

MODUL

REGRESI LINIER BERGANDA

Disusun oleh :
I MADE YULIARA



Jurusan Fisika
Fakultas Matematika Dan Ilmu Pengetahuan Alam
Universitas Udayana
Tahun 2016

Kata Pengantar

Puji syukur kami ucapkan ke hadapan Tuhan Yang Maha Kuasa atas rahmatNya modul ini dapat terselesaikan. Modul Regresi Linier Berganda ini merupakan bagian dari materi mata kuliah Statistika, FI29317 (3SKS) yang disusun untuk digunakan sebagai pedoman bagi mahasiswa FMIPA Fisika Unud yang mengambil mata kuliah Statistika pada semester genap tahun 2016.

Terimakasih kami ucapkan kepada rekan-rekan dosen Jurusan Fisika yang telah memberikan ide dan meluangkan banyak waktu dalam mendiskusikan modul ini. Modul ini tidaklah sempurna, untuk itu segala bentuk kritik dan saran yang konstruktif sangat diharapkan untuk memperbaiki modul ini.

Akhirnya kami ucapkan terimakasih semoga dapat menambah cakrawala ilmu pengetahuan dan bermanfaat bagi pembaca.

Maaret 2016

Penyusun,

I Made Yuliara

DAFTAR ISI

MODUL : Regresi Linier Berganda	Hal
Kata Pengantar	i
Daftar Isi	ii
1. Pendahuluan.....	1
2. Kegiatan Belajar 1 : Regresi Linier Berganda, Koefisien Korelasi, Koefisien Determinasi.....	2
3. Kegiatan Belajar 2 : Pengujian Hipotesis dan Koefisien Regresi, Uji-F...	6
4. Penutup.....	9
5. Daftar Pustaka.....	10

I. PENDAHULUAN

Penggunaan statistika dalam mengolah data penelitian akan mempengaruhi tingkat analisis dari suatu hasil penelitian. Penelitian dalam bidang ilmu pengetahuan alam secara umum yang menggunakan aspek penghitungan statistika, akan memperoleh data yang hampir mendekati benar jika memperhatikan analisis regresi yang dipergunakan. Untuk memprediksi dan mengukur besarnya pengaruh suatu variabel bebas (*independent/ predictor*) terhadap variabel tak bebas (*dependent/ response*) dapat digunakan uji regresi.

Analisis regresi merupakan kajian terhadap hubungan satu variabel sebagai variabel yang diterangkan (*the explained variabel*) dengan satu atau dua variabel yang menerangkan (*the explanatory*). Jika variabel bebas hanya satu, maka analisis regresi disebut regresi sederhana dan jika lebih dari satu, maka analisis regresi disebut regresi linear berganda. Disebut berganda karena pengaruh beberapa variabel bebas dikenakan kepada variabel tak bebas. Analisis regresi ini banyak digunakan dalam bidang penulisan karya ilmiah yang menyangkut perhitungan hasil akhir yang mana akan menentukan bahwa, berhasil atau tidaknya suatu penelitian yang sedang dilakukan dengan melihat kesimpulan dari hasil perhitungan analisis regresinya. Analisis perhitungan tidak hanya melibatkan satu analisis saja, tetapi menyangkut beberapa penghitungan statistika agar menunjang hasil analisis regresi, seperti uji-t, uji-F, penggunaan anova dan pendugaan hipotesis. Hasil analisis regresi berupa persamaan regresi yang merupakan fungsi prediksi suatu variabel dengan menggunakan variabel lain.

Dalam modul ini dibahas regresi linier berganda dan pengujian hipotesis menggunakan Uji-F.

