

ANALISIS REGRESI LINIER MULTIVARIAT PADA KANDUNGAN DAUN TEMBAKAU

Fitriyatus Shofiyah

S1 Matematika, FMIPA, Universitas Negeri Surabaya
e-mail : fitriyatusshofiyah@mhs.unesa.ac.id

A'yunin Sofro

S1 Matematika, FMIPA, Universitas Negeri Surabaya
e-mail : ayuninsofro@unesa.ac.id

Abstrak

Dalam beberapa kasus regresi biasanya ditemukan variabel respon yang berjumlah lebih dari satu. Model regresi linear yang dapat menggambarkan hubungan lebih dari satu variabel respon dengan satu atau lebih variabel prediktor adalah regresi linier multivariat. Dalam penelitian ini, dilakukan analisis regresi linier multivariat untuk mengetahui hubungan komponen kimia yang terkandung dalam daun tembakau, yaitu apakah unsur kimia seperti Nitrogen, Fosfor, Klorin, Kalsium, Potassium, dan Magnesium berpengaruh terhadap banyaknya gula dalam tembakau, banyaknya nikotin dalam tembakau, serta berpengaruh dalam cepat lambat tingkat terbakarnya rokok. Kemudian dilakukan uji wilks lambda secara serentak dan parsial. Berdasarkan pengujian secara parsial diperoleh variabel prediktor yang berpengaruh terhadap banyaknya gula dalam daun tembakau, banyaknya nikotin dalam daun tembakau, serta berpengaruh dalam cepat lambat tingkat terbakarnya rokok adalah 3 variabel prediktor yaitu x_1 (Nitrogen (%)), x_2 (Klorin (%)), dan x_3 (Potassium (%)).

Kata Kunci : Regresi linier, Regresi linier multivariat, Uji Wilks Lambda.

Abstract

In some cases, regression usually found a response variable of more than one. A linear regression model that can illustrate the relationship of more than one response variable with one or more predictor variables is a multivariate linear regression. In this research, multivariate regression analysis is done to find out the correlation of chemical components contained in tobacco leaf, ie whether chemical elements such as Nitrogen, Phosphorus, Chlorine, Calcium, Potassium, and Magnesium affect the amount of sugar in tobacco, the amount of nicotine in tobacco, fast slow rate burning cigarette. Then simultaneous and partial Wilks lambda test. Based on partial test, it is found that predictor variables affecting the amount of sugar in tobacco leaf, the amount of nicotine in tobacco leaf, and the effect in the fast-slow rate of cigarette burn are 3 predictor variables, ie x_1 (Nitrogen in tobacco leaf (%)), x_2 (Chlorine in tobacco leaf (%)), and x_3 (Potassium in tobacco leaves (%)).

Keywords: Linear regression, Multivariate linear regression, Wilks Lambda Test.

PENDAHULUAN

Analisis regresi merupakan suatu teknik analisis statistik yang berguna untuk mengetahui hubungan antar variabel yang dinyatakan dalam bentuk model matematik (Ritonga dan Setiawan, 2011). Dalam analisis regresi, terdapat dua jenis variabel, yaitu variabel prediktor (variabel bebas) dan variabel respon (variabel terikat). Variabel prediktor adalah variabel yang mempengaruhi variabel respon, sedangkan variabel respon merupakan variabel yang dipengaruhi atau variabel yang menyebabkan adanya variabel prediktor. Untuk menghubungkan antara variabel respon dan prediktor dapat digunakan model regresi linear. Model regresi linear yang menjelaskan hubungan antara satu variabel respon dengan satu atau lebih variabel prediktor adalah model regresi linear univariat (Chatterjee dan Hadi, 2006:14).

Pada beberapa kasus biasanya variabel respon bisa berjumlah lebih dari satu. Model regresi linear yang dapat menggambarkan hubungan lebih dari satu variabel respon dengan satu atau lebih variabel prediktor adalah regresi linier multivariat (Johnson dan Wichern, 2007). Pada regresi linier multivariat diperlukan pengujian tingkat signifikansi variabel. Pengujian dapat dilakukan dengan menggunakan uji wilks lambda. Dengan melakukan uji wilks lambda dapat diketahui tingkat signifikansi variabel regresi terhadap model secara serentak dan secara parsial. Salah satu contoh aplikasi data menggunakan regresi linier multivariat adalah data tentang kandungan daun tembakau.

