

MENGATASI HETEROSKEDASTISITAS PADA REGRESI DENGAN MENGGUNAKAN *WEIGHTED LEAST SQUARE*

Putu Ayu Maziyya^{§1}, I Komang Gde Sukarsa², Ni Made Asih³

¹Jurusan Matematika, Fakultas MIPA – Universitas Udayana [Email: nie.phoez.pam@gmail.com]

²Jurusan Matematika, Fakultas MIPA – Universitas Udayana [Email: sukarsakomang@yahoo.com]

³Jurusan Matematika, Fakultas MIPA – Universitas Udayana [Email: asihmath77@gmail.com]

[§]Corresponding Author

ABSTRACT

In the regression analysis we need a method to estimate parameters to fulfill the BLUE characteristic. There are assumptions that must be fulfilled homoscedasticity one of which is a condition in which the assumption of error variance is constant (same), infraction from the assumptions homoskedasticity called heteroscedasticity. The Consequence of going heteroscedasticity can impact OLS estimators still fulfill the requirements of not biased, but the variant obtained becomes inefficient. So we need a method to solve these problems either by Weighted Least Square (WLS). The purpose of this study is to find out how to overcome heteroscedasticity in regression with WLS. Step of this research was do with the OLS analysis, then testing to see whether there heteroscedasticity problem with BPG method, the next step is to repair the beginning model by way of weighting the data an exact multiplier factor, then re-using the OLS procedure to the data that have been weighted, the last stage is back with a heteroscedasticity test BPG method, so we obtained the model fulfill the assumptions of homoskedasticity. Estimates indicate that the WLS method can resolve the heteroscedasticity, with exact weighting factors based on the distribution pattern of data seen.

Keywords: *heteroscedasticity, OLS, WLS*

1. Pendahuluan

Dalam analisis regresi diperlukan suatu metode untuk menduga parameter agar memenuhi sifat BLUE (*Best Linear Unbiased Estimator*), salah satu metode yang paling sering digunakan adalah *Ordinary Least Square* (OLS) atau sering disebut dengan *Metode Kuadrat Terkecil* (MKT). Salah satu asumsi klasik yang harus dipenuhi dalam estimasi OLS agar hasil estimasinya dapat diandalkan, yaitu ragam sisaan homogen $E(u_i^2) = \sigma^2$ (*homoskedastisitas*). Pelanggaran terhadap asumsi homoskedastisitas disebut *heteroskedastisitas*, yang artinya galat bersifat tidak konstan. Konsekuensi dari terjadi heteroskedastisitas dapat mengakibatkan penduga OLS yang diperoleh tetap memenuhi persyaratan tak bias, tetapi varian yang diperoleh menjadi tidak efisien, artinya varian cenderung membesar

sehingga tidak lagi merupakan varian yang kecil. Dengan demikian model perlu diperbaiki dulu agar pengaruh dari heteroskedastisitas hilang (Gujarati [3]).

Mengingat secara statistik permasalahan heteroskedastisitas tersebut dapat mengganggu model yang akan diestimasi, bahkan dapat menyesatkan kesimpulan yang diambil maka diperlukan metode untuk mengatasi masalah tersebut. Metode *Weighted Least Square* (WLS) merupakan metode alternatif yang dapat mengatasi heteroskedastisitas. 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.

Metode WLS memiliki keunggulan dibandingkan dengan OLS yaitu metode WLS bisa mengatur pentingnya setiap observasi

karena pada OLS diasumsikan bahwa nilai duga parameter regresi bernilai sama untuk setiap observasi. Mekanisme kerja WLS yaitu apabila varian variabel gangguan tidak diketahui maka untuk menyelesaikan masalah heteroskedastisitas adalah dengan mengetahui pola heteroskedastisitas itu sendiri, jika variabel gangguan proporsional terhadap X_i maka model akan dibagi dengan $\sqrt{X_i}$, jika varian variabel gangguan adalah proporsional terhadap X_i^2 sehingga model akan dibagi dengan X_i , selain proporsional dengan X_i dan X_i^2 bisa juga diasumsikan bahwa pola varian variabel gangguan adalah proporsional terhadap $[E(Y_i)]^2$ sehingga dibagi dengan $E(Y_i)$ (Gujarati [3]).

