

PENERAPAN METODE *WEIGHTED LEAST SQUARE* PADA ANALISIS REGRESI BERGANDA (STUDI KASUS PADA BALITA GIZI BURUK DI PROVINSI SULAWESI SELATAN)

Wahidah Sanusi¹, Ahmad Zaki¹ dan Faisah^{1, a)}

¹Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Makassar, 90224

^{a)}e-mail: leefaisah@gmail.com

Abstrak. Analisis regresi merupakan kajian terhadap hubungan variabel terikat dengan variabel bebas. Salah satu asumsi yang harus terpenuhi dalam regresi agar mendapatkan penaksir koefisien regresi yang BLUE yaitu homoskedastisitas. Variabel-variabel penduganya yaitu variabel pemberian ASI eksklusif, rasio Puskesmas, rasio dokter, rasio bidan, rumah tangga ber-PHBS, balita ditimbang, dan pemberian vitamin A. Berdasarkan hasil yang diperoleh data melanggar asumsi karena terjadi heteroskedastisitas sehingga dilakukan metode OLS dan WLS memberikan hasil bahwa metode WLS efektif digunakan dalam mengatasi masalah heteroskedastisitas karena melalui prosedur pembobotan pada error OLS (ϵ) yang menghasilkan varian error WLS (f) menjadi homoskedastisitas. Hasil analisis regresi menunjukkan bahwa secara simultan variabel bebas berpengaruh signifikan terhadap jumlah balita gizi buruk di Provinsi Sulawesi Selatan.

Kata Kunci : *Weighted Least Square, heteroskedastisitas*

Abstract. Regression analysis is a study of the dependent variable relationship with independent variables. One of the assumptions that must be fulfilled in the regression to get the BLUE regression coefficient estimator is homoscedasticity. The variables are the exclusive breastfeeding variable, the ratio of Puskesmas, doctor's ratio, midwife ratio, PHBS household, toddlers weighed, and vitamin A administration. Based on the results obtained the data violate the assumption due to heteroscedasticity so that the OLS and WLS the result that the WLS method is effectively used in overcoming the problem of heteroscedasticity because it is through the weighting procedure in OLS error (ϵ) that results in a WLS error variant (f) being homoscedasticity. The result of regression analysis shows that simultaneously independent variable has significant influence to malnutrition balita number in South Sulawesi Province.

Keywords: *Weighted Least Square, heteroscedasticity*

PENDAHULUAN

Analisis regresi merupakan suatu metode yang digunakan untuk menganalisis hubungan antar variabel. Hubungan tersebut dapat dinyatakan dalam bentuk persamaan yang menghubungkan variabel dependen dengan satu atau lebih variabel independen [1]. Salah satu cara untuk mengukur pengaruh satu atau lebih variabel independen (variabel bebas) terhadap variabel dependen (variabel terikat) adalah dengan menggunakan analisis regresi berganda. Analisis regresi linear berganda adalah hubungan secara linear antara dua atau lebih variabel independen ($X_{1,2} \dots X_n$) dengan variabel dependen (Y). Analisis ini merupakan alat statistik yang memanfaatkan hubungan antara dua atau lebih peubah kuantitatif sehingga salah satu peubah bias diramalkan dari peubah lainnya. Analisis ini juga digunakan untuk mengetahui arah hubungan antara variabel independen dengan variabel dependen apakah masing-masing variabel independen berhubungan positif atau negatif dan untuk memprediksi nilai variabel apabila nilai variabel independen mengalami kenaikan atau penurunan [2].

Dalam analisis regresi linear berganda diperlukan suatu metode untuk menduga parameter agar memenuhi sifat BLUE (*Best Linear Unbiased Estimator*). Metode yang paling sering digunakan adalah *Ordinary Least Square* (OLS) atau sering disebut dengan Metode Kuadrat Terkecil (MKT). Terdapat beberapa asumsi yang harus dipenuhi dalam melakukan estimasi dengan metode kuadrat

terkecil. Asumsi tersebut antara lain: data harus mengikuti sebaran normal, homoskedastisitas, tidak ada multikolinearitas dan tidak ada autokorelasi. Metode kuadrat terkecil akan dapat memenuhi sifat BLUE, jika memenuhi semua asumsi tersebut. Namun jika terdapat salah satu atau lebih asumsi yang tidak terpenuhi, maka hasil estimasi yang diperoleh tidak dapat memenuhi sifat BLUE. Salah satu asumsi yang harus dipenuhi dalam melakukan estimasi adalah homoskedastisitas (*homoskedasticity*). Homoskedastisitas berarti varian *error* adalah konstan. Asumsi ini menyatakan peubah respon memiliki varian yang sama sepanjang nilai peubah bebas. Namun jika varian *error* menunjukkan adanya variasi (variantaksama) maka kondisi ini disebut heteroskedastisitas. Heteroskedastisitas adalah bentuk pelanggaran terhadap asumsi homoskedastisitas. Jika pada saat melakukan estimasi dengan metode kuadrat terkecil dan kemudian terjadi heteroskedastisitas, maka hasil estimasi yang diperoleh tidak lagi memenuhi sifat BLUE sehingga diperlukan metode alternatif lain dalam melakukan estimasi parameter yang dapat mengatasi adanya heteroskedastisitas [3].

