



ANALISIS REGRESI FAKTOR-FAKTOR YANG MEMPENGARUHI HASIL PRODUKSI PADI DI INDONESIA BAGIAN BARAT

Jeni Marianti Loban¹

Prodi Matematika FMIPA Universitas Tribuana Kalabahi

ABSTRACT

This study aims to measure the influence of factors affecting rice production in the western part of Indonesia. Data analysis used multiple linear regression analysis, using the independent variable X_1 = land area (Ha), X_2 = rainfall (mm) and X_3 = rainy days (days) on the dependent variable, namely the amount of rice production (Y). In this study, it was analyzed how much influence the land area (Ha), rainfall (mm), rainy days (days) had on the amount of rice production (tons). Data analysis using the help of RStudio software. The regression model obtained is $Y=61.6+0.000007917X_1-0.1541X_2+0.0002555X_3$ with a value of $R^2=0.9268$.

Keywords: *Multiple Regression Analysis, Rice Production*

PENDAHULUAN

Indonesia adalah negara agraris, secara geografis Negara Indonesia sangat diuntungkan karena letak Negara Indonesia terletak di garis katulistiwa dan secara teoritis Negara Indonesia sangat baik untuk pertanian dikarenakan seluruh wilayah yang ada di Negara Indonesia bisa terkena sinar matahari secara merata dan curah hujan yang merata. Sinar matahari dalam ilmu Biologi adalah bahan yang digunakan oleh tumbuhan untuk melakukan proses fotosintesis [1].

Posisi geografis Indonesia yang berada di wilayah tropis mempunyai

karakteristik iklim yang spesifik. Curah hujan merupakan unsur iklim yang paling tinggi keragaman dan fluktuasinya di Indonesia, sehingga merupakan unsur iklim yang paling dominan mencirikan Indonesia. Curah hujan di Indonesia bagian barat lebih besar daripada Indonesia bagian timur [2].

Daerah lumbung padi di Indonesia sebagian besar di Indonesia bagian barat yaitu Jawa Barat, Jawa Tengah, Jawa Timur, dan Sumatera Selatan. Hampir seluruh rakyat Indonesia mengkonsumsi beras sebagai makanan pokoknya. Oleh karena itu, sebagian

* Correspondence Address
E-mail: *yermialoban@yahoo.com

besar tanaman pangan yang ditanam di Indonesia adalah padi [3].

Dalam analisis regresi linier berganda membahas pola hubungan beberapa variabel yang ada dalam model, bagaimana pengaruh langsung dari variabel bebas (independent) terhadap variabel tidak bebas (dependent) [4]. Dalam penelitian ini dianalisa seberapa besar pengaruh luas lahan (Ha), curah hujan (mm), hari hujan (hari) terhadap jumlah produksi Padi (ton), sehingga dengan demikian dapat ditentukan faktor penyebab utama dan seberapa pengaruhnya

METODE PENELITIAN

Model yang digunakan adalah model regresi berganda, dengan menggunakan data sekunder. Yang menjadi variabel terikat (Y) yaitu jumlah produksi padi (ton) dan variabel bebas luas lahan (X_1), curah hujan (X_2), hari hujan (X_3). analisis diawali dengan melakukan uji asumsi klasik, selanjutnya melakukan uji regresi berganda. Bentuk umum persamaan regresi berganda yaitu:

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

Dimana Y : variabel terikat

a : Konstanta regresi

b : Koefisien regresi

X_1, X_2, \dots, X_n : Variabel bebas

PEMBAHASAN

1. Matriks Korelasi dan Linearitas

Matriks korelasi dipergunakan untuk melihat korelasi antara peubah baik respon dan prediktor. Matriks korelasi juga mampu menunjukkan apakah suatu hubungan antar peubah berkorelasi positif maupun negative [5]. Berikut matriks korelasi yang diperoleh dari data:

```
> cor(DataKP)
```

	Y	X1	X2	X3
Y	1.0000000	0.9074117	-0.9022860	0.9472407
X1	0.9074117	1.0000000	-0.8331805	0.8979495
X2	-0.9022860	-0.8331805	1.0000000	-0.8865410
X3	0.9472407	0.8979495	-0.8865410	1.0000000

Dari output diatas dapat diketahui bahwa hubungan linier antara setiap variabel prediktor (X_1, X_2 , dan X_3) dengan variabel respon (Y) terdapat hubungan positif dan negatif.