Pada penelitian ini penulis ingin mengetahui bagaimana cara mengatasi heteroskedastisitas pada regresi dengan metode *Weighted Least Square*. Pada penelitian ini digunakan data yang mengandung masalah heteroskedastisitas.

1.1 Metode Kuadrat Terkecil

Model persamaan regresi linier berganda:

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

Metode kuadrat terkecil adalah metode yang bertujuan untuk meminimumkan jumlah kuadrat galat (*sum square error*). Pendugaan koefisien regresi linear berganda dengan metode kuadrat terkecil dapat dilakukan dengan menggunakan persamaan matriks [4], sebagai berikut:

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

1.2 Weighted Least Square

Pada estimasi OLS, salah satu asumsi yang harus dipenuhi adalah

$$E\{\varepsilon\varepsilon'\} = \sigma^2$$

yaitu galat bersifat homoskedastisitas. Greene [1] menyatakan apabila terjadi pelanggaran asumsi tersebut, yakni kemungkinan variansinya tidak sama (heteroskedastisitas), maka metode yang dapat digunakan untuk menduga koefisien regresi adalah metode Generalized Least Square (GLS). Penaksir β pada metode GLS diperoleh dengan cara mentransformasikan model regresi

linier terlebih dahulu sehingga dapat memenuhi asumsi-asumsi pada OLS.

Asumsi yang diberikan pada metode GLS adalah: $E\{\varepsilon\} = 0$ dan $E\{\varepsilon\varepsilon^T\} = \sigma^2\Omega$. Dengan Ω merupakan matriks simetrik definit positif dan nonsingular yang diketahui dan berukuran $n \times n$, sehingga Ω dapat difaktorisasi menjadi:

$$\Omega = C\Lambda C^T$$

Dengan C merupakan matriks dengan kolom-kolomnya adalah vektor ciri dari Ω dan Λ merupakan matriks diagonal dengan unsur diagonalnya akar ciri dari Ω .

Didefinisikan variabel baru $T = C\Lambda^{1/2}$ dengan $\Lambda^{1/2}$ merupakan matriks diagonal yang entri pada diagonal ke- i adalah $\sqrt{\lambda_i}$ sehingga $\Omega = TT^T$. Diberikan $P^T = C\Lambda^{-1/2}$

Model regresi $Y = X\beta + \varepsilon$ dapat ditransformasi dengan mengalikan matriks P sehingga diperoleh:

$$PY = PX\beta + P\varepsilon$$

Atau

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

Variansi pada ε_* adalah

$$E\{\varepsilon^*\varepsilon^{*T}\} = P\sigma^2\Omega P^T = \sigma^2 I$$

Oleh karena itu, dalam mengestimasi dengan GLS dibutuhkan matriks P yang merupakan dekomposisi dari matriks Ω . Akan tetapi dalam prakteknya matriks Ω sulit untuk ditentukan, maka secara praktis menurut Gujarati [2] GLS diperkhusus dengan melihat pada pola yang ditunjukkan sisaan terhadap variabel X yang disebut dengan *Weighted Least Square (WLS)* yaitu beberapa melakukan pembobotan dengan $\frac{1}{\sqrt{X_i}}, \frac{1}{X_i}, \frac{1}{E(Y_i)}, \frac{1}{\sigma_i}$.