Metode *Weighted Least Square* merupakan salah satu metode yang dapat menyelesaikan masalah heteroskedastisitas. Metode estimasi WLS digunakan jika efisiensi estimator dianggap lebih penting dari pada sifat unbiased dan konsisten jika dalam kondisi heteroskedastisitas. WLS memiliki kemampuan untuk mempertahankan sifat efisiensi estimator tanpa harus kehilangan sifat tak bias dan konsistensinya. Metode WLS sama halnya seperti metode OLS dengan meminimumkan jumlah sisaan hanya saja pada metode WLS dilakukan pembobotan suatu faktor yang tepat kemudian baru menggunakan metode OLS terhadap data yang telah diboboti. Kelebihan dari metode ini adalah bias mengatur pentingnya setiap observasi dalam menentukan solusi akhir karena pada OLS diasumsikan bahwa nilai duga parameter regresi bernilai sama untuk setiap observasi [4].

Analisis Regresi

Dalam kasus parametrik, peneliti biasanya menggunakan Metode Kuadrat Terkecil untuk menduga parameter-parameter regresi dengan sampel yang teramati dan melandaskan kesimpulan-kesimpulan yang menyangkut parameter-parameter populasi pada asumsi-asumsi yang harus dipenuhi. Salah satu asumsi yang harus dipenuhi adalah kenormalan galat, yaitu bahwa galat berdistribusi normal dengan rata-rata nol dan simpangan baku tertentu. Analisis regresi terbagi menjadi 2, yakni analisis regresi sederhana dan analisis regresi berganda. Analisis regresi sederhana adalah suatu metode untuk mengetahui hubungan variabel dependen dengan satu variabel independen dalam bentuk persamaan model regresi. Sedangkan analisis regresi berganda adalah suatu metode untuk mengetahui hubungan variabel dependen dengan lebih dari satu variabel independen dalam bentuk persamaan model regresi [5].

Analisis Regresi Berganda

Secara umum model regresi berganda adalah sebagai berikut [6].

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_{p-1} X_{i,p-2} + \varepsilon_i \quad (1)$$

Dengan :

- | | | |
|---|---|---|
| $\beta_0, \beta_1, \beta_2, \dots, \beta_{p-1}$ | : | Parameter |
| $X_{i1}, X_{i2}, \dots, X_{i,p-2}$ | : | variabel bebas |
| ε_i | : | Sisa (error) untuk pengamatan ke-i yang diasumsikan berdistribusi normal yang saling identik dengan rata-rata nol (0) dan variansi σ^2 |

Asumsi Regresi

1. Uji Normalitas dengan tujuan untuk dapat mengetahui bahwa data yang ada terdistribusi normal dan independen
2. Uji Multikolinearitas merupakan keadaan dimana ada hubungan linier secara sempurna atau mendekati sempurna antara variabel independen dalam model regresi. [7]

3. Uji Autokorelasi adalah hubungan yang terjadi diantara residual dari pengamatan satu dengan pengamatan yang lain.[7]
4. Uji Heteroskedastisitas adalah kondisi dimana *error term* tidak memiliki suatu varian yang konstan untuk semua observasi.[8]

Metode Pendugaan Parameter

1. Metode Ordinary Least Square

Metode *ordinary least square* adalah prosedur penarikan garis regresi yang memilih suatu garis regresi dan membuat jumlah kuadrat jarak vertikal dari titik-titik yang dilalui garis lurus tersebut sekecil mungkin. Metode *ordinary least square* adalah metode yang digunakan untuk menaksir β pada persamaan regresi.[5]

Prinsip dasar metode ordinary least square adalah meminimumkan jumlah kuadrat error yaitu meminimumkan $\sum_{i=1}^n \varepsilon_i^2$ [9]

Jika memenuhi asumsi – asumsi regresi linier berganda, maka penaksir kuadrat terkecil mempunyai sifat linear, tak bias, dan variansi minimum atau biasa dikenal dengan sifat BLUE (*best, linier, unbiased, estimator*). Sifat BLUE adalah asumsi yang dikembangkan oleh Gauss dan Markov, yang kemudian teori tersebut terkenal dengan sebutan Gauss-Markov Theorem, dimana:

- a. Hasil regresi dikatakan *Best* apabila garis regresi yang dihasilkan guna melakukan estimasi atau peramalan dari sebaran data, menghasilkan *error* yang terkecil.
- b. Linear dalam model artinya model yang digunakan dalam analisis regresi telah sesuai dengan kaidah model OLS dimana variabel-variabel penduganya hanya berpangkat satu. Sedangkan linear dalam parameter menjelaskan bahwa parameter yang dihasilkan merupakan fungsi linear dari sampel.
- c. *Unbiased* atau tidak bias. Suatu estimator dikatakan *unbiased* jika nilai harapan dari estimator *b* sama dengan nilai yang benar dari *b*. Artinya, nilai rata-rata $b=b$. Bila rata-rata $b \neq b$, maka selisihnya itu disebut dengan bias.

2. Weighted Least Square

- a. WLS untuk Regresi Linear

Meminimumkan SSE pada regresi linear sederhana, yaitu :

$$SSE = \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_i)^2 \quad (2)$$

Sedangkan untuk WLS, masing-masing jumlah kuadrat error akan dikali dengan penimbang atau pembobot yaitu w_i , sehingga:

$$SSE_w = Q_w = \sum_{i=1}^n w_i (Y_i - \beta_0 - \beta_1 X_i)^2 \quad (3)$$

Sehingga diperoleh nilai dari koefisien regresinya dengan meminimumkan nilai Q_w .