2. Uji Asumsi Klasik

Sebelum kita melakukan uji analisis regresi linear berganda, maka yang harus dilakukan adalah menguji data-data yang akan dianalisis agar data tersebut valid, tidak bias dan memenuhi persyaratan, maka digunakan uji asumsi klasik. Adapun penjelasan uji asumsi klasik adalah sebagai berikut.

Uji Normalitas

Pengujian normalitas dilakukan untuk menguji apakah dalam sebuah model penelitian, data berdistribusi normal atau tidak. Identifikasi berdistribusi normal atau tidaknya data dapat dilakukan dengan uji Shapiro untuk sampel kurang dari 30 [6].

Hipotesis yang diuji pada uji Shapiro adalah:

- H_0 : berdistribusi normal

- H_1 : tidak berdistribusi normal

Kriteria pengujian:

- Tolak H_0 jika nilai p-value < 0.05
- Terima H_0 jika nilai p-value > 0.05

Setelah diolah dengan menggunakan RStudio maka diperoleh hasil sebagai berikut:

```
> shapiro.test(rnorm(RegressionModel$residuals))
Shapiro-Wilk normality test
data:  rnorm(RegressionModel$residuals)
W = 0.94947, p-value = 0.5902
```

Berdasarkan hasil diatas, terlihat bahwa p-value = 0.5902 > 0.05 sehingga H_0 diterima, yang berarti residu data berdistribusi normal.

Uji Multikolinearitas

Uji multikolinearitas dilakukan untuk mengetahui apakah terdapat hubungan (korelasi) yang tinggi atau tidak antar variabel bebas yang terdapat dalam model regresi [7]. Jika terjadi korelasi, berarti terdapat masalah multikolinearitas. Hal ini dapat dilihat dari nilai VIF (*Variance Inflation Factor*).

dimana :

$$VIF = (1 - R^2)^{-1}$$

Hipotesis yang akan diuji adalah:

- H_0 : $r_{x_1x_2} = r_{x_1x_3} = \dots = r_{x_2x_3} = 0$, $i \neq j$ (tidak terjadi multikolinearitas antara variabel – variabel bebas).
- H_1 : $r_{x_ix_j} \neq 0$, $i \neq j$ (terjadi multikolinearitas antara variabel – variabel bebas).

Kriteria pengujian:

- Tolak H_0 jika nilai VIF > 10 .
- Terima H_0 jika nilai VIF < 10 .

Setelah diolah dengan menggunakan RStudio maka diperoleh hasil sebagai berikut:

```
> library(car)
> vif(RegressionModel)
      X1      X2      X3
5.340391 4.832453 7.629922
```

Berdasarkan hasil diatas, terlihat bahwa VIF variabel tersebut diatas adalah kurang dari 10 sehingga H_0 diterima, yang berarti tidak terjadi Multikolinearitas antar variabel-variabel bebas.

Uji Heterokedastisitas

Homoskedastisitas berarti bahwa ragam dari error bersifat konstan. Artinya pada nilai variabel bebas berapapun, variannya konstan. Jika variannya berbeda-beda atau bervariasi, berarti terjadi heteroskedastisitas. Heteroskedastisitas adalah suatu kondisi dimana sebaran atau variance dari data hasil pengamatan tidak konstan sepanjang observasi [8]. Jika harga makin besar maka sebaran makin lebar atau makin sempit. Untuk menguji heterokedastisitas dapat dilakukan dengan uji Glejser.

Hipotesis yang akan diuji adalah

- H_0 : $var(\epsilon_1) = var(\epsilon_2) = \dots = var(\epsilon_{13}) = \sigma^2$ (terjadi homoskedastisitas).
- H_1 : Ada $var(\epsilon_i) \neq \sigma^2$, $i = 1, 2, \dots, 13$ (terjadi heteroskedastisitas).

Kriteria pengujian:

- Tolak H_0 jika nilai p-value < 0.05
- Terima H_0 jika nilai p-value > 0.05 .

Setelah diolah dengan menggunakan RStudio maka diperoleh hasil sebagai berikut:

```
> library(zoo)
> library(lmtest, pos=15)
> bptest(Y ~ X1 + X2 + X3,
data=DataKP)

          studentized Breusch
-Pagan test

data:  Y ~ X1 + X2 + X3
```

Berdasarkan hasil diatas, terlihat bahwa p-value = 0.06177 > 0.05 sehingga H_0 diterima, yang berarti terjadi homoskedastisitas dengan kata lain tidak terjadi heteroskedastisitas.