II. KEGIATAN BELAJAR 1

Regresi Linier Berganda

Model regresi linier berganda merupakan suatu persamaan yang menggambarkan hubungan antara dua atau lebih variabel bebas/ *predictor* (X_1, X_2, \dots, X_n) dan satu variabel tak bebas/ *response* (Y). Tujuan dari analisis regresi linier berganda adalah untuk memprediksi nilai variabel tak bebas/ *response* (Y) jika nilai variabel-variabel bebas/ *predictor* (X_1, X_2, \dots, X_n) diketahui. Disamping itu juga untuk mengetahui arah hubungan antara variabel tak bebas dengan variabel-variabel bebas.

Persamaan regresi linier berganda secara matematik diekspresikan oleh :

$$Y = a + b_1X_1 + b_2X_2 + \dots + b_nX_n$$

yang mana :

Y = variable tak bebas (nilai yang akan diprediksi)

a = konstanta

b_1, b_2, \dots, b_n = koefisien regresi

X_1, X_2, \dots, X_n = variable bebas

Bila terdapat 2 variable bebas, yaitu X_1 dan X_2 , maka bentuk persamaan regresinya adalah :

$$Y = a + b_1X_1 + b_2X_2$$

Keadaan-keadaan bila nilai koefisien-koefisien regresi b_1 dan b_2 adalah :

- bernilai 0, maka tidak ada pengaruh X_1 dan X_2 terhadap Y
- bernilai negatif, maka terjadi hubungan yang berbalik arah antara variabel bebas X_1 dan X_2 dengan variabel tak bebas Y
- bernilai positif, maka terjadi hubungan yang searah antara variabel bebas X_1 dan X_2 dengan variabel tak bebas Y

Konstanta a dan koefisien-koefisien regresi b_1 dan b_2 dapat dihitung menggunakan rumus :

$$a = \frac{(\sum Y) - (b_1 \times \sum x_1) - (b_2 \times \sum x_2)}{n}$$

$$b_1 = \frac{[(\sum x_2^2 \times \sum x_1 y) - (\sum x_2 y \times \sum x_1 x_2)]}{[(\sum x_1^2 \times \sum x_2^2) - (\sum x_1 \times x_2)^2]}$$

$$b_2 = \frac{[(\sum x_1^2 \times \sum x_2 y) - (\sum x_1 y \times \sum x_1 x_2)]}{[(\sum x_1^2 \times \sum x_2^2) - (\sum x_1 \times x_2)^2]}$$

yang mana :

$$\begin{aligned}\sum x_1^2 &= \sum X_1^2 - \frac{(\sum X_1)^2}{n} \\ \sum x_2^2 &= \sum X_2^2 - \frac{(\sum X_2)^2}{n} \\ \sum y^2 &= \sum Y^2 - \frac{(\sum Y)^2}{n} \\ \sum x_1 y &= \sum X_1 Y - \frac{\sum X_1 \sum Y}{n} \\ \sum x_2 y &= \sum X_2 Y - \frac{\sum X_2 \sum Y}{n} \\ \sum x_1 x_2 &= \sum X_1 X_2 - \frac{\sum X_1 \sum X_2}{n}\end{aligned}$$

Metode alternatif, yaitu metode matriks (metode kuadrat terkecil) a , b_1 dan b_2 dapat ditentukan dengan menyusun suatu persamaan :

$$\begin{aligned}an + b_1 \sum X_1 + b_2 \sum X_2 &= \sum Y \\ a \sum X_1 + b_1 \sum X_1^2 + b_2 \sum X_1 X_2 &= \sum X_1 Y \\ a \sum X_2 + b_1 \sum X_2 X_1 + b_2 \sum X_2^2 &= \sum X_2 Y\end{aligned}$$