Meminimumkan kuadrat error terhadap β_0 dan β_1 :

$$\frac{\partial Q_w}{\partial \beta_0} = 0 \rightarrow -2 \sum_{i=1}^n w_i y_i + 2\beta_0 \sum_{i=1}^n w_i + 2\beta_1 \sum_{i=1}^n w_i x_i = 0$$

dan

$$\frac{\partial Q_w}{\partial \beta_1} = 0 \rightarrow -2 \sum_{i=1}^n w_i x_i y_i + 2\beta_0 \sum_{i=1}^n w_i x_i + 2\beta_1 \sum_{i=1}^n w_i x_i^2 = 0$$

Sehingga diperoleh persamaan normal :

$$\sum_{i=1}^n w_i Y_i = \beta_0 \sum_{i=1}^n w_i + \beta_1 \sum_{i=1}^n w_i X_i \quad (4)$$

$$\sum_{i=1}^n w_i X_i Y_i = \beta_0 \sum_{i=1}^n w_i X_i + \beta_1 \sum_{i=1}^n w_i X_i^2 \quad (5)$$

Dari persamaan (4) diperoleh:

$$b_0 = \frac{\sum_{i=1}^n w_i Y_i - b_1 \sum_{i=1}^n w_i X_i}{\sum_{i=1}^n w_i} \quad (6)$$

Dari mensubstitusi persamaan (5) dan (6), diperoleh :

$$\begin{aligned} \sum_{i=1}^n w_i X_i Y_i &= \frac{\sum_{i=1}^n w_i Y_i - b_1 \sum_{i=1}^n w_i X_i}{\sum_{i=1}^n w_i} \sum_{i=1}^n w_i X_i + b_1 \sum_{i=1}^n w_i X_i^2 \\ &= \frac{\sum_{i=1}^n w_i Y_i - b_1 \sum_{i=1}^n w_i X_i}{\sum_{i=1}^n w_i} \sum_{i=1}^n w_i X_i + b_1 \sum_{i=1}^n w_i X_i^2 \\ b_1 &= \frac{\frac{\sum_{i=1}^n w_i X_i Y_i - \frac{\sum_{i=1}^n w_i Y_i \sum_{i=1}^n w_i X_i}{\sum_{i=1}^n w_i}}{\sum_{i=1}^n w_i X_i}}{\sum_{i=1}^n w_i X_i^2 - \frac{(\sum_{i=1}^n w_i X_i)^2}{\sum_{i=1}^n w_i}} \end{aligned} \quad (7)$$

b. WLS untuk Regresi Linear Berganda

Kriteria kuadrat terkecil untuk regresi linear berganda adalah sebagai berikut:

$$SSE = \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_{i1} - \dots - \beta_i X_{i,i})^2 \quad (8)$$

Sedangkan untuk WLS, masing-masing jumlah kuadrat error akan dikali dengan penimbang yaitu w_i , sehingga

$$SSE_w = Q_w = \sum_{i=1}^n w_i (Y_i - \beta_0 - \beta_1 X_{i1} - \beta_2 X_{i2} - \dots - \beta_i X_{i,i})^2 \quad (9)$$

Parameter $(\beta_0, \beta_1, \dots, \beta_p)$ diperoleh melalui diferensial Q_w terhadap setiap parameter β seperti pada kasus suatu variabel x. Sehingga diperoleh nilai dari koefisien regresinya dengan meminimalkan Q_w .

Meminimumkan kuadrat error terhadap β_0 dan β_1 yaitu dengan cara :

$$\frac{\partial Q_w}{\partial \beta_0}, \frac{\partial Q_w}{\partial \beta_1}, \dots, \frac{\partial Q_w}{\partial \beta_p} \quad (10)$$

Dari persamaan (9) akan menghasilkan p persamaan normal dan dengan melalui

proses aljabar matriks, maka diperoleh $\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}$. [10]

METODE PENELITIAN

Penelitian ini merupakan penelitian kajian teori dan terapan mengenai metode Weigthed Least Square pada kasus jumlah balita gizi buruk di provinsi Sulawesi Selatan tahun 2016. Data yang digunakan dalam penelitian ini adalah data sekunder yang diambil dari Dinas Kesehatan Provinsi Sulawesi Selatan. Data yang terdapat dalam penelitian ini adalah data mengenai gizi buruk yang ada di provinsi Sulawesi Selatan 2016.

Variabel Penelitian

1. Variabel Respon (Y)

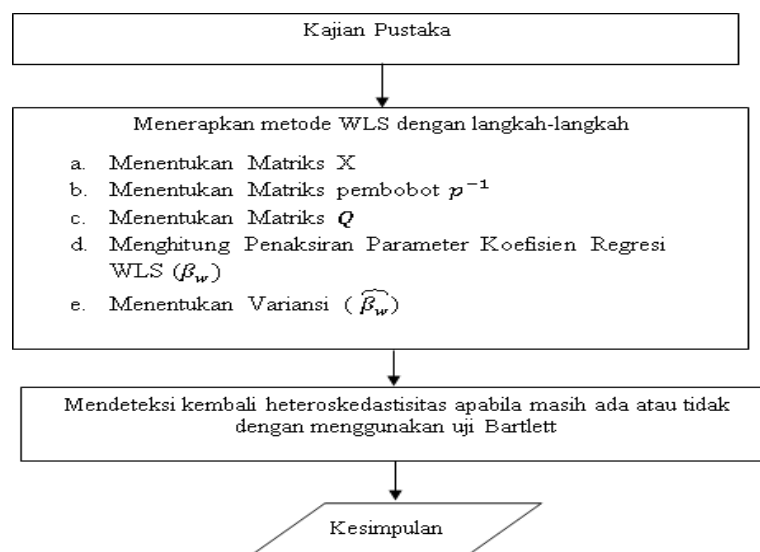
Variabel respon pada penelitian ini adalah jumlah balita gizi buruk di provinsi Sulawesi Selatan pada tahun 2016.