Uji Autokorelasi

Uji autokorelasi bertujuan untuk menguji apakah dalam suatu model regresi linier ada korelasi antar error-nya. Cara untuk mendiagnosis adanya autokorelasi dalam suatu model regresi dilakukan melalui pengujian terhadap nilai Uji Durbin-Watson (Uji Dw)[9].

Hipotesis untuk uji Durbin-Watson:

- Jika $DW < dL$ atau $DW > 4 - dL$ maka H_0 ditolak, berarti terdapat autokorelasi.
- Jika DW terletak diantara dU dan $4 - dU$, maka H_0 diterima, berarti tidak terdapat autokorelasi.
- Jika DW terletak diantara dL dan dU atau diantara $4 - dU$ dan $4 - dL$, maka

tidak menghasilkan kesimpulan yang pasti.

Nilai dL (nilai batas bawah) dan dU (nilai batas atas) dapat dilihat pada tabel Durbin Watson. Dimana n adalah jumlah sampel yakni 13, k adalah jumlah variabel bebas yakni 3, dan α adalah taraf signifikansi yaitu sebesar 0.05.

Setelah diolah dengan menggunakan RStudio maka diperoleh hasil sebagai berikut:

```
> dwtest(Y~X1+X2+X3, alternative="two.sided", data=DataKP)

Durbin-Watson test

data:  Y ~ X1 + X2 + X3
DW = 0.73312, p-value = 0.0009941
alternative hypothesis: true autocorrelation is not 0
```

Berdasarkan hasil diatas, terlihat bahwa $DW = 0.73312$ dimana $0.7147(dL) < 0.73312(DW) < 1.8159(dU)$, berarti tidak menghasilkan kesimpulan yang pasti. Apabila terjadi kasus seperti diatas, maka autokorelasi dapat diuji dengan menggunakan Runs Test.

Hipotesis untuk uji Runs Test adalah

- H_0 : Terjadi gejala autokorelasi.
- H_1 : Tidak terjadi gejala autokorelasi.

Kriteria pengujian:

- Tolak H_0 jika nilai Asymp.Sig. (2-tailed) < 0.05
- Terima H_0 jika nilai Asymp.Sig. (2-tailed) > 0.05 .

Setelah menggunakan software SPSS 22, nilai asumsi signifikannya dapat diperoleh:

Runs Test

	Unstandardized Residual
Test Value ^a	-.12013
Cases < Test Value	6
Cases >= Test Value	7
Total Cases	13
Number of Runs	5
Z	-1.144
Asymp. Sig. (2-tailed)	.253

Berdasarkan hasil diatas, terlihat bahwa Asymp.Sig. (2-tailed) = 0.253 > 0.05 sehingga H_0 ditolak, yang berarti tidak terdapat gejala autokorelasi.

3. Estimasi Parameter dan Penduga Model Regresi

```
> RegressionModel <- lm(Y~X1+X2+X3, data=DataKP)
> RegressionModel

Call:
lm(formula = Y ~ X1 + X2 + X3, data = DataKP)

Coefficients:
(Intercept)      6.160e+01      7.917e-06     -1.541e-01      2.555e-05
```

Berdasarkan hasil output maka penduga model regresinya adalah sebagai berikut:

$$Y = 61.6 + 0.000007917X_1 - 0.1541X_2 + 0.0002555X_3$$

4. Uji Hipotesis

Uji F (Uji Simultan)

Uji F digunakan untuk melihat apakah variabel bebas secara simultan atau bersama-sama berpengaruh terhadap variabel terikat

Uji Hipotesis:

- $H_0 : \beta_1 = \beta_2 = \beta_3 = 0$ (Tidak ada hubungan linier).
- $H_1 : \beta_i \neq 0, i = 1,2,3$ (Setidaknya satu variabel prediktor mempengaruhi variabel respon).