Matriks dengan 3 persamaan 3 variabel :

$$\begin{aligned}m_{11}a + m_{12}b_1 + m_{13}b_2 &= h_1 \\ m_{21}a + m_{22}b_1 + m_{23}b_2 &= h_2 \\ m_{31}a + m_{32}b_1 + m_{33}b_2 &= h_3\end{aligned}$$

$$\begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} a \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix}$$

$$a = \frac{\det M_1}{\det M} \quad b_1 = \frac{\det M_2}{\det M} \quad b_2 = \frac{\det M_3}{\det M}$$

$$M_1 = \begin{bmatrix} h_1 & m_{12} & m_{13} \\ h_2 & m_{22} & m_{23} \\ h_3 & m_{32} & m_{33} \end{bmatrix}$$

$$M_2 = \begin{bmatrix} m_{11} & h_1 & m_{13} \\ m_{21} & h_2 & m_{23} \\ m_{31} & h_3 & m_{33} \end{bmatrix}$$

$$M_3 = \begin{bmatrix} m_{11} & m_{12} & h_1 \\ m_{21} & m_{22} & h_2 \\ m_{31} & m_{32} & h_3 \end{bmatrix}$$

$$M = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix}$$

Contoh :

$$\left. \begin{array}{l} 2a + b_1 + 4b_2 = 16 \\ 3a + 2b_1 + b_2 = 10 \\ a + 3b_1 + 3b_2 = 16 \end{array} \right\} \rightarrow \begin{bmatrix} 2 & 1 & 4 \\ 3 & 2 & 1 \\ 1 & 3 & 3 \end{bmatrix} \begin{bmatrix} a \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 16 \\ 10 \\ 16 \end{bmatrix} \Rightarrow M = \begin{bmatrix} 2 & 1 & 4 \\ 3 & 2 & 1 \\ 1 & 3 & 3 \end{bmatrix}$$

$$M_1 = \begin{bmatrix} 16 & 1 & 4 \\ 10 & 2 & 1 \\ 16 & 3 & 3 \end{bmatrix} \quad M_2 = \begin{bmatrix} 2 & 16 & 4 \\ 3 & 10 & 1 \\ 1 & 16 & 3 \end{bmatrix} \quad M_3 = \begin{bmatrix} 2 & 1 & 16 \\ 3 & 2 & 10 \\ 1 & 3 & 16 \end{bmatrix}$$

Nilai a , b_1 dan b_2 diperoleh dari determinan, yaitu :

$$a = \frac{\det M_1}{\det M} = \frac{26}{26} = 1; \quad b_1 = \frac{\det M_2}{\det M} = \frac{52}{26} = 2; \quad b_2 = \frac{\det M_3}{\det M} = \frac{78}{26} = 3$$

Koefisien Determinasi (r^2)

- Koefisien determinasi digunakan untuk mengetahui prosentase pengaruh variable bebas X_1 dan X_2 terhadap variable tak bebas Y
- Besarnya r^2 dihitung dengan rumus :

$$r^2 = \frac{(b_1 \sum x_1 y) + (b_2 \sum x_2 y)}{\sum y^2}$$

- Jika $r^2 = 0$, maka variasi variable-variable bebas X_1 dan X_2 tidak sedikitpun dapat menjelaskan variasi variable tak bebas Y dalam model persamaan regresi
- Jika $r^2 = 1$, maka variasi variable-variable bebas X_1 dan X_2 dapat menjelaskan dengan **sempurna** variabel tak bebas Y dalam model persamaan regresi

Koefisien Korelasi Ganda (r)

- Koefisien korelasi ganda digunakan untuk mengetahui seberapa besar korelasi yang terjadi antara variable-variable X_1, X_2, \dots, X_n secara serentak/ simultan dengan variabel Y .

- Besarnya nilai koefisien korelasi ganda dapat dihitung dengan rumus :

$$r = \sqrt{r^2} = \sqrt{\frac{(b_1 \sum x_1 y) + (b_2 \sum x_2 y)}{\sum y^2}}$$

- Nilai r : $-1 \leq r \leq +1$. Semakin mendekati nilai $+1$ atau -1 , maka semakin kuat hubungan yang terjadi dan sebaliknya jika r mendekati 0 maka semakin lemah hubungan yang terjadi

