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 saya ucapkan ke hadapan Tuhan Yang Maha Kuasa karena atas berkat 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, maka dari itu, untuk itu untuk memperbaiki modul ini semua bentuk kritik maupun saran yang konstruktif sangat kami harapkan.

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

Maret 2016

Penyusun,

I Made Yuliara

DAFTAR ISI

I. PENDAHULUAN

Penggunaan statistika dalam mengolah data penelitian berpengaruh terhadap tingkat analisis hasil penelitian. Penelitian-penelitian dalam bidang ilmu pengetahuan alam (IPA) yang menggunakan perhitungan-perhitungan statistika, akan menghasilkan data yang mendekati benar jika memperhatikan tata cara analisis data yang digunakan. Dalam memprediksi dan mengukur nilai dari pengaruh satu variabel (bebas/*independent/ predictor*) terhadap variabel lain (tak bebas/*dependent/response*) dapat digunakan uji regresi.

Analisis/uji regresi merupakan suatu kajian dari hubungan antara satu variabel, dengan satu atau lebih variabel. Apabila variabel bebasnya hanya satu, maka uji/analisis regresinya dikenal dengan regresi linier sederhana. Apabila variabel bebasnya lebih dari pada satu, maka uji/analisis regresinya dikenal dengan regresi linear berganda. Dikatakan linier berganda karena terdapat dua atau lebih variabel bebas yang mempengaruhi variabel tak bebas.

Perhitungan-perhitungan hasil akhir untuk penulisan karya ilmiah/penelitian banyak menggunakan analisis/uji regresi. Hasil perhitungan analisis/uji regresi akan dimuat dalam kesimpulan penelitian dan akan menentukan apakah penelitian yang sedang dilakukan berhasil atau tidak. Analisis perhitungan pada uji regresi menyangkut beberapa perhitungan statistika seperti uji signifikansi (uji-t, uji-F), anova dan penentuan hipotesis. Hasil dari analisis/ uji regresi berupa suatu persamaan regresi. Persamaan regresi ini merupakan suatu fungsi prediksi variabel yang mempengaruhi variabel lain.

Dalam modul ini dibahas regresi linier berganda. Pengujian signifikansi hipotesis akan menggunakan Uji-F.

II. KEGIATAN BELAJAR 1

Regresi Linier Berganda

Regresi linier berganda merupakan model persamaan yang menjelaskan hubungan satu variabel tak bebas/ *response* (Y) dengan dua atau lebih variabel bebas/ *predictor* (X1, X2,…Xn). Tujuan dari uji regresi linier berganda adalah untuk memprediksi nilai variable tak bebas/ *response* (Y) apabila nilai-nilai variabel bebasnya/ *predictor* (X_1, X_2,..., X_n) diketahui. Disamping itu juga untuk dapat mengetahui bagaimanakah arah hubungan variabel tak bebas dengan variabel - variabel bebasnya.

Persamaan regresi linier berganda secara matematik diekspresikan oleh :

$$Y = a + b_1 X_1 + b_2 X_2 + \dots + bn\ Xn$$

yang mana :

Y = variable tak bebas (nilai variabel yang akan diprediksi)

a = konstanta

$b_1, b_2, \dots, b_n$ = nilai 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 koefisien-koefisien regresi, yaitu b_1 dan b_2 mempunyai nilai :

- Nilai=0. Dalam hal ini variabel Y tidak dipengaruh oleh X_1 dan X_2

- Nilainya negative. Disini terjadi hubungan dengan arah terbalik antara variabel tak bebas Y dengan variabel-variabel X_1 dan X_2
- Nilainya positif. Disni terjadi hubungan yang searah antara variabel tak bebas Y dengan variabel bebas X_1 dan X_2