Tembakau merupakan tanaman yang sering dimanfaatkan untuk bahan baku pembuatan rokok di Indonesia dan juga di luar negeri. Dalam pembakaran rokok, dihasilkan asap

yang mengandung senyawa dan gas-gas tertentu. Nikotin yang terkandung dalam tembakau ketika terbakar dapat menghasilkan senyawa yang berbahaya. Selain itu Nikotin juga dapat mengakibatkan ketagihan, gangguan pada jantung dan paru-paru. Sehingga kadar nikotin didalam tembakau perlu dikendalikan, meskipun tidak sampai batas terendah.

Pembakaran rokok bisa terjadi cepat atau lambat. Semakin cepat rokok terbakar, maka semakin banyak pula asap berbahaya yang dihasilkan. Sehingga cepat lambatnya pembakaran rokok juga harus dibatasi. Selain nikotin, di dalam tembakau juga terdapat senyawa gula, rasa, dan lain-lain yang memenuhi selera konsumen (perokok). Kandungan gula dalam tembakau tidak berbahaya, sehingga perlu ditambah jumlahnya.

Usaha menekan bahan berbahaya dan tidak berbahaya dalam tembakau dapat dilakukan menggunakan bermacam-macam cara. Salah satunya adalah pada saat proses penanaman tembakau. Tembakau ditumbuh kembangkan pada lahan dengan tingkat kesuburan berbeda berdasarkan indikasi ketersediaan hara nitrogen, fosfor, klorin, kalsium, potassium dan magnesium pada lahan tersebut (Hidayati dan Djumali, 2011). Masing-masing unsur kimia tersebut memiliki pengaruh yang berbeda terhadap kandungan dalam daun tembakau. Sehingga pemberian unsur-unsur kimia itulah yang harus diperbanyak atau dikurangi.

Berdasarkan uraian diatas, pada penelitian ini akan dibahas tentang analisis regresi multivariat untuk mengetahui pengaruh pemberian unsur kimia seperti Nitrogen, Fosfor, Klorin, Kalsium, Potassium, dan Magnesium terhadap banyaknya kandungan gula dalam tembakau, banyaknya nikotin dalam tembakau, dan cepat lambat tingkat terbakarnya rokok

KAJIAN TEORI

Regresi Linier Multivariat

Analisis regresi linier multivariat adalah suatu analisis regresi linier yang memiliki variabel respon (y) berjumlah lebih dari satu dan berkorelasi dengan satu atau lebih variabel prediktor (x) (Johnson dan Wichern, 2007). Misalkan terdapat p variabel respon yaitu y_1, y_2, \dots, y_p dan q variabel prediktor yaitu x_1, x_2, \dots, x_q , maka model regresi linier multivariat dengan p respon adalah sebagai berikut:

$$y_1 = \beta_{01} + \beta_{11}x_1 + \beta_{21}x_2 + \dots + \beta_{q1}x_q + \varepsilon_1$$

$$y_2 = \beta_{02} + \beta_{12}x_1 + \beta_{22}x_2 + \dots + \beta_{q2}x_q + \varepsilon_2$$

$$y_p = \beta_{0p} + \beta_{1p}x_1 + \beta_{2p}x_2 + \dots + \beta_{qp}x_q + \varepsilon_p$$

Estimasi Kuadrat Terkecil Model Regresi Multivariat

Analisis regresi memiliki salah satu tujuan yaitu untuk memperoleh persamaan yang akan memprediksi nilai variabel respon melalui nilai dari variabel prediktor. Agar diperoleh model yang tepat pada persamaan regresi, maka terlebih dahulu harus ditentukan nilai dari koefisien regresi (β) yang cocok dengan data. Estimasi kuadrat terkecil untuk $\hat{\beta}$ diberikan dalam persamaan sebagai berikut :

$$\hat{\beta} = (X^T X)^{-1} X^T Y$$

Dimana

Y : Matriks dari variabel respon yang berukuran $n \times q$

X : Matriks dari variabel prediktor yang berukuran $n \times (p + 1)$

$\hat{\beta}$: Matriks dari parameter yang berukuran $(p + 1) \times q$

Uji Signifikansi Parameter

Ada dua uji yang dapat dilakukan untuk mengetahui tingkat signifikansi, yaitu parameter regresi signifikan terhadap model secara serentak dan parameter regresi signifikan terhadap model secara parsial.

