

Tingkat Efisiensi Metode Regresi Robust dalam Menaksir Koefisien Garis Regresi Jika Ragam Galat Tidak Homogen

Harmi Sugiarti dan Andi Megawarni
e-mail: harmi@mail.ut.ac.id dan mega@mail.ut.ac.id

Abstract

This paper aims to compare the relative efficiency of weighted least square (WLS), ordinary least square (OLS) and robust regression method in regression coefficient estimation when the error term is not homogen.

The assumption of homogeneous error variance underlying the ordinary least square (OLS) is very important to get the best linear unbiased estimation of the regression coefficients. The investigation compares the methods in calculating efficiency of booth simulation and experimental data.

In conclusion, the WLS method is relatively more efficient than OLS and Robust Regression methods.

Key words: Efisiensi Relatif, OLS, WLS, Regresi Robust, Homogenitas,

PENDAHULUAN

Asumsi ragam galat homogen ($\sigma^2(\varepsilon_i) = \sigma^2$) diperlukan oleh metode kuadrat terkecil (ordinary least square, OLS) untuk mendapatkan penaksir parameter yang bersifat tak bias linear terbaik (best linear unbiased estimator, BLUE) dari model regresi $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_p X_{pi} + \varepsilon_i$, $i = 1, 2, \dots, n$. Y_i adalah nilai peubah respons pada pengamatan ke- i , X_i adalah nilai peubah bebas pada pengamatan ke- i dan $\beta_0, \beta_1, \dots, \beta_p$ adalah koefisien regresi yang tidak diketahui nilainya dan akan dicari nilai taksirannya.

Dalam hal asumsi ragam galat homogen tidak dipenuhi, salah satu metode alternatif yang dapat digunakan adalah metode kuadrat terkecil tertimbang (weighted least square, WLS). Metode ini menghasilkan lebar selang kepercayaan untuk koefisien garis regresi lebih sempit dibanding metode OLS (Sugiarti & Megawarni, 2001).

Penanganan kasus ketidakhomogenan ragam galat kadangkala diikuti munculnya penyimpangan asumsi lainnya, diantaranya adalah munculnya pengamatan pencilan (outlier) dalam data. Adanya pengamatan pencilan dalam data dapat mengakibatkan penaksir koefisien garis regresi yang diperoleh tidak baik. Namun demikian tindakan membuang (menolak) begitu saja suatu pengamatan pencilan bukanlah tindakan yang bijaksana, karena adakalanya pengamatan pencilan memberikan informasi yang cukup berarti.

Disamping alasan faktor kesulitan menentukan pembobot $w_i = \frac{1}{\sigma^2}$ karena besarnya σ^2 tidak diketahui, penggunaan metode WLS masih belum dapat mereduksi munculnya pengamatan pencilan. Untuk mengatasi kelemahan-kelemahan dari metode yang ada, baik metode OLS maupun metode WLS, perlu dicoba metode lain yang bersifat tidak sensitif terhadap pelanggaran asumsi-asumsi, yaitu metode regresi robust (robust regression). Metode regresi robust ini menghasilkan lebar selang kepercayaan untuk koefisien garis regresi lebih sempit dibanding metode OLS tetapi

lebih lebar dibanding metode WLS (Sugiarti & Megawarni, 2003). Untuk memberi rekomendasi metode mana yang tepat untuk digunakan menaksir koefisien garis regresi, perlu mengkaji tingkat efisiensi dari metode tersebut.

Efisiensi relatif dari dua penaksir adalah rasio dari ukuran sampel yang diperlukan untuk mendapatkan keakuratan yang sama. Efisiensi tergantung pada distribusi data sesungguhnya, sehingga dalam praktek efisiensi yang sesungguhnya tidak diketahui. Jika $\hat{\beta}$ dan $\hat{\beta}^*$ adalah

penaksir untuk parameter β , maka efisiensi relatif $\hat{\beta}$ terhadap $\hat{\beta}^*$ adalah $eff(\hat{\beta}, \hat{\beta}^*) = \frac{mse(\hat{\beta}^*)}{mse(\hat{\beta})}$, dengan $mse(\hat{\beta}) = E(\hat{\beta} - \beta)^2 = \text{var}(\hat{\beta}) + [E(\hat{\beta}) - \beta]^2$ dan $mse(\hat{\beta}^*) =$

$E(\hat{\beta}^* - \beta)^2 = \text{var}(\hat{\beta}^*) + [E(\hat{\beta}^*) - \beta]^2$. Jika $\hat{\beta}$ dan $\hat{\beta}^*$ adalah penaksir tak bias untuk parameter β ,

maka efisiensi relatif $\hat{\beta}$ terhadap $\hat{\beta}^*$ menjadi $eff(\hat{\beta} / \hat{\beta}^*) = \frac{\text{var}(\hat{\beta}^*)}{\text{var}(\hat{\beta})} \times 100\%$. Penaksir $\hat{\beta}^*$ dikatakan lebih efisien dibanding $\hat{\beta}$ jika $\text{var}(\hat{\beta}^*) < \text{var}(\hat{\beta})$ (Fleming, 2003).

