

Seemingly Unrelated Regression (SUR) Penderita Penyakit DBD RS. Wahidin Sudirohusodo Dan RS. Stella Maris Makassar

A n i s a[†]

Abstrak

Hubungan antar variabel adalah salah satu hal yang selalu menarik dalam penerapan Statistika. Hubungan ini biasa dituliskan dalam suatu model/persamaan yang dikenal dengan model/persamaan regresi. Makalah ini difokuskan pada kajian mengenai hubungan antar variabel yang dilakukan secara simultan dalam suatu sistem persamaan yang menggambarkan ketergantungan di antara variabel-variabel dalam persamaan tersebut dengan menggunakan *Seemingly Unrelated Regression (SUR)*. Untuk penerapan, digunakan data penderita penyakit DBD dua rumah sakit terbesar yang ada di Kota Makassar, yaitu RS. Wahidin Sudirohusodo dan RS. Stella Maris.

Keywords: Hubungan antar variabel, sistem persamaan, *Seemingly Unrelated Regression (SUR)*.

1. Pendahuluan

Analisis regresi adalah analisis statistika yang memanfaatkan hubungan antara dua atau lebih peubah kuantitatif sehingga salah satu peubah dapat diramalkan dari peubah lainnya. Analisis tersebut merupakan salah satu alat statistik yang dapat memodelkan fenomena-fenomena suatu bidang ke dalam persamaan matematis. Analisis ini telah diterapkan secara luas ke berbagai bidang, misalnya pada bidang ekonomi, analisis ini digunakan untuk memodelkan faktor-faktor yang mempengaruhi permintaan sehingga dapat menduga fungsi permintaan, sedangkan pada bidang pertanian dan beberapa bidang lain, analisis regresi digunakan dalam menganalisis fungsi produksi. Sedangkan dalam bidang kesehatan, analisis ini banyak digunakan untuk mengetahui faktor-faktor yang mempengaruhi timbulnya suatu penyakit tertentu, dan bahkan dalam perkembangannya analisis ini bisa digunakan untuk mengetahui seberapa besar peluang seorang pasien untuk bisa bertahan hidup dari suatu penyakit tertentu.

Namun dengan semakin berkembangnya ilmu pengetahuan dewasa ini menuntut kajian hubungan antar variabel juga semakin kompleks, dimana model hubungan antar variabel tidak lagi hanya dalam suatu persamaan tunggal, namun dalam beberapa persamaan yang akan dianalisis secara simultan. Sebagai contoh, jika terdapat dua gugus pengamatan yang sama, maka kajian hubungan antar variabel untuk pengamatan tersebut tidak hanya bisa dilakukan secara terpisah, namun juga bisa dilakukan secara simultan dengan menggunakan analisis yang tepat (Setiawan, 1992; Manurung, 2005).

Pada model tunggal atau model yang hanya terdiri dari satu gugus pengamatan, pendugaan parameter model dilakukan dengan menggunakan *Ordinary Least Square (OLS)*. Namun ketika sebuah persamaan membentuk sistem persamaan yang terdiri dari beberapa persamaan maka metode OLS tidak dapat digunakan karena menghasilkan dugaan yang tidak efisien. Hal tersebut disebabkan metode OLS mengasumsikan persamaan-persamaan yang terdapat dalam sistem tersebut sebagai sebuah persamaan yang berdiri sendiri sehingga tidak dapat menjelaskan korelasi-korelasi yang terdapat antar persamaan tersebut. Oleh karena itu

[†] Staf Pengajar pada Jurusan Matematika FMIPA Universitas Hasanuddin Makassar

dibutuhkan metode yang lain yang bisa menjelaskan keterkaitan antar model pada gugus-gugus persamaan tersebut.

Salah satu model analisis yang bisa digunakan untuk kajian hubungan antar variabel secara simultan adalah dengan menggunakan *Seemingly Unrelated Regression* atau SUR (Setiawan, 1992). Model atau metode ini pertama kali dikemukakan oleh Arnold Zellner pada 1962. Metode SUR merupakan pengembangan dari metode *Generalized Least Square* (GLS), dimana GLS merupakan transformasi dari *Ordinary Least Square* atau OLS (Manurung, 2005).