1.3 Heteroskedastisitas

Menurut Widarjono[6] terdapat konsekuensi terhadap estimator OLS jika ada masalah heteroskedastisitas namun tetap mempertahankan asumsi metode OLS dan ada kemungkinan tidak akan menghasilkan estimator linear tidak bias yang terbaik, apabila varian estimator β_1 mengandung heteroskedastisitas maka variansinya sebagai berikut:

$$\text{Var}(\beta_1) = \frac{\sum x_i^2 \sigma_i^2}{(\sum x_i^2)^2}; i = 1, 2, \dots, n$$

sementara itu varian estimator β_1 tanpa masalah heteroskedastisitas adalah sebagai berikut:

$$\text{Var}(\beta_1) = \frac{\sigma^2}{\sum x_i^2}$$

Dengan demikian adanya heteroskedastisitas menyebabkan estimator β_1 tidak lagi mempunyai varians yang minimum jika menggunakan metode OLS. Oleh karena itu, estimator β_1 yang didapatkan akan mempunyai karakteristik sebagai berikut:

- Estimator metode OLS masih tidak bias
- Estimator metode OLS masih linier
- Namun estimator metode OLS tidak lagi mempunyai varians yang minimum (*no longer best*)

Jadi dengan adanya heteroskedastisitas, estimator OLS tidak menghasilkan estimator yang *Best Linear Unbiased Estimator* (BLUE).

Dalam kenyataannya, asumsi *homoskedastisitas* dari kesalahan pengganggu ε mungkin tidak bisa dipenuhi, dengan kata lain varian dari kesalahan pengganggu bersifat *heteroskedastisitas*, yaitu $E(\varepsilon_i^2) = \sigma_i^2$. Hal ini dapat dipahami jika diperhitungkan faktor-faktor yang menjadi penyebab adanya kesalahan pengganggu ε dalam model regresi. Faktor kesalahan pengganggu ε dimasukkan kedalam model untuk dapat memperhitungkan kesalahan-kesalahan yang mungkin terjadi dalam pengukuran dan kesalahan karena mengabaikan variabel-variabel tertentu. Untuk memperkirakan bahwa ε bervariasi secara sistematis dengan variabel bebas X .

Konsekuensi dari pelanggaran asumsi *homoskedastisitas* adalah sebagai berikut:

- Penduga (estimator) yang diperoleh tetap memenuhi persyaratan tidak bias. Sifat ketidakbiasan tidak tergantung pada varian galat. Jika dalam model regresi ada *heteroskedastisitas*, maka kita tetap memperoleh nilai parameter yang tidak bias karena sebagai penduga tidak bias tidak memerlukan asumsi bahwa varian galat harus konstan.

- Varian penduga yang diperoleh akan menjadi tidak efisien, artinya penduga tersebut tidak memiliki varian terkecil diantara penduga-penduga tidak bias lainnya.

Namun asumsi ini sangat penting artinya dalam analisis regresi mengingat kaitannya dengan estimasi *standar error* koefisien regresi. Sebagaimana diketahui bahwa *standar error* ini memiliki peran dalam pembentukan t hitung maupun F hitung yang berkaitan akan menjadi terlalu (*overestimated*) yang mungkin selanjutnya menghasilkan kesimpulan bahwa kasus spesifik β_1 adalah kelihatannya signifikan, walaupun sebenarnya tidak signifikan. Oleh karena itu jika asumsi homoskedastisitas tidak dipenuhi maka hasil uji t tidak menentu. Berdasarkan yang kita ketahui:

$$t = \frac{b_1}{S_{b_1}}$$

Dari penjelasan di atas, *heteroskedastisitas* merupakan masalah yang berpotensi serius, karena mungkin merusak seluruh bangun standar, demikian pula dengan prosedur pengujian hipotesis dan estimator OLS yang begitu rutin digunakan. Untuk itu, dalam segala studi kongkret, khususnya yang melibatkan data lintas-sektoral (*cross section*), kita mesti menentukan apakah menghadapi masalah *heteroskedastisitas*.

