakhir dilanjutkan dengan melihat nilai *adjusted R*².

ANALISIS REGRESI LINEAR BERGANDA

Analisis regresi adalah salah satu metode yang biasa digunakan untuk menganalisis hubungan antar dua variabel atau lebih. Hubungan tersebut dapat dituliskan kedalam bentuk persamaan yang menghubungkan variabel dependen (Y) dan variabel independen (X). Model umum dari regresi linier berganda k variabel independen dapat dinyatakan sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \varepsilon_i \quad (1)$$

Penduga parameter $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ dapat diperoleh menggunakan metode kuadrat terkecil yang merupakan salah satu penduga parameter dalam model regresi. Tujuan dari metode kuadrat terkecil yaitu meminimumkan jumlah kuadrat (*error*) pada model yang terbentuk. Penduga parameter metode kuadrat terkecil menggunakan persamaan berikut [2]:

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (2)$$

HETEROSKEDASTISITAS

Heteroskedastisitas berarti variansi antar *error* yang satu dengan *error* yang lain berbeda ($\text{var}(\varepsilon_i) = \sigma_i^2$). Pada kasus heteroskedastisitas, metode kuadrat terkecil akan menjadi tidak efisien dalam menduga koefisien regresi, karena menghasilkan varians yang besar[3]. Jika diasumsikan $E(\varepsilon_i) = 0$ dan $E(\varepsilon_i, \varepsilon_j) = 0$ untuk $i \neq j$, maka bentuk matriks varians kovarians vektor ε pada kondisi heteroskedastisitas adalah sebagai berikut [4]:

$$\text{var}(\varepsilon) = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{bmatrix}$$

WEIGHTED LEAST SQUARE

Metode *Weighted Least Square* (WLS) merupakan bentuk pengembangan dari OLS yang digunakan untuk mengatasi masalah heteroskedastisitas. Heteroskedastisitas dalam model regresi terjadi karena ketidaksamaan varians dari vektor *error* yaitu $\text{var}(\varepsilon_i)$ tidak sama untuk setiap i , dinotasikan $\text{var}(\varepsilon_i) = \sigma_i^2$. WLS merupakan kasus khusus dari *Generalized Least Squares*. Metode WLS memiliki kemampuan untuk mempertahankan sifat efisiensi estimatornya tanpa harus kehilangan sifat bias dan konsistensinya [5].

Metode WLS diperoleh dengan meminimumkan:

$$\sum \varepsilon_i^2 = \sum w_i (Y_i - \beta_0^* - \beta_1^* X_{1i} - \beta_2^* X_{2i} - \dots - \beta_k^* X_{ki})^2 \quad (3)$$

dimana w_i sebagai pembobot, $\beta_0^*, \beta_1^*, \beta_2^*, \dots, \beta_k^*$ adalah penduga kuadrat terkecil tertimbang.

Metode WLS pada prinsipnya sama dengan metode kuadrat terkecil, bedanya pada metode WLS terdapat penambahan variabel baru yaitu w sebagai pembobot. Penduga parameter $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ untuk regresi linear berganda dengan metode WLS sebagai berikut:

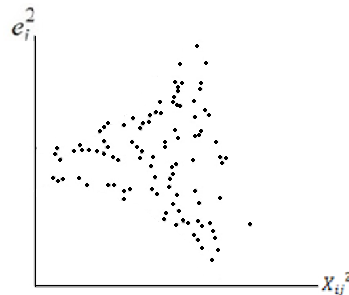
$$\beta = (X'WX)^{-1}X'WY \quad (4)$$

Dalam kenyataannya σ_i^2 tidak diketahui sehingga w juga tidak diketahui. Oleh karena itu, untuk menentukan pembobot dapat dilihat dari *error* terhadap variabel independen, antara lain:

- a. Varians *error* proporsional terhadap X_i^2

$$E(\varepsilon_i^2) = \sigma^2 X_i^2$$

Pendeteksian heteroskedastisitas dapat dilihat menggunakan metode grafik bahwa varians *error* proporsional terhadap nilai kuadrat dari variabel X_i dapat dilihat pada Gambar 1 sebagai berikut:



Gambar 1. Varians *error* proporsional terhadap kuadrat X_i

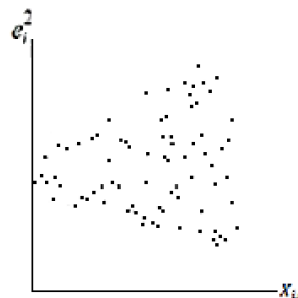
Jika pola menunjukkan hubungan kuadrat seperti pada Gambar 1 maka dapat diasumsikan varians dari ε_i proporsional terhadap kuadrat X_i sehingga pembobot yang digunakan dalam metode WLS adalah $\frac{1}{X_i}$, persamaan regresinya adalah:

$$\frac{Y_i}{X_i} = \frac{(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \varepsilon_i)}{X_i}$$

- b. Varians *error* proporsional terhadap X_i

$$E(\varepsilon_i^2) = \sigma^2 X_i$$

Pendeteksian heteroskedastisitas dapat dilihat menggunakan metode grafik bahwa varians *error* proporsional terhadap variabel X_i itu sendiri dapat dilihat pada Gambar 2 sebagai berikut:



Gambar 2. Varians *error* proporsional terhadap X_i

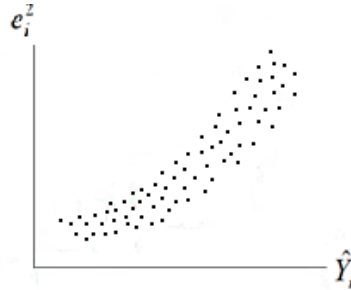
Jika pola menunjukkan hubungan linier seperti pada Gambar 2 maka dapat diasumsikan varians dari ε_i proporsional terhadap X_i itu sendiri, sehingga pembobot yang digunakan dalam metode WLS adalah $\frac{1}{\sqrt{X_i}}$, persamaan regresinya adalah

$$\frac{Y_i}{\sqrt{X_i}} = \frac{(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \varepsilon_i)}{\sqrt{X_i}}$$

- c. Varians *error* proporsional terhadap nilai kuadrat rata-rata:

$$E(\varepsilon_i^2) = \sigma^2[E(Y_i)]^2$$

Pendeteksian heteroskedastisitas dapat dilihat menggunakan metode grafik bahwa varians *error* proporsional terhadap $[E(Y_i)]^2$ dapat dilihat pada Gambar 3 sebagai berikut:



Gambar 3. Varians *error* proporsional terhadap $[E(Y_i)]^2$

Jika varians *error* proporsional terhadap $[E(Y_i)]^2$ maka langkah yang akan dilakukan yaitu melakukan regresi metode kuadrat terkecil biasa tanpa memperhatikan heteroskedastisitas untuk mendapatkan \hat{Y} , sehingga persamaan regresinya adalah

$$\frac{Y_i}{\hat{Y}_i} = \frac{\beta_0}{\hat{Y}_i} + \beta_1 \frac{\beta_{1i}}{\hat{Y}_i} + \beta_2 \frac{\beta_{2i}}{\hat{Y}_i} + \dots + \beta_k \frac{\beta_{ki}}{\hat{Y}_i} + \frac{\varepsilon_i}{\hat{Y}_i}$$

STUDI KASUS

Data pada penelitian ini merupakan data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) Republik Indonesia Tahun 2014. Penelitian ini terdiri dari 33 provinsi di Indonesia dengan variabel dependen yaitu tingkat pengangguran terbuka (TPT) dan variabel independen yaitu pertumbuhan ekonomi, jumlah penduduk, tingkat partisipasi angkatan kerja (TPAK), dan kebutuhan hidup minimum (KHM). Data yang diperoleh dilakukan analisis statistik deskriptif setiap variabel seperti yang disajikan pada Tabel 1 sebagai berikut:

Tabel 1 Statistik Deskriptif

	N	Minimum	Maximum	Mean	Std. Deviation
Pertumbuhan Ekonomi (%)	33	-0,58	6,79	3,73	1,69
Jumlah Penduduk (ribu jiwa)	33	849,80	46.029,6	7.621,41	10.783,88
TPAK (%)	33	59,99	78,67	66,83	3,86
KHM (ribu rupiah)	33	825,00	2.399,86	1.579,783	378,42
TPT (%)	33	1,90	10,51	5,40	2,10