a. Pengujian Serentak

Pengujian secara serentak dilakukan dengan tujuan untuk mengetahui apakah seluruh parameter signifikan terhadap model. Hipotesis yang digunakan dalam pengujian ini adalah sebagai berikut :

$$H_0 : \hat{\beta}_{11} = \hat{\beta}_{12} = \dots = \hat{\beta}_{p1} = \dots = \hat{\beta}_{pq} = 0 \quad (\text{model tidak signifikan})$$

$$H_1 : \text{paling sedikit ada satu } \hat{\beta}_{pq} \neq 0 \quad (\text{model signifikan})$$

Dimana q merupakan jumlah variabel respon dan p merupakan jumlah variabel prediktor

Statistik uji yang digunakan adalah Wilk's lamda:

$$\Lambda = \frac{|E|}{|E + H|} = \frac{|Y^T Y - \hat{B}^T X^T Y|}{|Y^T Y - n\bar{y}\bar{y}^T|}$$

Tolak H_0 jika $\Lambda \leq \Lambda_{\alpha, q, p, (n-p-1)}$. Nilai $\Lambda_{\alpha, q, p, (n-p-1)}$ merupakan nilai kritis dari tabel Wilk's Lamda.

b. Pengujian Parsial

Pengujian parsial digunakan dengan tujuan untuk mengetahui pengaruh dari setiap variabel prediktor terhadap variabel-variabel respon apakah signifikan atau tidak.

$H_0 : \hat{\beta}_{p1} = \hat{\beta}_{p2} = \dots = \hat{\beta}_{pq} = 0$ (parameter regresi prediktor p terhadap respon q tidak berpengaruh secara signifikan)

$H_1 :$ paling sedikit ada satu $\hat{\beta}_{pq} \neq 0$ (parameter regresi prediktor p terhadap respon q berpengaruh secara signifikan)

Statistik Uji menggunakan persamaan dan kriteria pengujian yang sama dengan uji serentak.

METODE

Data

Data yang digunakan dalam penelitian ini adalah data mengenai kandungan kimia dalam daun tembakao. Dalam penelitian ini digunakan sampel daun tembakao sebanyak 25 daun. Terdapat 3 variabel respon yaitu :

- y_1 : Tingkat rokok terbakar per inci dalam 1000 detik
- y_2 : Gula dalam daun tembakau (%)
- y_3 : Nikotin dalam daun tembakau (%)

Terdapat 6 variabel prediktor yaitu :

- x_1 : Nitrogen (%)
- x_2 : Klorin (%)
- x_3 : Potassium (%)
- x_4 : Fosfor (%)
- x_5 : Kalsium (%)
- x_6 : Magnesium (%)

Metode Analisis

Tahapan analisis data yaitu :

1. Menjelaskan deskripsi variabel respon dan variabel prediktor
2. Mengestimasi parameter model regresi multivariat
3. Melakukan uji signifikansi model regresi multivariat secara serentak dan parsial

HASIL DAN PEMBAHASAN

Statistik Deskriptif

Berikut ini adalah statistik deskriptif pada variabel respon. Variabel respon dalam penelitian ini yaitu tingkat rokok terbakar per inci dalam 1000 detik (y_1), Gula dalam daun tembakau (%) (y_2), dan Nikotin dalam daun tembakau (%) (y_3).

Tabel 1. Statistik deskriptif untuk variabel respon

Variabel	Mean	Min.	Max
y_1	1.688	1.400	2.090
y_2	16.62	12.58	20.05
y_3	2.161	1.240	3.280

Variabel prediktor dalam penelitian ini terdiri atas 6 variabel prediktor yaitu Nitrogen (%) (x_1), Klorin (%) (x_2), Potassium (%) (x_3), Fosfor (%) (x_4), Kalsium (%) (x_5), dan Magnesium (%) (x_6)

Berikut ini statistik deskriptif untuk variabel prediktor

Tabel 2. Statistik deskriptif untuk variabel prediktor

Variabel	Mean	Min.	Max.
x_1	2.157	1.720	2.870
x_2	2.481	0.740	3.300
x_3	2.24	1.64	2.64
x_4	0.49	0.42	0.56
x_5	3.592	2.790	4.570
x_6	0.964	0.780	1.250