Menurut Gujarati [3], untuk mendeteksi masalah *heteroskedastisitas* dapat dilakukan dengan metode formal dan informal. Metode formal dapat dilakukan dengan uji statistika diantaranya *Uji park*, *Uji White*, *Uji Glejser*, *Uji Korelasi rank dari Spearman*, dan *Uji Breusch Pagan Godfrey (BPG)*. Dalam tulisan ini akan digunakan *Breusch Pagan Godfrey (BPG)* dalam mendeteksi masalah *heteroskedastisitas*.

2. METODE PENELITIAN

Langkah-langkah yang dilakukan pada penelitian ini adalah:

- Membuat model awal pada regresi linier berganda dengan menggunakan metode OLS.

- Setelah memperoleh persamaan regresi, kemudian dilakukan pengujian untuk melihat

Tabel 1. Anova

Sumber Variansi	JK	Df
Regresi	7210,980	2
Sisaan	2112,408	82
Total	9323,388	84

- Dependent Variable: Y
- Predictors: (Constant), X2, X1

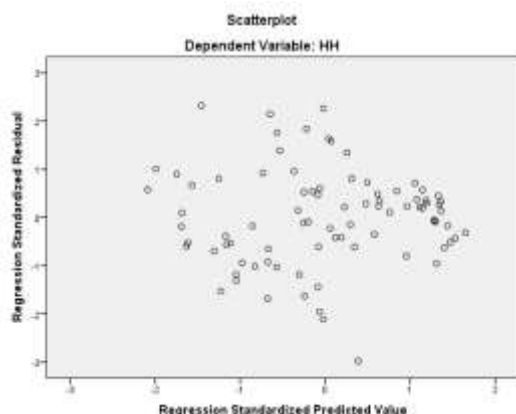
apakah terdapat masalah heteroskedastisitas atau tidak dengan menggunakan uji BPG.

- Tahap berikutnya adalah melakukan perbaikan pada model awal menggunakan metode WLS yaitu pembobotan data dengan suatu faktor pengali (pembobot) yang tepat, dalam penelitian ini pembobotan yang diberikan adalah $\frac{1}{\sqrt{X_i}}, \frac{1}{X_i}, \frac{1}{E(Y_i)}$, dan $\frac{1}{\sigma_i}$ (residual kuadrat).
- Kemudian kembali menggunakan prosedur OLS terhadap data yang telah diboboti. Lalu menganalisa kembali pada data yang telah diboboti tersebut apakah masih mengandung heteroskedastisitas.
- Tahap terakhir yaitu melakukan uji heteroskedastisitas kembali dengan uji BPG, sehingga model memenuhi asumsi homoskedastisitas.

3. HASIL DAN PEMBAHASAN

3.1 Uji Heteroskedastisitas

Heteroskedastisitas secara nonformal, digunakan untuk mendeteksi adanya heteroskedastisitas dengan melihat ada tidaknya pola tertentu, grafiknya sebagai berikut:



Dari gambar [3] dapat disimpulkan bahwa terdapat pola sebaran yang menunjukkan titik-titik yang terus-menerus bergerak mendekati nol. Berdasarkan pola tersebut mengindikasikan adanya heteroskedastisitas. Jika diketahui bentuk spesifik dari heteroskedastisitas pada gambar, maka dapat melakukan pembobotan terhadap variabel terikat dan variabel bebas sesuai dengan bentuk heteroskedastisitasnya. Namun pada gambar [3] sulit untuk mengidentifikasi pola sebarannya, oleh karena itu usaha perbaikan model untuk menghilangkan heteroskedastisitas akan diboboti dengan $\frac{1}{\sqrt{X_i}}, \frac{1}{X_i}, \frac{1}{E(Y_i)}$, dan $\frac{1}{\sigma_i}$.

Agar mendapat hasil yang lebih pasti untuk melihat adanya heteroskedastisitas atau tidak, terdapat beberapa uji yang dapat dilakukan. Salah satunya yang digunakan di dalam penelitian ini adalah *Uji Breusch Pagan Godfrey (BPG)*.