Korelasi Parsial

Korelasi parsial merupakan suatu korelasi yang menjelaskan korelasi antara satu variabel dengan satu variabel dan variabel lainnya dianggap konstan. Terdapat 3 macam bentuk korelasi parsial, yaitu :

- 1) korelasi antara X_1 dengan X_2 yang mana Y dianggap konstan ($r_{12.Y}$)

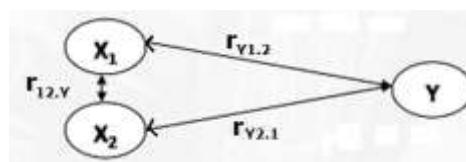
$$r_{12.Y} = \frac{r_{12} - (r_{Y1}r_{Y2})}{\sqrt{(1 - r_{Y1}^2)(1 - r_{Y2}^2)}}$$

- 2) korelasi antara Y dengan X_1 yang mana X_2 dianggap konstan ($r_{Y1.2}$)

$$r_{Y1.2} = \frac{r_{Y1} - (r_{Y2}r_{12})}{\sqrt{(1 - r_{Y2}^2)(1 - r_{12}^2)}}$$

- 3) korelasi antara Y dengan X_2 yang mana X_1 dianggap konstan ($r_{Y2.1}$)

$$r_{Y2.1} = \frac{r_{Y2} - (r_{Y1}r_{12})}{\sqrt{(1 - r_{Y1}^2)(1 - r_{12}^2)}}$$



yang mana

$$r_{Y1} = \frac{n \times \sum X_1 Y - (\sum Y \times \sum X_1)}{\sqrt{[(n \times \sum Y^2) - (\sum Y)^2] \times [(n \times \sum X_1^2) - (\sum X_1)^2]}}$$

$$r_{Y2} = \frac{n \times \sum X_2 Y - (\sum Y \times \sum X_2)}{\sqrt{[(n \times \sum Y^2) - (\sum Y)^2] \times [(n \times \sum X_2^2) - (\sum X_2)^2]}}$$

$$r_{12} = \frac{n \times \sum X_1 X_2 - (\sum X_1 \times \sum X_2)}{\sqrt{[(n \times \sum X_1^2) - (\sum X_1)^2] \times [(n \times \sum X_2^2) - (\sum X_2)^2]}}$$

Kesalahan Baku Estimasi (*Standart Error Estimate*)

Kesalahan baku estimasi digunakan untuk melihat apakah persamaan regresi yang terbentuk tepat/ kurang tepat dipakai untuk mengestimasi/ memprediksi variabel *response* Y. Jika kesalahan bakunya besar, maka persamaan regresi yang dibentuk kurang tepat dipakai untuk mengestimasi. Hal ini disebabkan karena selisih nilai antara variable *response* Y estimasi dengan Y kenyataan akan besar. Secara matematik kesalahan baku estimasi diekspresikan oleh :

$$S_e(S_{yx}) = \sqrt{\frac{\sum Y^2 - (a \sum Y) - (b_1 \sum X_1 Y) - (b_2 \sum X_2 Y)}{N - 3}}$$

III. KEGIATAN BELAJAR 2

Pengujian Hipotesis

Pengujian hipotesis dimaksudkan untuk menentukan apakah suatu hipotesis dapat diterima atau ditolak. Hipotesis merupakan asumsi atau pernyataan yang mungkin benar atau salah mengenai suatu populasi. Dengan mengamati seluruh populasi, maka suatu hipotesis akan dapat diketahui apakah suatu penelitian itu benar atau salah. Untuk keperluan praktis, pengambilan sampel secara acak dari populasi akan sangat membantu. Dalam pengujian hipotesis terdapat asumsi/ pernyataan istilah hipotesis nol. Hipotesis nol merupakan hipotesis yang akan diuji, dinyatakan oleh H_0 dan penolakan H_0 dimaknai dengan penerimaan hipotesis lainnya yang dinyatakan oleh H_1 .