2. Variabel Prediktor (X)

Variabel prediktor adalah variabel yang dianalisis untuk mengamati efeknya terhadap variabel terikat. Yang menjadi variabel prediktor atau variabel independent pada penelitian ini adalah:

- a. Pemberian Asi Eksklusif (X_1)
- b. Rasio Puskesmas (X_2)
- c. Rasio dokter (X_3)
- d. Rasio Bidan (X_4)
- e. Rumah Tangga Ber-PHBS (X_5)
- f. Balita ditimbang (X_6)
- g. Pemberian Vitamin A (X_7)

Prosedur Penelitian



HASIL DAN PEMBAHASAN

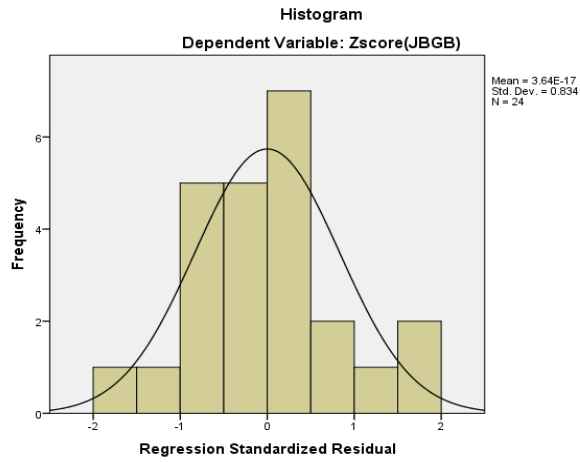
Pengujian Asumsi Klasik Regresi

1. Uji Multikolinearitas

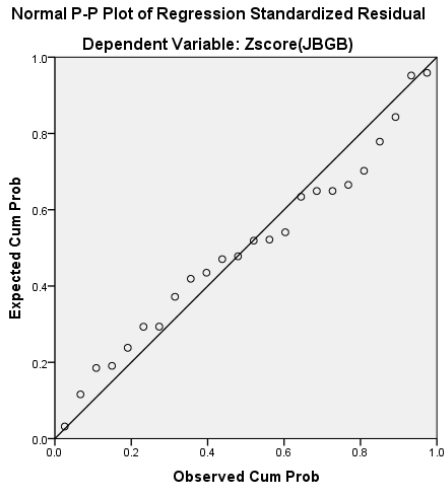
Tabel 1 Uji Multikolinearitas

Variabel	Collinearity Statistics	
	Tolerance	ViF
Pemberian asi eksklusif (X_1)	0.388	2.579
Rasio puskesmas (X_2)	0.626	1.597
Rasio dokter (X_3)	0.430	2.327
Rasio bidan (X_4)	0.298	3.352
Rumah tangga ber-PHBS (X_5)	0.765	1.306
Balita di timbang (X_6)	0.366	2.730
Pemberian vitamin A (X_7)	0.580	1.723

2. Uji Normalitas



Gambar 2 Grafik Histogram



Gambar 3 Grafik Normal P-Plot

Tabel 2 Uji Normalitas menggunakan Kolmogorov Smirnov

One-Sample Kolmogorov-Smirnov	Test	Unstandardized Residual
N		24
Normal Parameters ^{a,b}	Mean	0E-7
	Std. Deviation	.86748529
Most Extreme Differences	Absolute	.096
	Positive	.096
	Negative	-.070
Kolmogorov-Smirnov Z		.471
Asymp. Sig. (2-tailed)		.980

a. Test distribution is Normal.

b. Calculated from data.

3. Uji Autokorelasi

Tabel 3 Uji Korelasi

Model Summary ^b					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.497 ^a	.247	-.082	1.04007832	1.605

a. Predictors: (Constant), Zscore(PVA), Zscore(Bidan), Zscore(PHBS), Zscore(Puskesmas), Zscore(Dokter), Zscore(ASI), Zscore(BDT)

b. Dependent Variable: Zscore(JBGB)

Nilai Durbin-Watson yang tertera pada output SPSS disebut dengan DW hitung yaitu 1,605. Angka ini akan dibandingkan dengan kriteria penerimaan atau penolakan yang akan dibuat dengan nilai d_L dan d_U ditentukan berdasarkan jumlah variabel bebas dalam model regresi (k) dan jumlah sampelnya (n). Nilai d_L dan d_U dapat dilihat pada Tabel DW dengan tingkat signifikansi (*error*) 5% ($\alpha = 0,05$).

Jumlah variabel bebas: $k = 7$

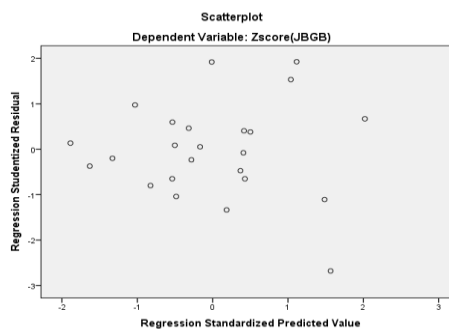
Jumlah sampel: $n = 24$

Dari hasil output di atas didapat nilai DW yang dihasilkan dari model regresi adalah 1,605. Sedangkan dari tabel DW dengan signifikansi 0,05 dan jumlah data (n) = 24, seta $k = 7$ (k adalah jumlah variabel bebas) diperoleh nilai $d_L = 0,7505$ dan nilai $d_U = 2,1743$. Karena nilai DW (1,605) berada pada daerah antara d_L dan d_U , maka tidak menghasilkan kesimpulan yang pasti (berada di daerah keragu-raguan) tapi disini kami mengasumsikan tidak terjadi gejala autokorelasi.