Hipotesis:

- H_0 : tidak ada heteroskedastisitas
- H_1 : ada heteroskedastisitas

Kriteria uji BPG adalah jika:

$P\text{-value} \leq \alpha$, maka H_0 ditolak

$P\text{-value} > \alpha$, maka H_0 diterima

Dengan menggunakan SPSS 20 diperoleh tabel Anova berikut:

Berdasarkan tabel 1 diperoleh nilai varian:

$$\begin{aligned} \delta^2 &= \frac{RSS}{n} \\ &= \frac{2112,408}{85} \\ &= 24,85 \end{aligned}$$

Kemudian diregresikan terhadap sisaan kuadrat yang dibagi dengan nilai varian, modelnya sebagai berikut:

$$\frac{e^2}{\delta^2} = \beta_0 + \beta_1 PP_i + \beta_2 IPK_i + v_i$$

Dengan menggunakan SPSS 20 diperoleh tabel Anova berikut:

Tabel 2. Anova

Sumber Variansi	JK	Df	R^2
Regression	14,909	2	0,071
Residual	194,946	82	
Total	209,854	84	

a. Dependent Variable: ujiBPG

b. Predictors: (Constant), X2, X1

Sehingga diperoleh χ^2_{hitung} dan χ^2_{tabel} sebagai berikut

$$\begin{aligned}\chi^2_{hitung} &= \frac{TSS \times R^2}{2} \\ &= \frac{209,854 \times 0,071}{2} \\ &= 7,45 \\ \chi^2_{tabel} &= \chi^2_{\alpha, db} \\ &= \chi^2_{0,05; 2} \\ &= 5,95\end{aligned}$$

Dapat disimpulkan bahwa hasil pengujian terhadap heteroskedastisitas dengan uji BPG pada taraf α sebesar 0,05 diperoleh nilai $\chi^2_{hitung} > \chi^2_{tabel}$. Berdasarkan kriteria pengujian, maka keputusan yang diambil adalah tolak H_0 atau data mengalami heteroskedastisitas.

3.2 Estimasi Dengan WLS

Berdasarkan perhitungan dengan metode BPG diperoleh bahwa H_0 ditolak yang artinya terdapat masalah Heteroskedastisitas dalam model, sehingga diperlukan adanya perbaikan pada model agar tidak menyesatkan kesimpulan.

Persoalan heteroskedastisitas dapat ditangani dengan melakukan pembobotan suatu faktor yang tepat kemudian menggunakan metode OLS terhadap data yang telah diboboti. Pemilihan terhadap suatu faktor untuk pembobotan tergantung bagaimana sisaan berkorelasi dengan X atau Y, jika sisaan proporsional terhadap X_i maka model akan dibagi dengan $\sqrt{X_i}$, jika sisaan adalah proporsional dengan X_i^2 sehingga model akan dibagi dengan X_i , selain proporsional dengan X_i dan X_i^2 bisa juga diasumsikan bahwa pola varian sisaan adalah proporsional dengan $[E(Y_i)]^2$ sehingga dibagi dengan $E(Y_i)$. Namun dalam prakteknya tidak selalu dengan pembobotan $\frac{1}{\sqrt{X_i}}, \frac{1}{X_i}, \frac{1}{E(Y_i)}$ dapat mengatasi heteroskedastisitas

karena sesungguhnya pembobot yang diberikan bergantung pada pola sebaran sisaan terhadap variabel bebas maupun variabel terikat. Oleh karena itu, dalam penelitian ini faktor pembobot yang akan dianalisis adalah $\frac{1}{\sqrt{X_i}}, \frac{1}{X_i}, \frac{1}{E(Y_i)}$, dan $\frac{1}{\sigma_i}$ (residual kuadrat).