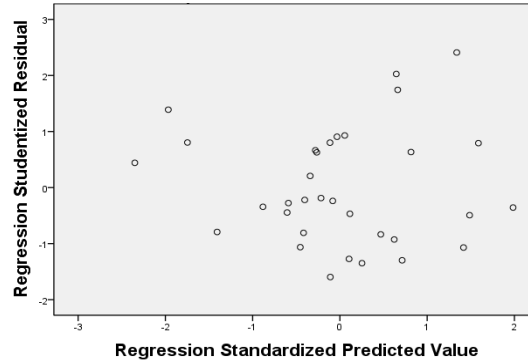
Selanjutnya melakukan standarisasi data karena data yang digunakan pada penelitian ini memiliki ukuran satuan yang berbeda. Kemudian melakukan analisis linear berganda dengan menggunakan metode kuadrat terkecil. Model regresi linear berganda dengan metode kuadrat terkecil persamaan regresinya sebagai berikut:

$$\hat{Y} = -2,860 \times 10^{-15} - 0,236X_1 + 0,391X_2 - 0,640X_3 + 0,306X_4$$

Tahap berikutnya mendeteksi adanya heteroskedastisitas pada data. Untuk mengetahui apakah terdapat permasalahan heteroskedastisitas pada data tingkat pengangguran terbuka dan faktor-faktor yang mempengaruhinya dilakukan uji heteroskedastisitas dengan menggunakan Uji Breusch Pagan (BP). Kriteria pengambilan keputusan uji signifikansi menyatakan tolak H_0 apabila nilai BP < 0,05. Hasil analisis diperoleh nilai BP sebesar $0,0334 < 0,05$ sehingga keputusan yang dapat diambil yaitu tolak H_0 . Dapat disimpulkan bahwa variabel tersebut teridentifikasi masalah heteroskedastisitas.

Berdasarkan hasil uji BP, diketahui bahwa asumsi homoskedastisitas belum terpenuhi, sehingga selanjutnya melakukan estimasi dengan regresi WLS.

Setelah diketahui data mengandung masalah heteroskedastisitas, selanjutnya menentukan pembobot yang tepat untuk digunakan dalam mengatasi masalah heteroskedastisitas. Pendeteksian untuk pembobot yang tepat dapat dilihat pada Gambar 4 sebagai berikut:



Gambar 4. Diagnostik plot

Pada kasus ini, pembobot yang tepat untuk digunakan dalam mengatasi heteroskedastisitas adalah $\frac{1}{X_3}$. Setelah mengestimasi parameter dengan menggunakan metode WLS dengan diberi pembobot yang tepat, langkah selanjutnya menduga model regresi dengan pembobot yang sudah ditentukan dengan metode kuadrat terkecil, sehingga diperoleh hasil sebagai berikut:

Tabel 2 Penduga parameter setelah diberi pembobot

Variabel	Estimasi Parameter	Standar error	Sig
β_0^*	-1,198	0,525	0,030
β_1^*	0,018	0,092	0,839
β_2^*	0,520	0,226	0,047
β_3^*	-0,822	0,109	0,000
β_4^*	0,466	0,128	0,001

Tahap selanjutnya melakukan uji signifikansi regresi pada persamaan yaitu uji signifikansi regresi secara simultan dan secara parsial. Kriteria pengambilan keputusan uji signifikansi menyatakan tolak H_0 apabila $F_{hitung} > F_{(\alpha, k-1, n-k)}$. Berdasarkan hasil analisis diperoleh nilai F_{hitung} sebesar 227,317 dengan taraf signifikansi sebesar 0,05 diperoleh nilai $F_{(0,05;4;28)}$ sebesar 2,71 sehingga keputusan yang dapat diambil yaitu tolak H_0 . Dapat disimpulkan bahwa terdapat hubungan linear antara variabel-variabel independen terhadap variabel dependen.

Kemudian uji signifikansi secara parsial diperoleh bahwa variabel independen yang berpengaruh secara signifikan terhadap tingkat pengangguran terbuka karena nilai $|t_{hitung}| \geq t_{(0,025;28)}$ adalah variabel jumlah penduduk, TPAK dan kebutuhan hidup minimum, sedangkan variabel pertumbuhan ekonomi tidak berpengaruh signifikan terhadap variabel tingkat pengangguran terbuka karena nilai $|t_{hitung}| \leq t_{(0,025;28)}$.

Setelah diberi pembobot pada data kemudian diuji kembali tanpa memasukkan variabel kebutuhan hidup minimum apakah data tersebut masih terdapat masalah heteroskedastisitas. Kriteria pengambilan keputusan uji signifikansi menyatakan tolak H_0 apabila nilai BP $< 0,05$. Diperoleh nilai BP sebesar $0,2771 > 0,05$, sehingga keputusan yang dapat diambil yaitu terima H_0 . Dapat disimpulkan bahwa asumsi homoskedastisitas telah terpenuhi. Berdasarkan hasil tersebut dapat disimpulkan bahwa metode WLS dapat mengatasi masalah heteroskedastisitas dengan faktor-faktor pembobot tertentu.