4. Uji Heteroskedastisitas

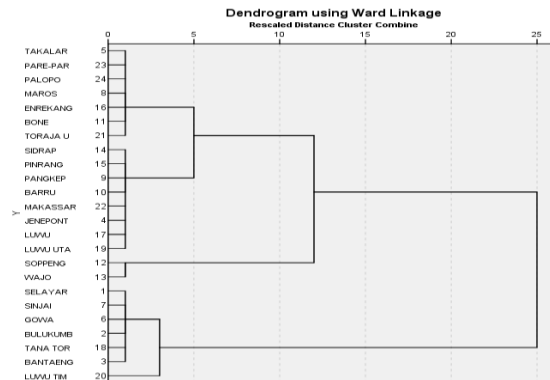
a. Uji Non-Formal

Gambar 4 Plot antara Estimasi Y dengan Residual



b. Uji Formal

Gambar 5 Dendogram metode Werd dengan Jarak Euclid dari 24 Kabupaten/Kota



Berdasarkan hasil dendrogram, diperoleh 4 kelompok jumlah balita gizi buruk dan pengelompokan ini akan digunakan juga pada tindakan perbaikan.

Tabel 4 Pengelompokan data Kabupaten/Kota Provinsi Sulawesi Selatan Berdasarkan Jumlah Balita Gizi Buruk Tahun 2016

Kelompok	Kabupaten/Kota	S^2
1	Selayar, Bulukumba, Bantaeng, Gowa, Sinjai, Tana Toraja, dan Luwu Timur.	0,230
2	Jeneponto, Pangkep, Barru, Sidrap, Pinrang, Luwu, Luwu Utara, dan Makassar.	0,023
3	Takalar, Maros, Bone, Enrekang, Toraja Utara, Pare-pare, dan Palopo.	0,067
4	Soppeng dan Wajo	0,003

Hipotesis statistik untuk pengujian homogenitas variansi adalah:

$$H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 = \sigma_5^2$$

H_a : Paling sedikit ada satu tanda sama dengan tidak berlaku.

Langkah-langkah perhitungan :

- Variansi dari setiap kelompok

$$S_1^2 = \frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n - 1} = 0,230$$

Dengan cara yang sama, diperoleh variansi untuk S_2^2, S_3^2, S_4^2 dan S_5^2 berturut-turut sebagai berikut:

$$S_2^2 = 0,023, S_3^2 = 0,067 \text{ dan } S_4^2 = 0,003$$

- Tabel proses perhitungan homogenitas variansi

Kelompok	Dk	1/dk	S^2	$\log S^2$	dk $\log S^2$	dk S^2
1	6	0.167	0.230	0.6383	3.8296	1.3800
2	7	0.143	0.023	1.6383	11.4679	0.1610
3	6	0.167	0.067	1.1739	7.0436	0.4020
4	1	1.000	0.003	2.5229	2.5229	0.0030
Jumlah	20	1.476			24.8640	1.9460

- Menghitung variansi gabungan

$$S^2 = \frac{\sum (dk S_k^2)}{\sum dk} = \frac{1,94600}{20} = 0,0973$$

$$\log S^2 = 1,4933$$

4. Menghitung nilai B

$$B = (\sum dk) \log S^2 = 29,8667$$

5. Menghitung nilai X^2

$$\begin{aligned} X^2 &= (\ln 10) \{B - \sum dk \log S_i^2\} = 2,303 \times \{29,8667 - 24,8640\} \\ &= 2,303 \times 5,0028 = 11,5214 \end{aligned}$$

Jadi, untuk $\alpha = 5\%$ dari daftar distribusi X^2 dengan $dk = 4-1=3$ diperoleh $X_{tabel}^2 = 7,81$ yang ternyata bahwa $X_{hitung}^2 (11,5214) > X_{tabel}^2 (7,81)$ sehingga hipotesis yang menyatakan variansi homogen ditolak pada taraf $\alpha = 5\%$.

Weighted Least Square (WLS)

Salah satu asumsi penting dalam membuat model regresi berganda adalah $var(\varepsilon_i)$ harus sama dengan σ^2 (konstan), atau dengan kata lain, semua residual atau error mempunyai variansi yang sama. [10] Sedangkan apabila variansi residual tidak identik mengakibatkan $var(\varepsilon_i)$ tidak sama untuk setiap i , dinotasikan $var(\varepsilon_i) = \sigma_i^2$ disebut heteroskedastisitas. Salah satu cara untuk mengatasi heteroskedastisitas adalah dengan pembobotan.

Pada kondisi heteroskedastisitas, mula-mula $var(\varepsilon_i) = \sigma_i^2$ pada setiap $\varepsilon_i; i = 1, 2, \dots, n$, jika ditulis dalam bentuk matriks, yaitu:

$$var(b_w) = \begin{bmatrix} var(\varepsilon_1) & cov(\varepsilon_1, \varepsilon_2) & \dots & cov(\varepsilon_1, \varepsilon_n) \\ cov(\varepsilon_2, \varepsilon_1) & var(\varepsilon_2) & \dots & cov(\varepsilon_2, \varepsilon_n) \\ \vdots & \vdots & \ddots & \vdots \\ cov(\varepsilon_n, \varepsilon_1) & cov(\varepsilon_n, \varepsilon_2) & \dots & var(\varepsilon_n) \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{bmatrix} = \mathbf{V} \quad (11)$$

Karena adanya asumsi tentang homoskedastisitas, maka setiap kesalahan pengganggu mempunyai varian yang sama dan tidak ada korelasi serial, artinya antar kesalahan pengganggu yang satu dengan yang lainnya bebas, yaitu $cov(\varepsilon_n, \varepsilon_m) = 0$.

