

Berkenalan dengan Homoskedastisitas dan Heterokedastisitas

Wahyu Widhiarso

Fakultas Psikologi UGM, Yogyakarta
Bahan Perkuliahan. Tahun 2011

Menurut Foster dkk. (2006) ada tiga asumsi yang perlu dipenuhi untuk melakukan analisis regresi yaitu distribusi normal, linieritas hubungan dan homoskedastisitas (*homoscedastic*). Tulisan berikut ini akan menjelaskan pengertian dan prosedur memverifikasi asumsi homoskedastisitas. Karena bersifat pengenalan maka banyak dilakukan penyederhanaan dalam pemaparan konsepnya.

1 Sekilas Tentang Regresi

Sebelum membahas mengenai homoskedastisitas, saya ajak pembaca untuk merefresh pemahaman mengenai analisis regresi. Analisis regresi bertujuan untuk melihat seberapa besar peranan variabel bebas (prediktor) terhadap variabel tergantung (kriterium).

Contoh Kasus. Misalnya kita hendak meneliti peranan ketabahan individu. Dalam hal ini prediktor kita adalah ketabahan, sedangkan rasa syukur menjadi kriteriumnya. Misalnya hasil analisis mendapatkan persamaan regresi sebagai berikut :

$$Y = B_1X_1 + C$$

$$Y = 4X_1 + 2$$

$$\text{Tingkat Rasa Syukur} = 4 (\text{Ketabahan}) + 2$$

Gambar SPSS di bawah ini adalah contoh persamaan dari analisis regresi melalui SPSS. Dari tabel di bawah ini didapatkan persamaan regresi :

$$Y = 0.256 X_1 + 19.30$$

Coefficients^a

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constant)	19.306	1.657		11.654	.000
Ketabahan	.256	.051	.469	4.984	.000

a. Dependent Variable: Syukur

Persamaan Regresi. Persamaan regresi di atas menunjukkan bahwa koefisien prediksi (b) ketabahan adalah 4. Di sisi lain ada konstanta yang menjelaskan residu persamaan regresi. Residu adalah hal-hal yang dapat memprediksi tingkat rasa syukur selain ketabahan. Dari persamaan di atas kita mendapatkan bahwa meningkatnya 1 poin ketabahan akan diikuti dengan peningkatan 4 poin tingkat rasa syukur.

- Misalnya seorang individu memiliki ketabahan sebesar 1. Kita masukkan angka 1 ke dalam ketabahan $Y=4(1)+2$. Hasilnya, tingkat rasa syukur individu adalah 7.
- Bagaimana kalau ketabahannya adalah 4. Kita dapatkan persamaan regresi $Y=4(4)+2$. Tingkat syukur individu sama dengan 18.

Apakah Residu itu ? Dari persamaan di atas, bagaimana jika ketabahan seseorang adalah 0 (nol). Tinggal kita masukkan ke dalam persamaan $Y=4(0)+2 = 2$. Ternyata meskipun ketabahan seseorang nihil dia masih memiliki rasa syukur sebesar 2. Berarti ada variabel lain yang bisa memprediksi rasa syukur selain ketabahan, buktinya tanpa ketabahan orang masih bisa memiliki rasa syukur. Bisa jadi variabel tersebut adalah tingkat keimanan, pengalaman hidup atau lainnya. Inilah yang dinamakan dengan residu, yaitu variabel-variabel lain yang terlibat akan tetapi tidak termuat di dalam model.

Residu. Nilai kuantitatif yang tidak dapat dijelaskan oleh model/persamaan regresi

Contoh tabel di bawah ini merupakan data dan hasil analisis regresi melalui SPSS. PRE_1 adalah prediksi dari model sedangkan RES_1 adalah residu dari model. Kita bisa mengeluarkan nilai PRE_1 dan RES_1 melalui menu Regresi di SPSS.

	Tabah	Syukur	PRE_1	RES_1	var
1	32	23	28	-5	
2	30	24	27	-3	
3	32	29	28	1	
4	25	24	26	-2	
5	32	30	28	2	
6	40	31	30	1	
7	33	28	28	0	
8	37	30	29	1	
9	34	24	28	-4	
10	42	30	30	0	
11	32	27	28	-1	
12	23	26	25	1	

Gambar 1. Hasil Keluaran Analisis Melalui SPSS

Prediksi tingkat rasa syukur individu dengan menggunakan persamaan regresi tidak sepenuhnya akurat karena setiap prediksi menghasilkan residu. Misalnya subjek 1.

Rasa syukur subjek 1 tersebut adalah 23, tapi dengan memasukkan skor ketabahan sebesar 32 ke dalam persamaan $Y=0.256X_1 + 19.30$ kita memprediksi rasa syukur subjek 1 sebesar 28. Karena prediksinya meleset ($23 - 28$) residunya terlihat 5 poin.