Koefisien-koefisien regresi b_1 dan b_2 serta konstanta a dapat dihitung dengan 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 :

$$\sum x_1^2 = \sum X_1^2 - \frac{\left(\sum X_1\right)^2}{n}$$

$$\sum x_2^2 = \sum X_2^2 - \frac{\left(\sum X_2\right)^2}{n}$$

$$\sum y^2 = \sum Y^2 - \frac{\left(\sum Y\right)^2}{n}$$

$$\sum x_1 y = \sum X_1 Y - \frac{\sum X_1 \sum Y}{n}$$

$$\sum x_2 y = \sum X_2 Y - \frac{\sum X_2 \sum Y}{n}$$

$$\sum x_1 x_2 = \sum X_1 X_2 - \frac{\sum X_1 \sum X_2}{n}$$

Metode alternatif, yaitu metode matriks (metode kuadrat terkecil) dapat digunakan untuk menentukan nilai *a*, b_1 dan b_2. Metode ini dilakukan dengan cara membuat dan menyusun suatu persamaan sebagi berikut :

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

Matriks dengan 3 persamaan dan 3 variabel :

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

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

$$a = \frac{\det M_1}{\det M}$$

$$b_1 = \frac{\det M_2}{\det M}$$

$$b_2 = \frac{\det M_3}{\det M}$$

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

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

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

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

Contoh :

$$\left.\begin{aligned} 2a + b_1 + 4b_2 &= 16 \\ 3a + 2b_1 + b_2 &= 10 \\ a + 3b_1 + 3b_2 &= 16 \end{aligned}\right\}$$

$$\begin{bmatrix} 2 & 1 & 4 \\ 3 & 2 & 1 \\ 1 & 3 & 3 \end{bmatrix}\begin{bmatrix} a \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 16 \\ 10 \\ 16 \end{bmatrix}$$

$$M = \begin{bmatrix} 2 & 1 & 4 \\ 3 & 2 & 1 \\ 1 & 3 & 3 \end{bmatrix}$$

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

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

$$M_3 = \begin{bmatrix} 2 & 1 & 16 \\ 3 & 2 & 10 \\ 1 & 3 & 16 \end{bmatrix}$$

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

$$a = \frac{\det M_1}{\det M} = \frac{26}{26} = 1$$

$$b_1 = \frac{\det M_2}{\det M} = \frac{52}{26} = 2$$

$$b_2 = \frac{\det M_3}{\det M} = \frac{78}{26} = 3$$

Koefisien Determinasi (r^2)

- Untuk mengetahui prosentase pengaruh variable-variable X_1 dan X_2 terhadap variable Y digunakan koefisien determinasi
- Besarnya r^2 dihitung dengan rumus :

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

- Apabila r^2 bernilai 0 , maka dalam model persamaan regresi yang terbentuk, variasi variable tak bebas Y tidak sedikitpun dapat dijelaskan oleh variasi variable-variable bebas X_1 dan X_2
- Apabila r^2 bernilai 1, maka dalam model persamaan regresi yang terbentuk, variable tak bebas Y secara **sempurna** dapat dijelaskan oleh variasi variable-variable bebas X_1 dan X_2.

Koefisien Korelasi Ganda (r)

- Untuk mengetahui seberapa besar korelasi secara serentak/ simultan antara variable-variable X_1, X_2, ..., X_n dengan variabel Y dapat digunakan koefisien korelasi ganda.
- 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 : $-1 \leq r \leq +1$.

  Apabila nilai r mendekati nilai +1 atau – 1, maka dapat dikatakan bhawa semakin kuatnya hubungan/korelasi yang terjadi. Sebaliknya, apabila nilai r mendekati 0, maka semakin lemahnya hubungan/korelasi yang terjadi.