Agar $var(\varepsilon_i) = \sigma^2$ maka dilakukan pembobotan/transformatasi pada setiap $\varepsilon_i; i = 1, 2, \dots, n$ yaitu:

$$\varepsilon_i \rightarrow \frac{\varepsilon_i}{\sigma_i}; i = 1, 2, \dots, n \quad (12)$$

Jika ditulis dalam bentuk matriks menjadi:

$$\begin{bmatrix} \frac{1}{\sigma_1^2} & 0 & \dots & 0 \\ 0 & \frac{1}{\sigma_2^2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sigma_n^2} \end{bmatrix} X \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Katakan misalnya:

$$V^{-1} = \begin{bmatrix} \frac{1}{\sigma_1^2} & 0 & \dots & 0 \\ 0 & \frac{1}{\sigma_2^2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sigma_n^2} \end{bmatrix} \text{ dan } \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Matriks V^{-1} berupa matriks diagonal yang berelemenkan nilai-nilai pembobot, yaitu $\frac{1}{\sigma_i^2} = w_i$, matriks ini disebut matriks pembobot. Dari matriks V^{-1} maka diperoleh matriks pembobot lainnya, yaitu matriks P^{-1} yang juga merupakan matriks diagonal, dan elemennya merupakan akar elemen V^{-1} , yaitu $\frac{1}{\sigma_i} = w_i^{0,5}$.

$$P^{-1} = \begin{bmatrix} w_1^{0,5} & 0 & \dots & 0 \\ 0 & w_2^{0,5} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_n^{0,5} \end{bmatrix} \quad (13)$$

Atau matriks \mathbf{P} jika ditulis dalam matriks, yaitu:

$$P = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n \end{bmatrix}$$

dimana matriks P ditentukan sedemikian rupa sehingga memenuhi :

$$P^T P = P P = P^2 = V \quad (14)$$

Diketahui model regresi linear dirumuskan sebagai berikut:

$Y = X\beta + \varepsilon$ kemudian diberikan penimbang sehingga diperoleh model regresi terboboti yaitu :

$$P^{-1}Y = P^{-1}X\beta + P^{-1}\varepsilon \quad (15)$$

Jika vektor residual $P^{-1}\varepsilon$ disimbolkan dengan f , atau:

$$f = P^{-1}\varepsilon \quad (16)$$

Kemudian $P^{-1}Y$ disimbolkan menjadi Y_w , atau:

$$Y_w = P^{-1}Y \quad (17)$$

Kemudian $P^{-1}X$ disimbolkan menjadi Q_w , atau:

$$Q_w = P^{-1}X \quad (18)$$

Maka model regresi terboboti adalah

$$Y_w = Q_w\beta + f$$

Dengan

$$P = \begin{bmatrix} q_{01} & q_{11} & \dots & q_{k1} \\ q_{02} & q_{12} & \dots & q_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ q_{0n} & q_{1n} & \dots & q_{kn} \end{bmatrix}$$

Error terboboti yaitu $= P^{-1}\varepsilon$, memiliki sifat seperti error pada regresi OLS, yaitu $var(\varepsilon) = I\sigma^2$ (homoskedastisitas). Penjabaran $var(f)$ secara rinci adalah sebagai berikut :

Diketahui $var(X) = E(X - E(X))(X - E(X))^T$

$E(\varepsilon) = 0$

$$var(\varepsilon) = \sigma^2$$

dan diperoleh:

$$\begin{aligned} var(f) &= E(f - E(f))(f - E(f))^T \\ &= E(f - 0)(f - 0)^T \\ &= E(ff^T) \\ &= E(P^{-1}\varepsilon(P^{-1}\varepsilon)^T) \\ &= E(P^{-1}\varepsilon\varepsilon^T P^{-1}) \\ &= P^{-1}E(\varepsilon\varepsilon^T)P^{-1} \\ &= P^{-1}\sigma^2 P^{-1} \\ &= I\sigma^2 \end{aligned} \quad (19)$$

Melalui penjabaran diatas, menunjukkan bahwa error terboboti f telah memenuhi asumsi homoskedastisitas.

Jika menggunakan notasi matriks, meminimumkan jumlah kuadrat galat pada WLS sama dengan meminimumkan jumlah kuadrat galat pada OLS yang diberi pembobot, yaitu:

$$(P^{-1}\varepsilon)^T (P^{-1}\varepsilon) = f^T f$$

$$\begin{aligned}
&= [f_1 f_2 \dots f_n] \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix} \\
&= f_1^2 + f_2^2 + \dots + f_n^2 \\
&= \sum_{i=1}^n f_i^2
\end{aligned} \tag{20}$$

Penaksir kuadrat terkecil b_w didapatkan dengan cara meminimumkan jumlah kuadrat galatnya, yaitu:

$$Y_w = Qb_w + f$$

Maka $f = Y_w - Qb_w$, sehingga

$$\begin{aligned}
L &= \sum_{i=1}^n f_i^2 \\
&= f^T f \\
&= (Y_w - Qb_w)^T (Y_w - Qb_w) \\
&= Y_w^T Y_w - b_w^T Q^T Y_w - Y_w^T Q b_w + b_w^T Q^T Q b_w \\
&= Y_w^T Y_w - 2b_w^T Q^T Y_w + b_w^T Q^T Q b_w
\end{aligned}$$