2 Heteroskedastisitas

Pengertian. Residu adalah variabel tidak diketahui sehingga diasumsikan bersifat acak. Seperti contoh di atas, kita tidak tahu variabel selain ketabahan yang mempengaruhi rasa syukur. Karena diasumsikan acak, maka besarnya residu tidak terkait dengan besarnya nilai prediksi.

Misalnya dalam nilai prediksi rasa syukur 5 subjek dari 100 subjek dalam penelitian adalah 20. Maka nilai residu dari 5 subjek ini harus berbeda-beda. Demikian juga untuk subjek-subjek yang nilainya 21, residu subjek-subjek ini harus berbeda-beda, dan seterusnya. Tabel 1 berikut menunjukkan contoh data yang terkangkit heterokedastisitas dan yang tidak.

Tabel 1. Perbandingan Dua Data

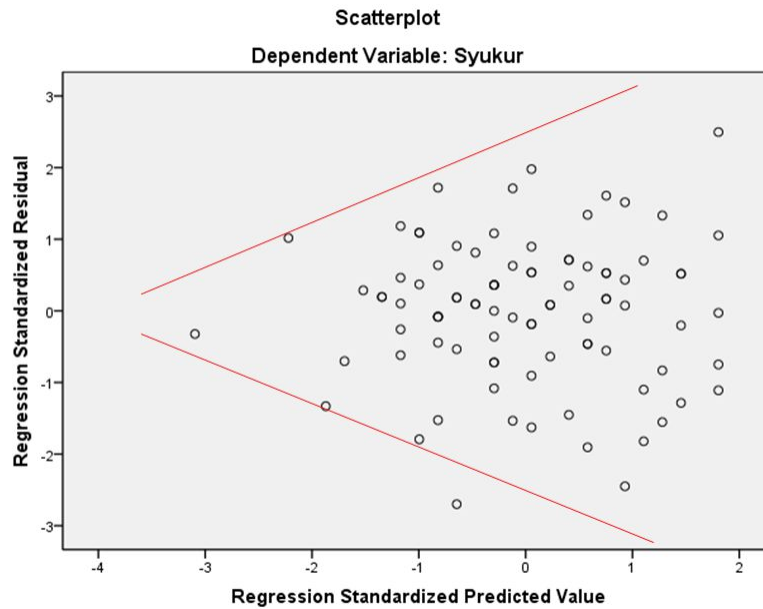
PREDIKSI	RESIDU	PREDIKSI	RESIDU	PREDIKSI	RESIDU
...	...	...	...	...	...
20	2	20	2	20	1
20	2	20	2	20	13
20	1	20	1	20	5
...	...	...	...	...	...
31	15	31	8	31	1
31	16	31	2	31	8
31	19	31	17	31	10
...	...	...	...	...	...
1(A)		1(B)		1(C)	
Heteroskedastisitas		Heteroskedastisitas		Homoskedastisitas	

Tabel 1A. Menunjukkan bahwa data terjangkit heteroskedastisitas karena porsi nilai residu pada setiap nilai prediksi tidak tersebar secara acak. Data membentuk pola : semakin besar nilai prediksi, nilai residunya semakin besar.

Tabel 1B. Menunjukkan bahwa data terjangkit heteroskedastisitas karena membentuk pola : semakin tinggi nilai prediksi, nilai residunya semakin bervariasi.

Tabel 1C. Menunjukkan bahwa data tidak terjangkit heteroskedastisitas karena porsi nilai residu pada setiap nilai prediksi secara acak. Pada nilai prediksi yang nilainya kecil (20), residunya bisa bernilai kecil atau besar. Demikian juga pada yang nilai prediksinya tinggi (31). Kesimpulannya, residu tersebar acak pada tiap nilai prediksi.