Penelitian ini bertujuan untuk membandingkan tingkat efisiensi antara metode OLS, metode WLS, dan metode regresi robust dalam menaksir koefisien garis regresi, jika ragam galat tidak homogen.

METODOLOGI

Data yang dipergunakan dalam penelitian ini adalah data simulasi, yaitu data yang dibangkitkan dengan bantuan program paket MINITAB versi 11.12 dan data eksperimen, yaitu berupa data rata-rata panjang daun (cm) tanaman temulawak (*Curcuma xanthorrhiza roxb.*) pada umur 17 minggu yang diberi pupuk kandang pada berbagai taraf yaitu: tanpa pupuk, 0.5 kg/lubang, 1 kg/lubang dan ditanam pada variasi jarak tanam 60 x 40 cm dan 60 x 60 cm (Priono, 1988).

Langkah pertama yang harus dilakukan adalah mencari galat (ε) yang memenuhi kriteria nilai rata-rata nol dan ragam tidak homogen. Galat tersebut dapat diperoleh dengan cara membangkitkan data. Pembangkitan galat dikelompokkan ke dalam tiga kelompok dengan setiap kelompok mempunyai rata-rata nol dan ragam berbeda-beda.

Langkah ke dua adalah menentukan nilai-nilai peubah bebas X_1 dan X_2 , karena X_1 dan X_2 adalah konstanta yang diketahui. Nilai-nilai dari koefisien regresi yaitu β diasumsikan dengan nilai tertentu. Dari nilai-nilai yang telah diketahui, dapat dicari nilai peubah respons sebagai $Y = X\beta + \varepsilon$.

Langkah ke tiga adalah menguji apakah dari pasangan data X_1 , X_2 dan Y diperoleh ragam galat yang tidak homogen. Apabila hasil uji menyatakan ragam galat tidak homogen, maka pasangan data X_1 , X_2 dan Y dengan rata-rata galat nol dan ragam galat tidak homogen telah diperoleh.

Langkah ke empat, setelah pasangan data X_1 , X_2 dan Y yang mempunyai rata-rata galat nol dan ragam galat tidak homogen telah diperoleh, selanjutnya dilakukan penaksiran koefisien regresi β dan variansi koefisien regresi β dengan metode OLS, metode WLS, dan metode regresi robust.

Langkah ke lima, menghitung dan membandingkan tingkat efisiensi metode WLS terhadap metode OLS, menghitung dan membandingkan tingkat efisiensi metode regresi robust terhadap metode OLS, serta menghitung dan membandingkan tingkat efisiensi metode regresi robust terhadap metode WLS.

Langkah ke enam adalah mengulang langkah ke empat dan ke lima untuk data eksperimen.

HASIL DAN PEMBAHASAN

Hasil Simulasi

Dengan menggunakan pasangan data X_1 , X_2 dan Y pada Lampiran, metode OLS, metode WLS, dan metode regresi robust memberikan nilai taksiran variansi untuk masing-masing koefisien regresinya seperti tampak pada Tabel 1 berikut.

Tabel 1. Penaksir Variansi Koefisien Regresi Data Simulasi

Koefisien Regresi	OLS	WLS	Robust
$\hat{\beta}_0$	2.424249	0.5909	2.97566
$\hat{\beta}_1$	0.000568	0.00013	0.000496
$\hat{\beta}_2$	0.00042	0.000101	0.000301

Nilai efisiensi relatif penaksir koefisien regresi yang diperoleh dengan metode WLS terhadap metode OLS untuk masing-masing koefisien adalah $eff(\hat{\beta}_0 / \hat{\beta}_0^*) = 410.26\%$, $eff(\hat{\beta}_1 / \hat{\beta}_1^*) = 436.96\%$, dan $eff(\hat{\beta}_2 / \hat{\beta}_2^*) = 415.67\%$.

Nilai efisiensi relatif penaksir koefisien regresi yang diperoleh dengan metode regresi robust terhadap metode OLS untuk masing-masing koefisien regresi adalah $eff(\hat{\beta}_0 / \hat{\beta}_0^{**}) = 81.47\%$, $eff(\hat{\beta}_1 / \hat{\beta}_1^{**}) = 114.60\%$, dan $eff(\hat{\beta}_2 / \hat{\beta}_2^{**}) = 139.31\%$.