Jika telah ditentukan Koefisien Determinasi (r^2), maka selanjutnya dilakukan pengujian hipotesis yang diajukan (misalkan dalam suatu penelitian) yang dapat menggunakan Uji-t ; Uji-F ; Uji-z atau Uji Chi Kuadrat. Uji ini dimaksudkan untuk mengetahui apakah variabel bebas/ *predictor*/ independen (X) berpengaruh secara signifikan terhadap variabel terikat/ *response*/ dependen (Y). Signifikan berarti bahwa pengaruh yang terjadi berlaku untuk seluruh populasi. Dalam modul ini hanya dibahas Uji-t.

Uji-F

Uji-F digunakan untuk mengetahui apakah variabel bebas X_1 dan X_2 secara bersama-sama berpengaruh secara signifikan terhadap variabel tak bebas Y.

Langkah-langkah yang dilakukan dalam Uji-F adalah:

1. Menentukan Hipotesis

$H_0 : \beta_1 = \beta_2 = 0$; (tidak ada pengaruh variable X_1 dan X_2 terhadap Y)

$H_1 : \beta_1 \neq \beta_2 \neq 0$; (ada pengaruh variabel X_1 dan X_2 terhadap Y)

2. Menentukan Tingkat/Taraf Signifikansi (α)

Nilai yang biasa digunakan adalah $\alpha = 5\%$

3. Menentukan F hitung

$$\text{Rumus F hitung : } F_{hit} = \frac{r^2/k}{(1-r^2)/(n-k-1)} = \frac{r^2(n-k-1)}{k(1-r^2)}$$

4. Menentukan F table (mempergunakan table Uji-F)

Tabel Uji-F untuk $\alpha = 5\%$ dengan derajat kebebasan pembilang (*Numerator*, df) = $k - 1$; dan untuk penyebut (*Denominator*, df) = $n - k$

n = jumlah sample/ pengukuran, k = jumlah variabel bebas dan terikat).

5. Kriteria Pengujian nilai F_{hit} dan t_{tab}

Bila nilai $F_{hit} < F_{tab}$, maka H_0 diterima, H_1 ditolak

Bila nilai $F_{hit} > F_{tab}$, maka H_0 ditolak, H_1 diterima

6. Kesimpulan : akan disimpulkan ada/tidaknya pengaruh antara variable-variabel bebas X_1 dan X_2 terhadap variable tak bebas Y .

Uji Koefisien Regresi secara Parsial (Uji-t)

Uji koefisien regresi secara parsial digunakan untuk mengetahui apakah model regresi yang terbentuk variable-variabel bebasnya (X_1 dan X_2) secara parsial berpengaruh signifikan terhadap variable tak bebas Y .

Langkah-langkah dalam melakukan Uji-t sama dengan pada regresi linear sederhana. (Lihat Modul Regresi Linier Sederhana)

Soal latihan :

Diberikan data tentang IQ dan tingkat kehadiran sepuluh siswa di kelas yang diperkirakan mempengaruhi nilai UAS.

Siswa	IQ (X_2)	Tingkat kehadiran (%) (X_1)	Nilai UAS (Y)
1	110	60	65
2	120	70	70
3	115	75	75
4	130	80	75
5	110	80	80
6	120	90	80
7	120	95	85
8	125	95	95
9	110	100	90
10	120	100	98

Pertanyaan :

1. Buatlah persamaan regresi linier berganda !
2. Variabel yang mana memberikan pengaruh lebih besar terhadap nilai UAS ?
Jelaskan mengapa demikian ?
3. Berapa koefisien determinasinya? Interpretasi hasil ini !
4. Lakukan Uji-F

Jawaban :