Kajian SUR pada makalah ini diarahkan ke kajian terapan, dengan memperlihatkan hasil analisis SUR pada kasus penderita DBD dua rumah sakit terbesar yang ada di Kota Makassar, yaitu RS. Wahidin Sudirohusodo dan RS. Stella Maris. Untuk itu pembahasan pada makalah ini akan mengulas landasan teoritis metode SUR pada bagian 2, selanjutnya penerapan metode ini pada data yang digunakan dan pembahasannya pada bagian 3, dan terakhir kesimpulan yang diperoleh dari hasil tersebut pada bagian 4.

2. Model *Seemingly Unrelated Regression* (SUR)

Metode SUR yang dikemukakan pertama kali oleh Arnold Zellner pada tahun 1962 akan menghasilkan model SUR. Model ini merupakan model yang belum pernah dianalisis dalam literatur statistik, dimana variabel-variabel respon dapat dijelaskan oleh beberapa kumpulan variabel penjelas yang berbeda, dan dugaan statistik dan teknik pengujian yang dikembangkan Zellner berbeda dari model regresi klasik.

Model yang menunjukkan adanya galat pada persamaan struktural, dimana galat tersebut mengalami autokorelasi atau saling berkorelasi antara satu galat dengan galat persamaan lainnya adalah model *Seemingly Unrelated Regression* (SUR).

Model SUR terdiri dari M persamaan regresi, yang memenuhi asumsi dari model regresi standar dapat dituliskan dalam bentuk

$$y_i = X_i \beta_i + \varepsilon_i, \quad i = 1, \dots, M, \quad (1)$$

dimana $y_i (M \times 1)$, $X_i (M \times M)$, $\beta_i (M \times 1)$, $\varepsilon_i (M \times 1)$. Jika ε_{it} dimana t adalah elemen dari ε_i , dengan asumsi bahwa $(\varepsilon_{1t}, \varepsilon_{2t}, \dots, \varepsilon_{mt})$ berdistribusi bebas identik, dimana

$$\begin{aligned} - E(\varepsilon_{it}) &= 0 \\ - E(\varepsilon_{it}, \varepsilon_{js}) &= \sigma_{ij} \quad \text{jika } t = s \\ &= 0 \quad \text{jika } t \neq s \end{aligned}$$

maka persamaan dapat ditulis dalam bentuk matriks :

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix} = \begin{bmatrix} X_1 & 0 & \cdots & 0 \\ 0 & X_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & X_M \end{bmatrix} \times \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_M \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_M \end{bmatrix} \quad (2)$$

sehingga

$$y = X\beta + \varepsilon \quad (3)$$

dan

$$\begin{aligned} - E(\varepsilon) &= 0 \\ - E(\varepsilon, \varepsilon') &= V = \Sigma \otimes I_N \end{aligned}$$

dimana

$$V(y)_{(MN \times MN)} = \begin{pmatrix} \sigma_{11} I_N & \sigma_{12} I_N & \cdots & \sigma_{1M} I_N \\ \sigma_{21} I_N & \sigma_{22} I_N & \cdots & \sigma_{2M} I_N \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{M1} I_N & \cdots & \cdots & \sigma_{MM} I_N \end{pmatrix}$$

Matriks kovarian non-skalar ini merupakan gabungan tertentu dari matriks

$$\Sigma_{(M \times M)} \equiv \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1M} \\ \sigma_{21} & \sigma_{22} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{M1} & \cdots & \cdots & \sigma_{MM} \end{pmatrix}$$

dan matriks identitas ($N \times N$) I_N . Maka penaksir GLS dari β yaitu

$$\begin{aligned} \hat{\beta}_{GLS} &= (X'V^{-1}X)^{-1}X'V^{-1}y && \text{dimana } V = (\Sigma \otimes I_N) \\ &= (X'(\Sigma \otimes I_N)^{-1}X)^{-1}X'(\Sigma \otimes I_N)^{-1}y \\ &= (X'(\Sigma^{-1} \otimes I_N)X)^{-1}X'(\Sigma^{-1} \otimes I_N)y \\ &= \begin{pmatrix} \sigma_{11}(X_1'X_1) & \sigma_{12}(X_1'X_2) & \cdots & \sigma_{1M}(X_1'X_M) \\ \sigma_{21}(X_2'X_1) & \sigma_{22}(X_2'X_2) & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{M1}(X_M'X_1) & \cdots & \cdots & \sigma_{MM}(X_M'X_M) \end{pmatrix}^{-1} \begin{pmatrix} X_1'(\Sigma_j \sigma_{1j} y_j) \\ X_2'(\Sigma_j \sigma_{2j} y_j) \\ \cdots \\ X_M'(\Sigma_j \sigma_{Mj} y_j) \end{pmatrix} \quad (4) \end{aligned}$$