Nilai efisiensi relatif penaksir koefisien regresi yang diperoleh dengan metode WLS terhadap metode regresi robust untuk masing-masing koefisien regresi adalah $eff(\hat{\beta}_0^{**} / \hat{\beta}_0^*) = 503.58\%$, $eff(\hat{\beta}_1^{**} / \hat{\beta}_1^*) = 381.28\%$, dan $eff(\hat{\beta}_2^{**} / \hat{\beta}_2^*) = 298.38\%$. Secara ringkas nilai efisiensi tersebut dapat dilihat pada Tabel 2 berikut.

Tabel 2. Nilai Efisiensi Relatif untuk Data Simulasi

Koefisien Regresi	$eff(\hat{\beta} / \hat{\beta}^*)$	$eff(\hat{\beta} / \hat{\beta}^{**})$	$eff(\hat{\beta}^{**} / \hat{\beta}^*)$
$\hat{\beta}_0$	410.264050	81.469301	503.581158
$\hat{\beta}_1$	436.956679	114.603469	381.277008
$\hat{\beta}_2$	415.672978	139.310700	298.378357

Hasil Eksperimen

Dengan menggunakan data eksperimen tentang rata-rata panjang daun tanaman temulawak berumur 17 minggu (Y) yang diberi pupuk kandang pada tiga taraf (X_1) dan ditanam pada dua variasi jarak tanam (X_2), metode OLS, metode WLS, dan metode regresi robust memberikan nilai taksiran variansi masing-masing koefisien regresinya seperti pada Tabel 3 dan Tabel 4 berikut.

Tabel 3. Penaksir Variansi Koefisien Regresi Data Eksperimen dengan Pembobot Konsentrasi Pupuk

Koefisien Regresi	OLS	WLS	Robust
$\hat{\beta}_0$	742.5625	18.41268	166.37388
$\hat{\beta}_1$	162.0529	10.80437	45.36965
$\hat{\beta}_2$	0.270088	0.006757	0.06027

Tabel 4. Penaksir Variansi Koefisien Regresi Data Eksperimen dengan Pembobot Jarak Tanam

Koefisien Regresi	OLS	WLS	Robust
$\hat{\beta}_0$	742.5625	439.3216	166.37388
$\hat{\beta}_1$	162.0529	6.630625	45.36965
$\hat{\beta}_2$	0.270088	0.270192	0.06027

Jika digunakan pembobot konsentrasi pupuk, nilai efisiensi relatif penaksir koefisien regresi yang diperoleh dengan metode WLS terhadap metode OLS untuk masing-masing koefisien adalah $eff(\hat{\beta}_0 / \hat{\beta}_0^*) = 4032.89\%$, $eff(\hat{\beta}_1 / \hat{\beta}_1^*) = 1499.88\%$, dan $eff(\hat{\beta}_2 / \hat{\beta}_2^*) = 3997.25\%$.

Jika digunakan pembobot jarak tanam, nilai efisiensi relatif penaksir koefisien regresi yang diperoleh dengan metode WLS terhadap metode OLS untuk masing-masing koefisien regresi adalah $eff(\hat{\beta}_0 / \hat{\beta}_0^*) = 169.02\%$, $eff(\hat{\beta}_1 / \hat{\beta}_1^*) = 2444.01\%$, dan $eff(\hat{\beta}_2 / \hat{\beta}_2^*) = 99.96\%$.

Nilai efisiensi relatif penaksir koefisien regresi yang diperoleh dengan metode regresi robust terhadap metode OLS untuk masing-masing koefisien regresi adalah $eff(\hat{\beta}_0 / \hat{\beta}_0^{**}) = 446.32\%$, $eff(\hat{\beta}_1 / \hat{\beta}_1^{**}) = 357.18\%$, dan $eff(\hat{\beta}_2 / \hat{\beta}_2^{**}) = 448.13\%$.

Jika digunakan pembobot konsentrasi pupuk, nilai efisiensi relatif untuk masing-masing penaksir koefisien regresi yang diperoleh dengan metode WLS terhadap metode regresi robust adalah $eff(\hat{\beta}_0^{**} / \hat{\beta}_0^*) = 903.58\%$, $eff(\hat{\beta}_1^{**} / \hat{\beta}_1^*) = 419.92\%$, dan $eff(\hat{\beta}_2^{**} / \hat{\beta}_2^*) = 891.99\%$.

Jika digunakan jarak tanam sebagai pembobot, nilai efisiensi relative penaksir koefisien regresi yang diperoleh dengan metode WLS terhadap metode regresi robust untuk masing-masing koefisien regresi adalah $eff(\hat{\beta}_0^{**} / \hat{\beta}_0^*) = 37.87\%$, $eff(\hat{\beta}_1^{**} / \hat{\beta}_1^*) = 684.24\%$, dan $eff(\hat{\beta}_2^{**} / \hat{\beta}_2^*) = 22.31\%$.