1. Persamaan regresi : $Y = 25.047 + 0.6705X_1 - 0.00343X_2$
2. Dilihat dari persamaan regresi, nilai b_1 lebih besar dibandingkan dengan nilai b_2 .
Nilai b_1 menandakan kemiringan X_1 (kehadiran dikelas) dan b_2 menandakan kemiringan X_2 (IQ). Dalam hal ini dapat disimpulkan bahwa presentase kehadiran dikelas lebih berpengaruh daripada IQ.
3. Koefisien Determinasi : $r^2 = (0.6935)^2 = 0.4809 = 48.09\%$
Nilai akhir (Y) yang dapat dijelaskan oleh tingkat kehadiran (X_1) dan IQ (X_2) pada persamaan regresi $Y = 25.047 + 0.6705X_1 - 0.00343X_2$ adalah 48.09%. Sisanya, sebesar 51.91% dijelaskan oleh faktor lain diluar variable-variabel pada persamaan regresi $Y = 25.047 + 0.6705X_1 - 0.00343X_2$.

Contoh pembacaan dan penjelasan mengenai :

R^2 , Uji-F, Uji-t parsial dan Persamaan regresi berganda dari hasil pengolahan data *software* statistik.

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	,729 ^a	,532	,477	7,4206	,532	9,658	2	17	,002	1,655

a. Predictors: (Constant), Nilai Bahasa, Nilai Matematika
b. Dependent Variable: Nilai Fisika

Dari tabel terlihat, r atau $R = 0,729$ dan $R^2 = 0,532$. Hal ini berarti bahwa kemampuan variabel bebas dalam menjelaskan varians variabel tak bebas adalah sebesar 53,2%. Masih terdapat 46,8% varians variabel tak bebas yang belum mampu dijelaskan oleh variabel bebas.

Anova (Uji-F; uji simultan) :

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1063,689	2	531,845	9,658	,002 ^a
	Residual	936,111	17	55,065		
	Total	1999,800	19			

a. Predictors: (Constant), Nilai Bahasa, Nilai Matematika
b. Dependent Variable: Nilai Fisika

Nilai F-hitung adalah 9,658 dengan taraf signifikan 0,002. Nilai signifikan ini di bawah 0,05 yang menunjukkan bahwa, variabel bebas secara serempak mempunyai pengaruh yang signifikan terhadap variabel tak bebas pada taraf signifikan 5%.

Uji-t Parsial

Uji-t (parsial) untuk melihat pengaruh variabel-variabel bebas secara parsial terhadap variabel tak bebasnya.

		Coefficients ^a						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95,0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	66,051	28,026		2,357	,031	6,921	125,180
	Nilai Matematika	,823	,203	,675	4,053	,001	,395	1,252
	Nilai Bahasa	-,664	,326	-,339	-2,037	,058	-1,351	,024

a. Dependent Variable: Nilai Fisika

Berdasarkan hasil pada tabel di atas, dapat disusun persamaan regresi linear berganda :

$$Y = 66,051 + 0,823 X_1 - 0,664 X_2$$

yang mana :

Y = hasil belajar fisika

X₁ = kemampuan matematika

X₂ = kemampuan berbahasa

IV. PENUTUP

Uji regresi linier berganda sangat membantu untuk mengetahui pengaruh secara serempak (simultan) baik kualitas maupun kuantitas dari variable-variabel bebas terhadap variable tak bebas. Hasil model persamaan regresi dapat dipergunakan sebagai pedoman untuk memprediksi hubungan antar variabel diluar data yang dijadikan sampel dalam suatu populasi.

V. DAFTAR PUSTAKA

- M. Nazir, 1983, *Metode Statistika dasar I* , Gramedia Pustaka Utama :Jakarta.
- Sudijono, Anas. 1996. *Pengantar Statistik Pendidikan*. Jakarta: Rajawali
- Spiegel. Murray. R. 2004. *Statistika*. Jakarta : Erlangga
- Supranto. J. 2001. *Statistika Teori dan Aplikasi Edisi Keenam Jilid2*. Jakarta : Erlangga
- Walpole. Ronald.E. 1995. *Ilmu Peluang dan Statistika untuk Insinyur dan Ilmuawan*.
Bandung : ITB