dimana σ_{ij} didefinisikan menjadi unsur pada baris ke- i dan kolom ke- j , dari Σ^{-1} , atau $\Sigma^{-1} \equiv [\sigma_{ij}]$. Persamaan (4) merupakan model penaksir GLS yang diperluas yang merupakan model penaksir SUR. Untuk efisiensi dalam penulisan, Greene (2000) menyatakan model penaksir SUR yang merupakan pengembangan dari metode GLS dapat ditulis dalam notasi :

$$\delta_{SUR} = (X'(\Sigma^{-1} \otimes I_N)X)^{-1}X'(\Sigma^{-1} \otimes I_N)y \quad (5)$$

3. Aplikasi Model *Seemingly Unrelated Regression* (SUR)

Sebelum pembahasan mengenai hasil analisis model SUR dengan menggunakan bantuan paket program SAS 9.1, maka akan diberikan gambaran sekilas mengenai data yang digunakan.

Data

Data yang digunakan pada penelitian ini merupakan data rekaman medik penderita penyakit Demam Berdarah (DBD) yang diperoleh dari RS. Wahidin Sudirohusodo dan RS. Stella Maris tahun 2006. Indikator atau peubah yang diamati dan di ukur, secara ringkas diberikan pada tabel berikut.

Tabel 1. Jenis Peubah/Data yang Diamati.

Pengamatan	Jenis Peubah/Data yang Diamati	Keterangan
Karakteristik Pasien	1. Nama 2. Jenis Kelamin 3. Umur	Informasi Umum dan hanya sebagian yang

	4. Pekerjaan 5. Tingkat pendidikan 6. Bobot Badan Relatif (%kg/cm) 7. Lama Menderita Penyakit 8. Lama Dirawat di Rumah Sakit 9. Status Kelangsungan Hidup	akan dimodelkan dengan SUR
--	--	----------------------------

Penerapan dan Pembahasan Metode SUR pada Data Penderita DBD RS. Wahidin Sudirohusodo dan RS. Stella Maris Makassar

Hasil pertama pengolahan data pasien penderita penyakit DBD dengan bantuan perangkat lunak SAS 9 diperoleh matriks kovarians yang dapat digunakan untuk melihat metode yang terbaik untuk digunakan dalam penelitian ini, atau membandingkan metode *Ordinary Least Square*(OLS) dengan metode *Seemingly Unrelated Regression* (SUR) dengan melakukan pengujian kovarians. Adapun matriks kovarian yang diperoleh adalah :

$$\begin{bmatrix} 206.534 & -20.90 \\ -20.90 & 5780.97 \end{bmatrix}$$

Dari hasil perhitungan dengan menggunakan persamaan $\lambda_{LR} = N \left(\sum_{i=1}^m \ln MSE_i - \ln |\hat{\Omega}| \right)$, diperoleh nilai $\lambda_{LR} = 2307.921$, dimana nilai ini bila dibandingkan dengan nilai tabel $\chi^2_{(m(m-1)/2, \alpha)}$ pada taraf $\alpha = 5\%$, maka dapat ditarik kesimpulan bahwa metode SUR lebih baik digunakan daripada OLS.

Selanjutnya, hasil analisis regresi dengan berbagai statistik yang bersesuaian yang diperoleh dengan metode OLS dan SUR, diberikan pada Tabel 1 berikut.

Dari tabel di bawah, terlihat bahwa dengan metode SUR hasil analisis regresi dan signifikansi yang diperoleh lebih baik, meskipun hasil analisis mengindikasikan variabel bebas yang berpengaruh adalah sama untuk kedua rumah sakit tersebut. Terlihat bahwa untuk penderita DBD RS. Wahidin Sudirohusodo, variabel yang berpengaruh adalah umur dan jumlah hemoglobin, sedangkan untuk RS. Stella Maris yang berpengaruh adalah lama rawat dan jenis penyakit DBD apakah merupakan DBD suspect, DBD stadium I sampai IV, baik dengan metode OLS maupun dengan metode SUR.