Hasil Estimasi Parameter

Berikut ini hasil estimasi parameter untuk masing-masing variabel respon

Tabel 3. Hasil estimasi parameter

Parameter	Untuk y_1	Untuk y_2	Untuk y_3
β_0	1.41114	13.63291	-1.56482
β_1	0.06197	-4.31867	0.55216
β_2	-0.16013	1.32629	-0.27856
β_3	0.29212	1.58995	0.21759
β_4	-0.65798	13.95265	-0.72311
β_5	0.17303	0.55259	0.32309
β_6	-0.42835	-3.50211	2.00486

Berdasarkan Tabel diatas didapatkan model regresi multivariat untuk variabel respon y_1 (Tingkat rokok terbakar per inci dalam 1000 detik), y_2 (gula dalam daun tembakau (%)), dan y_3 (nikotin dalam daun tembakau (%)) adalah sebagai berikut:

$$y_1 = 1.41114 + 0.06197 x_1 - 0.16013 x_2 + 0.29212 x_3 - 0.65798 x_4 + 0.17303 x_5 - 0.42835 x_6$$

$$y_2 = 13.63291 - 4.31867 x_1 + 1.32629 x_2 + 1.58995 x_3 + 13.95265 x_4 + 0.55259 x_5 - 3.50211 x_6$$

$$y_3 = -1.56482 + 0.55216 x_1 - 0.27856 x_2 + 0.21759 x_3 - 0.72311 x_4 + 0.32309 x_5 + 2.00486 x_6$$

Pengujian Signifikansi Serentak dan Parsial

Pengujian signifikansi model pada analisis regresi multivariat dilakukan pengujian secara serentak dan parsial.

Berikut ini pengujian secara serentak dengan menggunakan uji *Wilk's Lambda*.

$H_0 : \hat{\beta}_{11} = \hat{\beta}_{12} = \dots = \hat{\beta}_{p1} = \dots = \hat{\beta}_{pq} = 0$ (model tidak signifikan)

$H_1 : \text{paling sedikit ada satu } \hat{\beta}_{pq} \neq 0$ (model signifikan)

Hasil yang diperoleh yaitu $\Lambda_{hitung} = 0.057499$ dan $\Lambda_{tabel} = \Lambda_{0.05, 3, 6, 18}$ sebesar 0.215. Kriteria uji yaitu $\Lambda_{hitung} < \Lambda_{tabel}$ ($0.057499 < 0.215$), maka keputusan yang dipilih adalah menolak H_0 , sehingga dapat diambil kesimpulan bahwa model signifikan.

Pengujian secara parsial ditunjukkan dalam tabel berikut

Tabel 4. Hasil pengujian parsial

Variabel	Nilai Wilk's lamda	p-value
x_1	0.16846	1.996×10^{-6}
x_2	0.37994	0.0011793
x_3	0.3604	0.0007844
x_4	0.81985	0.3513629
x_5	0.84527	0.4284186
x_6	0.64447	0.0647650

Hasil dari pengujian parsial pada Tabel 4 memperlihatkan bahwa variabel prediktor x_1 , x_2 dan x_3 memiliki pengaruh yang signifikan terhadap variabel respon y_1 , y_2 dan y_3 karena didapatkan nilai *p-value* untuk masing-masing variabel kurang dari α (0.05). Sedangkan *p-value* untuk variabel prediktor x_4 , x_5 , x_6 lebih dari α , dengan nilai $\alpha = 0.05$, sehingga variabel prediktor x_4 , x_5 , x_6 tidak signifikan terhadap terhadap variabel respon y_1 , y_2 dan y_3 .