Kriteria pengujian:

- H_0 diterima jika probabilitasnya (nilai signifikan) > 0.05
- H_1 diterima jika probabilitasnya (nilai signifikan) ≤ 0.05

```
> library(mixlm)
> anova_reg(RegressionModel)
Analysis of Variance Table
      Df Sum Sq Mean Sq F value Pr(>F)
Regression  3  51.565  17.1882  37.997 1.945e-05 ***
Residuals  9   4.071   0.4524
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Berdasarkan hasil yang didapatkan pada pengolahan terlihat bahwa p-value = 0.00001945 < 0.05 , sehingga H_0 ditolak artinya secara bersamaan variabel X_1, X_2, X_3 berpengaruh terhadap variable Y

Uji T (Uji Parsial)

Uji-t digunakan untuk memprediksi pengaruh masing-masing variabel prediktor terhadap variabel respon. Dalam pengujian ini dilakukan digunakan $\alpha = 5\%$.

Uji Hipotesis:

- $H_0 : \beta_i = 0, i = 1,2,3$
Luas lahan (Ha), curah hujan (mm), hari hujan (hari) secara parsial tidak berpengaruh signifikan terhadap variabel terhadap jumlah produksi Padi (ton)
- $H_1 : \beta_i \neq 0, i = 1,2,3$
Luas lahan (Ha), curah hujan (mm), hari hujan (hari) secara parsial berpengaruh signifikan terhadap variabel terhadap jumlah produksi Padi (ton)

Kriteria pengujian:

- H_0 diterima jika probabilitasnya (nilai signifikan) > 0.05
- H_1 diterima jika probabilitasnya (nilai signifikan) ≤ 0.05

```
> RegressionModel <- lm(Y~X1+X2+X3, data=DataKP)
> summary(RegressionModel)

Call:
lm(formula = Y ~ X1 + X2 + X3, data = DataKP)

Residuals:
    Min       1Q   Median       3Q      Max
-1.3664 -0.2119 -0.1201  0.4921  0.8027

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.160e+01  3.622e+00  17.005 3.78e-08 ***
X1           7.917e-06  6.716e-06   1.179  0.002
X2          -1.541e-01  1.225e-01  -1.259  0.240
X3           2.555e-04  1.259e-04   2.029  0.043 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
```

Berdasarkan output tersebut dapat diketahui sebagai berikut:

1. Uji t untuk variabel luas lahan (X_1) terhadap variabel jumlah produksi padi

(Y). Berdasarkan hasil diatas terlihat bahwa $Pr(> |t|) = 0.002 < 0.05$ sehingga H_1 diterima artinya secara parsial variabel luas lahan (X_1) berpengaruh terhadap variabel jumlah produksi padi (Y).

2. Uji t untuk variabel curah hujan (X_2) terhadap variabel jumlah produksi padi (Y). Berdasarkan hasil diatas terlihat bahwa $Pr(> |t|) = 0.240 > 0.05$ sehingga H_0 diterima artinya secara parsial variabel curah hujan tidak berpengaruh terhadap variabel jumlah produksi padi atau secara parsial tingginya curah hujan tidak berpengaruh terhadap jumlah produksi padi.

3. Uji t untuk variabel hari hujan (X_3) terhadap variabel jumlah produksi padi (Y). Berdasarkan hasil diatas terlihat bahwa $Pr(> |t|) = 0.043 < 0.05$ sehingga H_0 diterima artinya secara parsial jumlah hari hujan berpengaruh terhadap variabel jumlah produksi padi atau secara parsial tingginya/ banyaknya hari hujan berpengaruh terhadap jumlah produksi padi.

Koefisien Determinasi

Koefisien determinasi pada regresi linier sering diartikan sebagai seberapa besar kemampuan semua variabel prediktor dalam menjelaskan varians dari variabel responnya.

Jadi, koefisien determinasi menunjukkan pengaruh semua variabel independen terhadap nilai variabel dependen.

Residual standard error: 0.6726 on 9 degrees of freedom
 Multiple R-squared: 0.9268, Adjusted R-squared: 0.9024
 F-statistic: 38 on 3 and 9 DF, p-value: 1.945e-05

Berdasarkan hasil di atas karena pada regresi linier berganda untuk melihat baik tidaknya suatu model dilihat dari $R^2 - adjusted$ nya. Terlihat bahwa $R^2 - adjusted = 0.9024$ yang berarti 90.24% variabel prediktor dapat menjelaskan variabel respon sedangkan sisanya sebesar 9.76% dijelaskan oleh faktor lain di luar model.

Model Terbaik

Model terbaik ialah model yang sederhana dan mampu menjelaskan lebih baik dari pada model yang lain. Untuk menentukan model terbaik disini digunakan software Minitab 16. Adapun hasilnya sebagai berikut:

Response is y

Vars	R-Sq	R-Sq(adj)	Mallows Cp	S	x x
1	89.7	88.8	3.6	0.72084	
1	82.3	80.7	12.7	0.94511	X
2	91.6	89.9	3.4	0.68555	X
2	91.4	89.7	3.6	0.69194	X
3	92.7	90.2	4.0	0.67257	X X

Dimana kriteria yang digunakan pada metode Best Subset, yaitu:

- R^2 terbesar
- S residual terkecil

- Cp yang mendekati jumlah parameter

Pada hasil output diatas dapat disimpulkan bahwa model terbaiknya yaitu X1, X2 dan X3. Jadi persamaan regresinya yaitu:

```
> summary(RegressionModel)
Call:
lm(formula = Y ~ X1 + X2 + X3, data = DataKP)

Residuals:
    Min       1Q   Median       3Q      Max
-1.3664 -0.2119 -0.1201  0.4921  0.8027

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.160e+01  3.622e+00  17.005 3.78e-08 ***
X1           7.917e-06  6.716e-06   1.179  0.269
X2          -1.541e-01  1.225e-01  -1.259  0.240
X3           2.555e-04  1.259e-04   2.029  0.043 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6726 on 9 degrees of freedom
Multiple R-squared:  0.9268, Adjusted R-squared:  0.9024
F-statistic: 38 on 3 and 9 DF, p-value: 1.945e-05
```

Persamaan Regresi:

$$Y = 61.6 + 0.000007917X_1 - 0.1541X_2 + 0.0002555X_3$$

KESIMPULAN

Berdasarkan pembahasan dapat disimpulkan melalui uji F diketahui secara bersamaan (simultan) variable luas lahan (X_1), curah hujan (X_2) dan hari hujan (X_3) berpengaruh signifikan terhadap variable jumlah produksi padi (Y) dan secara parsial melalui uji T diketahui variabel luas lahan (X_1) berpengaruh terhadap variabel jumlah produksi padi (Y), variabel curah hujan (X_2) tidak berpengaruh terhadap variabel jumlah produksi padi (Y) dan variabel hari hujan (X_3)

berpengaruh terhadap variabel jumlah produksi padi (Y).

gresi-linier-berganda-1/

DAFTAR PUSTAKA

- [1] T. Wahyudie, *Pengelolaan Komoditas Hortikultura Unggulan Berbasis Lingkungan*. 2020.
- [2] S. B. Sipayung, L. Q. Avia, B. D. Dasanto, and Sutikno, "Analisis Pola Curah Hu]an Indonesia Berbasis Luaran Model Sirkulasi Global (Gcm)," *J. Sains Dirgant.*, vol. 4, no. 2, pp. 145–154, 2007, [Online]. Available: http://jurnal.lapan.go.id/index.php/jurnal_sains/article/viewFile/669/587
- [3] G. P. Dewi and A. M. Ginting, "Antisipasi Krisis Pangan Melalui Kebijakan Diversifikasi Pangan," *J. Ekon. Kebijakan. publik*, no. September 2011, pp. 67–75, 2012.
- [4] T. N. Padilah and R. I. Adam, "Analisis Regresi Linier Berganda Dalam Estimasi Produktivitas Tanaman Padi Di Kabupaten Karawang," *FIBONACCI J. Pendidik. Mat. dan Mat.*, vol. 5, no. 2, p. 117, 2019, doi: 10.24853/fbc.5.2.117-128.
- [5] I. M. Yuliara, "Regresi linier berganda 1.," *J. Artic.*, pp. 1–6, 2016, [Online]. Available: <http://www.mendeley.com/research/re>
- [6] W. Widana and P. L. Muliani, *Uji Persyaratan Analisis*. 2020.
- [7] E. Supriyadi, S. Mariani, and Sugiman, "Perbandingan Metode Partial Least Square (PLS) dan Principal Component Regression (PCR) Untuk Mengatasi Multikolinieritas Pada Model Regresi Linear Berganda.," *UNNES J. Math.*, vol. 6, no. 2, pp. 117–128, 2017, [Online]. Available: <http://journal.unnes.ac.id/sju/index.php/ujm>
- [8] L. P. Talangko, P. S. Statistika, U. Hasanuddin, M. K. Terkecil, and M. Eksponensial, "Estimasi Parameter dan Pengujian Breusch Pagan Godfrey Pada Regresi Eksponensial," pp. 1–9.
- [9] Anggito dan Setiawan, "Jenis Penelitian," *Angew. Chemie Int. Ed.* 6(11), 951–952., no. 2007, pp. 66–82, 2018.