Penaksir kuadrat terkecil itu harus memenuhi:

$$\frac{\partial L}{\partial \beta} = 0$$

$$-2Q^T Y_w + 2Q^T Q b_w = 0$$

Penyederhanaannya menjadi:

$$Q^T Q b_w = Q^T Y_w$$

Kedua sisi dikalikan $(Q^T Q)^{-1}$, maka diperoleh penyelesaian untuk b_w , yaitu:

$$b_w = (Q^T Q)^{-1} Q^T Y_w \tag{21}$$

Maka penaksir parameter koefisien regresi didapatkan dengan rumus berikut:

$$\begin{aligned}
b_w &= \begin{bmatrix} b_{0w} \\ b_{1w} \\ \vdots \\ b_{kw} \end{bmatrix} = (Q^T Q)^{-1} Q^T Y_w \\
&= ((P^{-1}X)^T P^{-1}X)^{-1} (X P^{-1})^T (P^{-1}Y) \\
&= (X^T P^{-1} P^{-1} X)^{-1} X^T P^{-1} P^{-1} Y \\
&= (X^T V^{-1} X)^{-1} X^T V^{-1} Y
\end{aligned} \tag{22}$$

Matrik V bukan matrik identitas, tetapi seperti ketentuan di atas pada persamaan (14).

Jika $Y = AX$, dimana $A = \text{konstan}$

Maka $E(Y) = E(AX) = AE(X)$

$$\begin{aligned}
\text{var}(Y) &= E(Y - E(Y))(Y - E(Y))^T \\
&= E(AX - E(AX))(AX - E(AX))^T \\
&= E(A(X - E(X)))(A(X - E(X)))^T \\
&= AE[(X - E(X))(X - E(X))^T A^T] \\
&= A\sigma_x^2 A^T
\end{aligned} \tag{23}$$

Dengan mengacu pada persamaan (23), variansi dari b_w , yaitu:

$$\begin{aligned}
\text{var}(b_w) &= \text{var}(Q^T Q)^{-1} Q^T Y_w \\
&= (Q^T Q)^{-1} Q^T \text{var}(Y_w) ((Q^T Q)^{-1} Q^T)^T \\
&= (Q^T Q)^{-1} Q^T \sigma^2 Q (Q^T Q)^{-1} \\
&= (Q^T Q)^{-1} (Q^T Q) (Q^T Q)^{-1} \sigma^2 = I (Q^T Q)^{-1} \sigma^2 \\
&= (Q^T Q)^{-1} \sigma^2
\end{aligned} \tag{24}$$

Sehingga variansi penaksir parameter regresi OLS dalam notasi matriks, yaitu:

$$\begin{aligned}
\text{var}(b_w) &= \begin{bmatrix} \text{var}(b_{0w}) & \text{cov}(b_{0w}, b_{1w}) & \dots & \text{cov}(b_{0w}, b_{kw}) \\ \text{cov}(b_{1w}, b_{0w}) & \text{var}(b_{1w}) & \dots & \text{cov}(b_{1w}, b_{kw}) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(b_{kw}, b_{0w}) & \text{cov}(b_{kw}, b_{1w}) & \dots & \text{var}(b_{kw}) \end{bmatrix} \\
&= (Q^T Q)^{-1} \sigma^2
\end{aligned}$$

$$= (X^T V^{-1} X)^{-1} \sigma^2 \quad (25)$$

Jadi, diperoleh:

$$b_w = (X^T V^{-1} X)^{-1} X^T V^{-1} Y$$

$$\text{var}(b_w) = (X^T V^{-1} X)^{-1} \sigma^2$$

Analisis Statistika Deskriptif

Tabel 5 Statistika Deskriptif Data Jumlah Balita Gizi Buruk di Provinsi Sulawesi Selatan Tahun 2016

Indikator	N	Minimum	Maximum	Rerata	Simpangan Baku	Total
ASI	24	539	5194	2388.38	1319.56	57321
Puskesmas	24	6	35	17.79	6.91	427
Dokter	24	30	1239	119.46	240.39	2867
Bidan	24	74	267	152.96	53.91	3671
PHBS	24	2000	49429	21576.13	12404.57	517827
BDT	24	5609	8834	7629.33	839.91	183104
PVA	24	4754	15289	9246.21	2301.59	221909
JBGB	24	0	22	7.21	6.26	173

Analisis Inferensial

1. Standarisasi Data

Proses standarisasi dilakukan apabila diantara variabel-variabel yang diteliti terdapat perbedaan ukuran satuan yang besar. Perbedaan satuan yang mencolok dapat mengakibatkan perhitungan pada analisis regresi menjadi tidak valid. Untuk itu, perlu dilakukan proses standarisasi dengan melakukan transformasi (standarisasi) pada data asli sebelum dianalisis lebih lanjut. Transformasi dilakukan terhadap variabel yang relevan ke dalam bentuk z skor

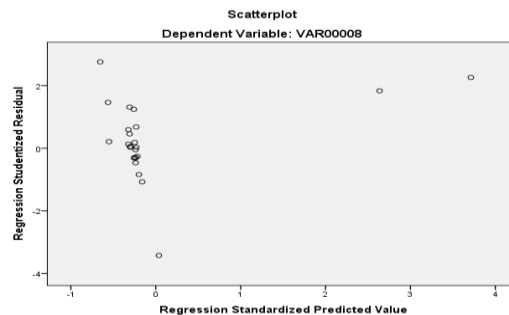
2. Mengatasi Heterokedastisitas Menggunakan Metode WLS

- Menentukan matriks rancangan X sesuai persamaan
- Menentukan matriks pembobot P^{-1}
- Merancang Matriks Q
- Menghitung Penaksiran Parameter Koefisien WLS

Pengujian Kembali Heteroskedastisitas

1. Uji Non-Formal

Gambar 6 Plot antara Estimasi dengan Residual



2. Uji Formal

Tabel 6 Pengelompokan data terbobot Kabupaten/Kota Provinsi Sulawesi Selatan Berdasarkan Jumlah Balita Gizi Buruk Tahun 2016