Hasil estimasi parameter untuk variabel yang signifikan adalah sebagai berikut :

Tabel 5. Hasil estimasi parameter yang signifikan

Parameter	Untuk y_1	Untuk y_2	Untuk y_3
β_0	1.20870	20.0904	1.0311
β_1	0.10791	-4.9764	1.4138
β_2	-0.15625	1.3257	-0.2014
β_3	0.28312	1.7721	-0.6338

Model Regresi Multivariat

Hasil yang diperoleh dari pengujian model dengan uji wilks secara parsial adalah terdapat 3 variabel prediktor yang memiliki pengaruh yang signifikan terhadap variabel

respon, yaitu x_1 , x_2 dan x_3 , sehingga diperoleh model regresi multivariat sebagai berikut :

$$y_1 = 1.20870 + 0.10791 x_1 - 0.15625 x_2 + 0.28312 x_3$$

$$y_2 = 20.0904 - 4.9764 x_1 + 1.3257 x_2 + 1.7721 x_3$$

$$y_3 = 1.0311 + 1.4138 x_1 - 0.2014 x_2 - 0.6338 x_3$$

Interpretasi dari model diatas adalah sebagai berikut :

1. Untuk setiap penambahan satu persen nitrogen, maka akan menambah tingkat rokok terbakar per inci dalam 1000 detik sebesar 1.20870. Setiap penambahan satu persen klorin, maka akan mengurangi tingkat rokok terbakar per inci dalam 1000 detik sebesar 0.15625. Untuk setiap penambahan satu persen potassium, maka akan menambah tingkat rokok terbakar per inci dalam 1000 detik sebesar 0.28312
2. Untuk setiap penambahan satu persen nitrogen, maka akan mengurangi persentase gula dalam daun tembakau sebesar 4.9764. setiap penambahan satu persen klorin, maka akan menambah persentase gula dalam daun tembakau sebesar 1.3257. Untuk setiap penambahan satu persen potassium, maka akan menambah persentase gula dalam daun tembakau sebesar 1.7721
3. Untuk setiap penambahan satu persen nitrogen, maka akan menambah persentase nikotin dalam daun tembakau sebesar 1.0311. Setiap penambahan satu persen klorin, maka akan mengurangi persentase nikotin dalam daun tembakau sebanyak 0.2014. Untuk setiap penambahan satu persen potassium, maka akan mengurangi persentase nikotin dalam daun tembakau sebesar 0.6338

PENUTUP

Kesimpulan

Berdasarkan dari hasil uji wilks lamda, diperoleh 3 variabel prediktor yang memiliki pengaruh signifikan terhadap variabel respon. Variabel prediktor tersebut adalah x_1 (Nitrogen (%)), x_2 (Klorin (%)), dan x_3 (Potassium (%)). Model regresi linier multivariat yang terbentuk adalah sebagai berikut

$$y_1 = 1.20870 + 0.10791 x_1 - 0.15625 x_2 + 0.28312 x_3$$

$$y_2 = 20.0904 - 4.9764 x_1 + 1.3257 x_2 + 1.7721 x_3$$

$$y_3 = 1.0311 + 1.4138 x_1 - 0.2014 x_2 - 0.6338 x_3$$

DAFTAR PUSTAKA

Aswi dan Sukarna, 2006. *Analisis Deret Waktu*. Makasar : Penerbit Andira.

Chatterjee, S. and Hadi, A.S, 2006. *Regression Analysis by Example, Fourth Edition*. New York: A Wiley-Interscience Publication.

Draper, N dan Smith, H, 1992. *Analisis Regresi Terapan*. Jakarta : Gramedia.

Hidayati, S dan Djumali. 2011. *Produksi dan Kadar Nikotin Tembakau Temanggung pada Tiga Seri Tanah*. Balai Penelitian Tanaman Pemanis dan Serat Alam.

Johnson, R.A dan Wichern, D., 2007. *Applied Multivariate Statistical Analysis*. New Jersey : Prentice Hall.

Mansur, A, Syahputra, E., 2014. *Analisis Regresi Multivariat Pada Intention To Buy Berdasarkan Efektivitas Pemasaran*. Seminar Nasional IENACO. Pp 1-6.

Nurani, V., Sudarno, Rahmawati R., (2015). *Penerapan Regresi Linier Multivariat Pada Distribusi Ujian Nasional 2014 (Pada Studi Kasus Nilai Ujian Nasional 2014 Smp Negeri 1 Sayung)*. Jurnal Gaussian. Volume 4, Nomor 3. pp 697-704.

Setiawan, I. A. dan F. Ritonga, 2011. *Analisis Jalur (Path Analysis) dengan menggunakan Program AMOS. Edisi Pertama. Cetakan Pertama*. Tangerang : Penerbit Sulu Media.