Tabel 1. Deskripsi Hasil Pengolahan Data dengan Menggunakan Perangkat Lunak SAS 9.

RS	Variabel Bebas	Dengan Metode OLS			Dengan Metode OLS		
		Dugaan Parameter	Statistik t	Signifikansi	Dugaan Parameter	Statistik t	Signifikansi
Wahidin Sudirohusodo	Jenis Kelamin	18.98639	1.61	0.1092	19.20961	1.63	0.1051
	Umur	2.522876	2.07	0.0402	2.521650	2.07	0.0403
	Berat Badan	0.110982	0.18	0.8539	0.109797	0.18	0.8554
	Tinggi Badan	0.148989	0.59	0.5587	0.149064	0.59	0.5585
	Lama Rawat	0.582689	0.25	0.8029	0.560310	0.24	0.8103

	Jenis Penyakit	5.057012	0.87	0.3845	5.026959	0.87	0.3873
	Status Penyakit	8.542861	1.06	0.2890	8.463029	1.05	0.2934
	Hemoglobin	-0.40930	-3.04	0.0029	-0.40440	-3.00	0.0032
Stella Maris	Jenis Kelamin	3.061324	1.32	0.1880	3.013509	1.30	0.1949
	Umur	-0.48404	-1.38	0.1697	-0.48862	-1.39	0.1657
	Berat Badan	0.220839	1.11	0.2691	0.218756	1.10	0.2736
	Tinggi Badan	0.047646	0.79	0.4317	0.048526	0.80	0.4232
	Lama Rawat	-1.36268	-2.00	0.0472	-1.34246	-1.97	0.0495
	Jenis Penyakit	5.289948	2.98	0.0034	5.263154	2.97	0.0035
	Status Penyakit	-2.01151	-1.03	0.3049	-2.01106	-1.03	0.3049
	Hemoglobin	-0.01062	-1.27	0.2055	-0.01031	-1.24	0.2188

Dari hasil di atas, model SUR yang diperoleh untuk data penderita DBD kedua rumah sakit, berdasarkan variabel-variabel yang mempengaruhi model dituliskan dalam bentuk berikut.

RS. Wahidin Sudirohusodo :

$$\delta_{SUR} = -2.522 \text{ Umur} - 3.000 \text{ Hemoglobin}$$

RS. Stella Maris :

$$\delta_{SUR} = -1.342 \text{ LamaRawat} + 5.263 \text{ JenisPenyakit}$$

Dari model SUR untuk kedua rumah sakit di atas, maka diharapkan jika seorang penderita DBD masuk untuk menjalani rawat inap ke rumah sakit tersebut, maka diharapkan jumlah trombosit penderita tersebut dapat diprediksi, sehingga dapat dilakukan langkah penanganan yang tepat untuk pasien tersebut.

4. Kesimpulan

Hasil penelitian kecil dengan metode SUR ini diharapkan bisa menjadi masukan bagi pihak manajemen kedua rumah sakit, bahkan untuk pihak Departemen Kesehatan pada khususnya, untuk meningkatkan pelayanan kepada masyarakat. Tentu saja hasil ini masih perlu proses validasi dan pengolahan kembali dengan data yang terbaru dan berkesinambungan untuk bisa dijadikan sebagai model acuan prediksi jumlah trombosit untuk penderita DBD yang menjalani rawat inap di rumah-rumah sakit yang ada di Kota Makassar.

Referensi

- [1] Greene, W.H., 2000, "Ilmu Econometric Analysis, 4th", Prentice Hall International Editions. New York.

- [2] Manurung, dkk., 2005, "*Ekonometrika Teori Dan Aplikasi*", Elex Media Komputindo. Jakarta.
- [3] Setiawan, 1992, "Kajian tentang *Seemingly Unrelated Regression* dan penerapannya pada model *Almost Ideal Demand System*", *Tesis*, Program Pasca Sarjana, Institut Pertanian Bogor.
- [4] Zellner, A., 1962, "An efficient methods of estimating *Seemingly Unrelated Regression* and tests for aggregation bias", *Econometrica*, Vol.38, hal.57-68.