Kelompok	Kabupaten/Kota	S^2
1	Selayar dan Gowa	0,016
2	Bulukumba, Jeneponto, Takalar, Sinjai, Maros, Pangkep, Barru, Soppeng, Wajo, Sidrap, Luwu, Makassar, Pare-pare dan Palopo	0,039
3	Bantaeng, Pinrang, Tana Toraja dan Luwu Timur	0,019
4	Bone, Enrekang, Luwu Utara dan Toraja Utara	0,071

a. Variansi setiap kelompok

$$S_1^2 = \frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n-1} = 0,016$$

b. Variansi Gabungan

$$S^2 = \frac{\sum dk S_i^2}{\sum dk} = \frac{0,79300}{20} = 0,0396. \log S^2 = 1,5187$$

c. Nilai B

$$B = (\sum dk) \log S^2 = 30,3740$$

d. Nilai X^2

$$X^2 = (\ln 10) \{B - \sum dk \log S_i^2\} = 2,303 \times \{30,3740 - 28,7220\} = 2,303 \times 1,8039 = 3,8039$$

Jadi, untuk $\alpha = 5\%$ dari daftar distribusi X^2 dengan $dk = 4-1=3$ diperoleh $X_{tabel}^2 = 7,81$ yang ternyata bahwa $X_{tabel}^2(3,8039) < X_{tabel}^2(7,81)$ sehingga hipotesis yang menyatakan variansi homogen diterima pada taraf $\alpha = 5\%$

Setelah dilakukan perbaikan model dengan estimasi WLS, langkah selanjutnya adalah membahas model tersebut. Seperti yang telah dijelaskan pada bab sebelumnya, suatu asumsi yang kritis dari model regresi linear klasik adalah bahwa kesalahan pengganggu ε_i semuanya harus mempunyai varian yang sama.

Tabel 7 Analisa perbandingan R^2, S dan X^2 , S dengan dua uji estimasi

	Estimasi OLS	Estimasi WLS
	R^2	0,247
S	0,8675	0,7806
X^2	Ada heterokedastisitas	Homokedastisitas

Berdasarkan hasil perbandingan dari dua estimasi di atas, dapat dilihat bahwa nilai R^2 pada masing-masing estimasi yaitu dimana pada OLS 0,247 sangat rendah dan pada WLS 0,810 ini sangat baik. Namun jika dilihat dari nilai S , pada estimasi OLS mempunyai nilai S yang lebih tinggi dibandingkan pada estimasi WLS yaitu sebesar 0,8675 dan setelah diboboti pada estimasi WLS nilai S mengalami penurunan menjadi 0,7806 yang disertai dengan kenaikan nilai R^2 yang menunjukkan model baik. Jika dilihat juga pada nilai χ^2 , pada OLS mengalami

heteroskedastisitas dan setelah data mengalami estimasi WLS menjadi homoskedastisitas. Selain itu dapat dilihat juga pada tingkat signifikansi $< 0,05$ yang menunjukkan hasil perhitungan secara statistik bermakna. Jadi, dapat disimpulkan bahwa setelah data ditransformasi menggunakan metode WLS, maka model regresi berganda tersebut telah memenuhi asumsi homoskedastisitas

KESIMPULAN

Berdasarkan pembahasan pada Bab IV, maka kesimpulan pada penelitian ini adalah sebagai berikut:

1. Metode WLS efektif digunakan dalam mengatasi masalah heteroskedastisitas karena melalui prosedur pembobotan pada error OLS (ϵ) yang menghasilkan varian error WLS (f) menjadi homoskedastisitas.
2. Efektifitas WLS dapat juga dilihat dari studi kasus dimana pada OLS terdapat masalah heteroskedastisitas dan pada WLS menjadi homoskedastisitas. Serta nilai $R^2 OLS < R^2 WLS$ dan nilai $S^2 OLS > S^2 WLS$.

DAFTAR PUSTAKA

1. Nachrowi, D.N. 2008. *Penggunaan Teknik Ekonometrika*. Jakarta: Raja Grafindo Persada
2. Neter, 1997. *Penanganan Heteroskedastisitas Pada regresi Berganda Melalui Pendekatan Weighed Least Square*.
3. Uthami, I. A.P., Sukarsa, I.K.G., Kencana, I.P.K. 2013. Regresi Kuantil Median untuk Mengatasi Heteroskedastisitas pada Analisis Regresi. *E-Jurnal Matematika* Vol. 2. No. 1, Januari 2013, 6-13.
4. Gujarati, N. D. 2003. *Basic Econometrics*. 4th ed. New York: McGraw-Hill Companies, Inc.
5. Sembiring, R.K. 1995. *Analisis Regresi*. Bandung: ITB.
6. Kutner, Nachtsheim, Neter, (2004). *Model Umum Regresi Berganda*
7. Widarjono. (2005). *Analisis Statistika Multivariat Terapan*. UPP STIM YKPN. Yogyakarta.
8. Ghazali, Imam. (2012). *Aplikasi Analisis Multivariate Dengan Program IBM SPSS 20*. Edisi Keenam, Penerbit Universitas Diponegoro.
9. Notoatmojo, Soekidjo. (2010). *Metodologi Penelitian Kesehatan*. Jakarta: Rineka cipta.
10. Winahju, Wiwiek Setya. 2015. *Regresi Terboboti (Weighted Regression/ Weighted Least Square)*. (<http://oc.its.ac.id/jurusan.php?fid=1&jid=3>). Diakses tanggal 16 Februari 2015).