Selanjutnya dilakukan uji koefisien determinasi untuk mengetahui proporsi pengaruh variabel independen terhadap variabel dependen, diperoleh nilai *adjusted R*² sebesar 95,49% yang berarti

bahwa besarnya pengaruh variabel jumlah penduduk, TPAK dan kebutuhan hidup minimum terhadap tingkat pengangguran terbuka sebesar 95,49%. Selanjutnya dibentuk model regresi metode WLS tanpa menyertakan pertumbuhan ekonomi dikarenakan pada uji t dinyatakan bahwa pertumbuhan ekonomi tidak berpengaruh signifikan terhadap tingkat pengangguran terbuka. Koefisien penduga yang diperoleh pada WLS masih dalam bentuk baku, maka dikembalikan ke dalam bentuk awal sebagai berikut:

$$\beta_2 = \left(\frac{S_Y}{S_{x_2}} \right) \beta_2^* = 5,640 \times 10^{-5} \qquad \beta_3 = \left(\frac{S_Y}{S_{x_3}} \right) \beta_3^* = -0,439$$

$$\beta_4 = \left(\frac{S_Y}{S_{x_4}} \right) \beta_4^* = 2,365 \times 10^{-6} \qquad \beta_0 = \bar{Y} - \sum_{j=1}^p \beta_j \bar{X}_j = 30,616$$

KESIMPULAN

Hasil analisis faktor-faktor yang mempengaruhi tingkat pengangguran terbuka di Indonesia terdapat perbedaan yang signifikan pada variabel pertumbuhan ekonomi (X_1), jumlah penduduk (X_2), tingkat partisipasi angkatan kerja (X_3), kebutuhan hidup minimum (X_4). Berdasarkan dari ke empat variabel hanya tiga variabel yang berpengaruh signifikan, diperoleh persamaan regresi sebagai berikut:

$$\hat{Y} = 30,616 + 5,640 \times 10^{-5} X_2 - 0,439 X_3 + 2,365 \times 10^{-6} X_4$$

Model yang diperoleh menunjukkan variabel jumlah penduduk (X_2), tingkat partisipasi angkatan kerja (X_3) dan kebutuhan hidup minimum (X_4) berpengaruh secara signifikan terhadap tingkat pengangguran terbuka dengan nilai *adjusted R*² sebesar 95,49% yang berarti bahwa besarnya pengaruh variabel jumlah penduduk, TPAK dan kebutuhan hidup minimum terhadap tingkat pengangguran terbuka sebesar 95,49%.

DAFTAR PUSTAKA

- [1]. Maziyya, P.A., Sukarsa, I.K.G., Asih, N.M. Mengatasi Heteroskedastisitas pada Regresi dengan Menggunakan *Weighted Least Square* (WLS). *E-Jurnal Matematika*. Vol. 4. No. 1: 20-25; 2015.
- [2]. Widyaningsih, A., Susilawati, M., dan Sumajaya, I.W. Estimasi Model Seemingly Unrelated Regression (SUR) dengan Metode Generalized Least Square (GLS). *E-Jurnal Matematika*, Vol. 4 No. 2: 102-110; 2014.
- [3]. Sugiarti, H. dan Megawarni, A., Tingkat Efisiensi Metode Regresi Robust dalam Menaksir Koefisien Garis Regresi Jika Ragam Galat Tidak Homogen, *J. Matematika, Sains, dan Teknologi.*, Vol. 6:17-24; 2005.
- [4]. Rencher, A.C. *Linear Models in Statistics*, John Wiley & Sons, Hoboken, New Jersey; 2008.
- [5]. Gujarati, D.N. *Dasar-Dasar Ekonometrika Buku 1*, Ed ke-5, Salemba Empat, Jakarta; 2010.

HIDAYATUN NISA : Jurusan Matematika FMIPA UNTAN, Pontianak
hidayatunisa06@gmail.com

DADAN KUSNANDAR : Jurusan Matematika FMIPA UNTAN, Pontianak
dkusnand@untan.ac.id

SHANTIKA MARTHA : Jurusan Matematika FMIPA UNTAN, Pontianak
shantika.martha@math.untan.ac.id