Secara ringkas nilai efisiensi tersebut dapat dilihat pada Tabel 5 dan Tabel 6 berikut.

Tabel 5. Nilai Efisiensi Relatif untuk Data Eksperimen dengan Pembobot Konsentrasi Pupuk

Koefisien Regresi	$eff(\hat{\beta} / \hat{\beta}^*)$	$eff(\hat{\beta} / \hat{\beta}^{**})$	$eff(\hat{\beta}^{**} / \hat{\beta}^*)$
$\hat{\beta}_0$	4032.886357	446.3215568	903.5831445
$\hat{\beta}_1$	1499.883057	357.1834563	419.9195204
$\hat{\beta}_2$	3997.254486	448.1283718	891.9887107

Tabel 6. Nilai Efisiensi Relatif untuk Data Eksperimen dengan Pembobot Jarak Tanam

Koefisien Regresi	$eff(\hat{\beta} / \hat{\beta}^*)$	$eff(\hat{\beta} / \hat{\beta}^{**})$	$eff(\hat{\beta}^{**} / \hat{\beta}^*)$
$\hat{\beta}_0$	169.0248101	446.3215568	37.87063553
$\hat{\beta}_1$	2444.00641	357.1834563	684.2440115
$\hat{\beta}_2$	99.96152736	448.1283718	22.30644915

KESIMPULAN

Secara keseluruhan dapat disimpulkan bahwa baik metode WLS maupun metode regresi robust sangat efisien dibanding metode OLS dalam menaksir koefisien garis regresi jika ragam galat tidak homogen.

Sedangkan jika dibandingkan tingkat efisien antara metode WLS dengan metode regresi robust, dapat disimpulkan bahwa metode WLS lebih efisien dibanding metode regresi robust dalam menaksir koefisien garis regresi jika ragam galat tidak homogen.

DAFTAR PUSTAKA

- Draper, N.R. & Smith, H. 1981. *Applied regression analysis*. 2nd ed. New York: Wiley.
- Fleming, Ron. 2003. *Basic statistical concepts*. Diperoleh dari <http://www.uky.edu/Ag/AgEcon>. 17 Oktober 2003
- Montgomery, D.C. & Peck, E.A. 1992. *Introduction to linear regression analysis*. 2nd ed. New York: Wiley.
- Myers, R.H. 1990. *Classical and modern regression with applications*. 2nd ed. Boston: PWS-Kent.
- Neter, J & Wasserman, W. 1990. *Applied linear statistical models*. 3rd ed. Homewood, Illinois: Irwin.
- Priono, M. 1988. Pengaruh pemberian pupuk kandang & jarak tanam terhadap pertumbuhan tanaman temulawak (*curcuma xanthorrhiza roxb.*). Skripsi yang tidak dipublikasikan. Universitas Jenderal Soedirman, Purwokerto.
- Staudte, R.G. & Sheather, S.J. 1990. *Robust estimation and testing*. New York: Wiley.

Sugiarti, H. & Megawarni, A. 2003. Penggunaan metode regresi Robust untuk mencari slang kepercayaan koefisien garis regresi jika ragam galat tidak homogen. Jakarta: Lembaga Penelitian Universitas Terbuka.

Sugiarti, H. & Megawarni, A. 2001. Selang kepercayaan untuk koefisien garis regresi jika ragam galat tidak homogen dengan metode OLS dan WLS. Jakarta: Lembaga Penelitian Universitas Terbuka.

LAMPIRAN

Data Simulasi

Pengamatan	Y	X_1	X_2	ε
1.	100.972	48	53	-0.02759
2.	121.659	95	26	0.65870
3.	41.731	42	1	-1.26902
4.	109.832	62	46	1.83219
5.	55.494	34	22	-0.50565
6.	26.970	17	11	-1.03007
7.	111.646	23	90	-1.35382
8.	112.610	62	50	0.60968
9.	107.549	83	24	0.54867
10.	170.493	79	91	0.49285
11.	103.917	42	62	-0.08276
12.	32.976	18	13	1.97609
13.	121.193	89	36	-3.80707
14.	53.497	8	40	5.49658
15.	88.483	76	14	-1.51694
16.	34.910	2	36	-3.09022
17.	55.939	37	23	-4.06147
18.	95.829	5	89	1.82904
19.	50.646	21	28	1.64600
20.	138.479	48	89	1.47855
21.	170.862	83	88	-0.13793
22.	54.293	40	11	3.29348
23.	77.655	60	24	-6.34511
24.	67.161	30	28	9.16097
25.	134.472	37	100	-2.52823
26.	134.850	66	74	-5.15037
27.	144.231	69	82	-6.76911
28.	81.048	67	11	3.04840
29.	83.743	55	26	2.74334
30.	100.464	37	61	2.46426