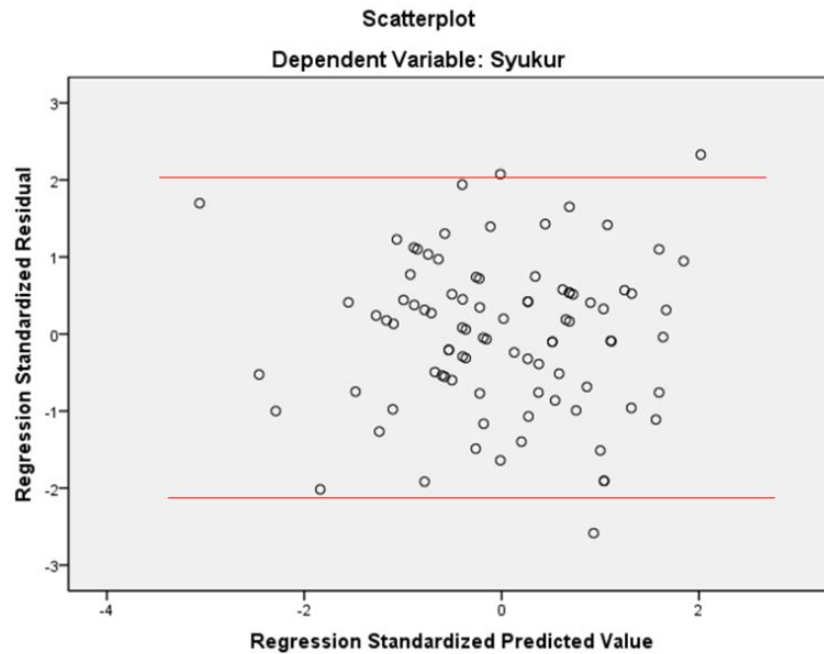
Gambar 2 dan 3 menunjukkan bagaimana data yang terjangkit heteroskedastisitas dan tidak.



Gambar 2. Hasil Keluaran Analisis Melalui SPSS

Gambar 2A. Gambar ini menunjukkan adanya heteroskedastisitas. Semakin besar nilai residu, eror semakin bervariasi.

Gambar 2B. Gambar ini menunjukkan tidak adanya heteroskedastisitas. Grafik menunjukkan bahwa pada semua setiap nilai prediksi, nilai residu memiliki variasi residu yang sama.



Gambar 3. Hasil Keluaran Analisis Melalui SPSS

Kita telah belajar bahwa heteroskedastisitas terjadi ketika hubungan antara prediksi dan residu membentuk sebuah pola. Di sisi lain homoskedastisitas adalah kondisi ketika nilai residu pada tiap nilai prediksi bervariasi dan variasinya cenderung konstan.

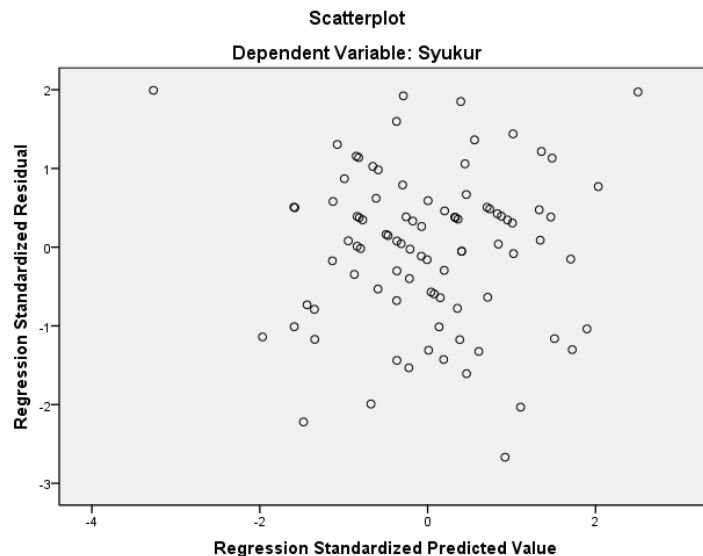
Mengapa Kok Heteroskedastisitas dihindari ? Kita tahu bahwa residu adalah variabel yang acak. Rasa syukur selain diprediksi oleh ketabahan, juga diprediksi oleh residu (variabel selain rasa syukur). Nah besarnya residu pada tiap orang harus bersifat acak.

Kalau antara nilai prediksi dan residu memiliki keterkaitan, berarti keduanya adalah variabel yang sama. Berarti antara ketabahan dan residu adalah hal yang sama. Tidak masuk akal bukan ? Nah kalau ini terjadi maka analisis regresi tidak dapat diterapkan.

3 Aplikasi pada SPSS

Pada SPSS cara memverifikasi homos/heteroskedastisitas masuk jadi satu dengan menu Regresi. Berikut ini prosedur baku memverifikasi heteroskedastisitas pada SPSS.

- Klik ANALYZE – LINEAR REGRESSION
- Masukkan Variabel Dependen dan Independen pada kotak yang tersedia
- Klik menu PLOTS. Masukkan ZPRED pada X dan ZRESID pada Y. OK



Gambar 4. Hasil Keluaran Analisis Melalui SPSS

Lihat pada plot yang dihasilkan dari analisis. Terlihat bahwa tidak ada pola tertentu yang menjelaskan hubungan antara nilai prediksi yang terstandarisasi (ZPRED) dan nilai residu terstandarisasi (ZRESID).

4 References

Foster, J., Barkus, E., & Yavorsky, C. (2006). *Understanding and using advanced statistics*. London: SAGE Publications Ltd.