Korelasi Parsial

Merupakan suatu korelasi yang menjelaskan korelasi antara 1 variable dengan 1 variable dan variable lainnya dianggap konstan. Terdapat 3 macam bentuk korelasi parsial, yaitu :

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

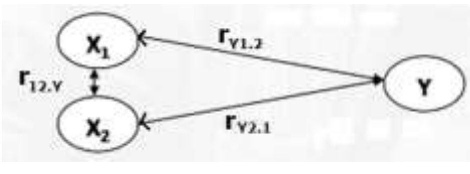
$$r_{12.Y} = \frac{r_{12} - (r_{Y1} r_{Y2})}{\sqrt{(1 - r_{Y1}^2)(1 - r_{Y2}^2)}}$$

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 melihat apakah suatu hipotesis yang diajukan ditolak atau dapat diterima 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/ hipotesis alternatif yang dinyatakan oleh H_1.

Jika telah ditentukan Koefisien Determinasi (r^2), maka selanjutnya dilakukan uji signifikan hipotesis yang diajukan. Uji ini dapat menggunakan Uji-t ; Uji-F ; Uji-z atau Uji Chi Kuadrat. Dengan uji signifikansi ini dapat diketahui apakah variable bebas/ *predictor*/ *independent* (X) berpengaruh secara signifikan terhadap variable tak bebas/ *response*/ *dependent* (Y). Arti dari signifikan adalah bahwa pengaruh antar varible berlaku bagi seluruh populasi. Dalam modul ini hanya dibahas uji signifikansi menggunakan Uji-F.

Uji - F

Penggunaan Uji-F bertujuan mengetahui apakah variabel-variabel bebas (X_1 dan X_2) secara signifikan bersama-sama berpengaruh terhadap variable tak bebas Y.

Tahapan yang dilakukan dalam Uji - F adalah:

1. Menentukan Hipotesis

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

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

2. Menentukan Taraf/tingkat Signifikansi (α)

   Nilai yang sering digunakan untuk adalah α = 5%

3. Menentukan F hitung

   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 α = 5% dengan derajat kebebasan pembilang (*Numerator*, df) = k - 1; dan untuk penyebut (*Denominator,* df) = n – k. n= jumlah sample/ pengukuran, k= jumlah variable bebas dan terikat).

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

   Apabila nilai $F_{hit} < F_{tab}$, maka hipotesis H_1 ditolak dan H_0 diterima.

   Apabila nilai $F_{hit} > F_{tab}$, maka hipotesis H_1 diterima dan H_0 ditolak.

6. Kesimpulan : akan disimpulkan apakah ada/ tidak pengaruh variable-variable bebas (X_1 dan X_2) terhadap variable tak bebas (Y).

Uji Koefisien Regresi Parsial (Uji-t)

Pengujian koefisien regresi secara parsial bertujuan mengetahui apakah persamaan model regresi yang terbentuk secara parsial variable-variable bebasnya (X_1 dan X_2) berpengaruh signifikan terhadap variable tak bebas (Y).

Tahapan 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 : $\mathbf{Y = 25.047 + 0.6705X_1 - 0.00343X_2}$

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 53,2% varians variabel tak bebas mampu dijelaskan oleh variabel bebas. Juga dapat dikatakan bahwa 46,8% variable bebas belum mampu menjelaskan varians variabel tak bebas.

<u>Anova (Uji-F; uji simultan) :</u>

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 lebih kecil dari 0,05 yang mengandung arti bahwa, secara serempak variable bebas berpengaruh signifikan terhadap variable tak bebas untuk taraf signifikan 5 %.

Uji - t Parsial

Uji-t parsial digunakan untuk mengetahui pengaruh variable-variable bebas terhadap variabel tak bebasnya secara parsial.

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 yang terdapat dalam table di atas, maka dapat dibentuk suatu persamaan regresi linear berganda, yaitu :

$$\mathbf{Y = 66{,}051 + 0{,}823X_1 - 0{,}664X_2}$$

yang mana :

Y = hasil pelajaran fisika

X_1 = nilai matematika

X_2 = nilai bahasa

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 , GramediaPustaka 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 Ke-6 Jilid 2. Jakarta : Erlangga

Walpole. R.,.E. 1995. Ilmu Peluang dan Statistika untuk Insinyur dan Ilmuawan. Bandung : ITB