Berikut ini koefisien determinasi R^2 setelah dilakukan pembobotan, yaitu:

Tabel 3. Hasil Estimasi WLS

No.	Faktor Pembobot	Koefisien Determinasi R^2
1	$\frac{1}{\sqrt{PP_i}}$	0,84
2	$\frac{1}{\sqrt{IPK_i}}$	0,49
3	$\frac{1}{PP_i}$	0,78
4	$\frac{1}{IPK_i}$	0,07
5	$\frac{1}{E(Y_i)}$	0,67
6	$\frac{1}{\sigma_i}$	0,99

Berikut ini hasil uji heteroskedastisitas dengan metode BPG yang telah diboboti.

Tabel 4. Hasil Uji Heteroskedastisitas dengan BPG

No.	χ^2 Hitung	χ^2 Tabel	Kesimpulan
1	16,42	5,95	Heteroskedastisitas
2	28,01	5,95	Heteroskedastisitas
3	19,63	5,95	Heteroskedastisitas
4	12,31	5,95	Heteroskedastisitas
5	7,30	5,95	Heteroskedastisitas
6	0,18	5,95	Homoskedastisitas

Dapat disimpulkan bahwa pembobot $\frac{1}{\sigma_i}$ pada taraf α sebesar 0,05 dapat mengatasi heteroskedastisitas dengan diperoleh nilai $\chi^2_{hitung} < \chi^2_{tabel}$. Berdasarkan kriteria pengujian, maka keputusan yang diambil adalah terima H_0 atau data tidak mengalami heteroskedastisitas.

4. KESIMPULAN DAN SARAN

Berdasarkan hasil yang telah diperoleh maka dapat disimpulkan bahwa faktor pembobot yang dicobakan pada metode WLS dalam penelitian ini adalah $\frac{1}{\sqrt{X_1}}, \frac{1}{\sqrt{X_2}}, \frac{1}{X_1}, \frac{1}{X_2}, \frac{1}{E(Y_i)}$, dan $\frac{1}{\sigma_i}$ (residual kuadrat). Berdasarkan hasil uji BPG dengan menggunakan nilai α sebesar 0,05, sesuai kriteria pengujian diperoleh faktor pembobot yang tidak signifikan adalah $\frac{1}{\sigma_i}$ yang artinya varian dari *error* bersifat konstan atau homoskedastisitas dengan R^2 sebesar 99%, sedangkan faktor pembobot $\frac{1}{\sqrt{X_1}}, \frac{1}{\sqrt{X_2}}, \frac{1}{X_1}, \frac{1}{X_2}, \frac{1}{E(Y_i)}$ masih signifikan. Maka pada penelitian ini faktor pembobot yang tepat adalah $\frac{1}{\sigma_i}$.

Pada penelitian ini model yang digunakan adalah regresi linier berganda, jika ingin melakukan penelitian serupa, peneliti menyarankan menggunakan model nonlinier.

DAFTAR PUSTAKA

- [1] Greene, W.H. 2003. *Econometrics Analysis*. Fifth Edition. New Jersey: Prentice Hall.
- [2] Gujarati, Damodar. 1995. *Basic Econometrics*. Third Edition. Singapore: McGraw-Hill
- [3] Gujarati, Damodar. 2007. *Dasar-dasar Ekonometrika*. Edisi Ketiga. Diterjemahkan oleh Julius A. Mulyadi dan Yelvi Andri. Jakarta: Erlangga.
- [4] Neter, J. et al. 1997. *Model Linear Terapan Buku I dan II: Analisis Regresi Linear Sederhana dan Analisis Regresi Ganda*, diterjemahkan oleh Bambang Sumantri. Bogor: Jurusan Statistika FMIPA IPB.
- [5] Uthami, P. 2013. *Regresi Kuantil Median Untuk Mengatasi Heteroskedastisitas Pada Analisis Regresi*. *Skripsi*. Universitas Udayana: Bali.
- [6] Widarjono. 2005. *Analisis Statistika Multivariat Terapan*. UPP STIM YKPN. Yogyakarta.