

MODUL 8

ANALISIS KORELASI DAN REGRESI LINEAR BERGANDA 1



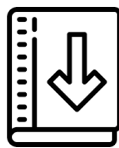
CAPAIAN PEMBELAJARAN

1. Praktikan mampu mengimplementasikan penggunaan software R untuk analisis korelasi dan regresi linear berganda



KEBUTUHAN ALAT/BAHAN/SOFTWARE

1. Komputer
2. Software R



DASAR TEORI

A. PENDAHULUAN

Regresi linear berganda (*multiple linear regression*) merupakan suatu teknik statistika yang menghasilkan suatu persamaan linear. Persamaan linear tersebut menerangkan atau menjelaskan hubungan antara variabel-variabel bebas ($X_1, X_1, X_1, \dots, X_n$) terhadap variabel tak bebas (Y).

Di samping itu, persamaan linear tersebut dapat digunakan untuk keperluan prediksi suatu nilai dari variabel tak bebas berdasarkan masukkan dari nilai-nilai variabel tak bebas.

Model persamaan regresi linear berganda, yaitu:

$$\hat{Y} = a + b_1X_1 + b_2X_2 + b_3X_3 + \dots + b_nX_n$$

Koefisien determinasi (r^2), adalah koefisien yang digunakan untuk mengetahui seberapa besar variabel bebas (X) mempengaruhi variabel terikat (Y). Nilai koefisien determinasi berkisar antara 0 sampai dengan 1.



PRAKTIK

Praktik 1 (mencari korelasi)

Sebuah toko olahraga “Spirit” menjual berbagai peralatan olahraga dan untuk mendapatkan profit yang maksimal toko tersebut gencar melakukan promosi dan membuka outlet di berbagai kota di Indonesia. Berikut ini adalah data mengenai penjualan, biaya promosi dan luas outlet yang berasal dari 6 kota di Indonesia

| Daerah | Keuntungan (dalam juta) | Promosi (dalam juta) | Outlet (dalam m ²) |
|----------|----------------------------|-------------------------|-----------------------------------|
| Jakarta | 30 | 6 | 40 |
| Surabaya | 20 | 5 | 40 |
| Medan | 35 | 7 | 60 |
| Bandung | 25 | 5 | 70 |
| Semarang | 45 | 8 | 55 |
| Yogya | 30 | 6 | 50 |

Analisislah data di atas!

Script

```

1 #input data
2 keuntungan <-c(30,20,35,25,45,30)
3 promosi <- c(6,5,7,5,8,6)
4 outlet <- c(40,40,60,70,55,50)
5
6 #membuat & menampilkan tabel
7 datakorelas13 <- data.frame(keuntungan,promosi,outlet)
8 datakorelas13
9
10 #menghitung nilai korelasi
11 cor(datakorelas13)

```

Gambar 1a

Output

```

> datakorelas13
  keuntungan promosi outlet
1         30         6     40
2         20         5     40
3         35         7     60
4         25         5     70
5         45         8     55
6         30         6     50
> cor(datakorelas13)
      keuntungan  promosi  outlet
keuntungan  1.0000000  0.9767088  0.222806
promosi      0.9767088  1.0000000  0.109423
outlet       0.2228060  0.1094230  1.000000
>

```

Gambar 1b

Analisis

Dari Gambar 1b, misalkan keuntungan = Y, sementara promosi = X_1 dan outlet = X_2 maka:

1. Korelasi promosi dan keuntungan (r_{1y}) sebesar 0,976, artinya positif dan kuat
2. Korelasi outlet dan keuntungan (r_{2y}) sebesar 0,223, artinya positif dan lemah
3. Korelasi promosi dan outlet (r_{12}) sebesar 0,109, artinya positif dan lemah

Praktik 2 (mencari persamaan regresi)

Aktifkan package `stats`, dengan maksud untuk menggunakan fungsi `lm`. Fungsi `lm` digunakan untuk membuat model regresi linear. Secara umum, perintah untuk menggunakan fungsi `lm` yaitu: `lm(formula, data)`

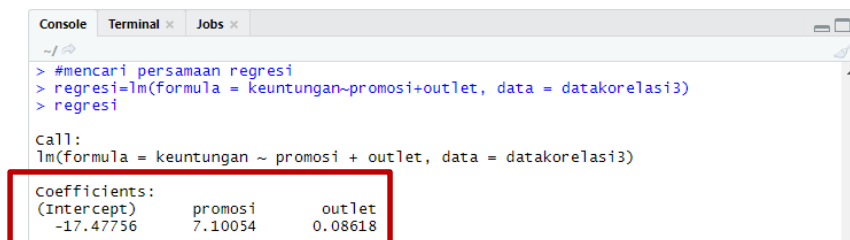
Script



```
13 #mencari persamaan regresi
14 regresi=lm(formula = keuntungan~promosi+outlet, data = datakorelasi3)
15 regresi
16
17 #menampilkan ringkasan regresi
18 summary(regresi)
```

Gambar 2a

Output dan Analisis



```
> #mencari persamaan regresi
> regresi=lm(formula = keuntungan~promosi+outlet, data = datakorelasi3)
> regresi

Call:
lm(formula = keuntungan ~ promosi + outlet, data = datakorelasi3)

Coefficients:
(Intercept)   promosi    outlet
-17.47756      7.10054      0.08618
```

Gambar 2b1

Pada Gambar 2b1 diperoleh persamaan regresi:

$$\hat{Y} = -17.4776 + 7.1005X_1 + 0.0862X_2$$

```
Console Terminal Jobs
~/
> summary(regresi)

Call:
lm(formula = keuntungan ~ promosi + outlet, data = datakorelasi3)

Residuals:
    1     2     3     4     5     6 
1.4273 -1.4722 -2.3968  0.9425  0.9336  0.5655 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -17.47756    5.94363   -2.941  0.0605 .
promosi      7.10054    0.77063    9.214  0.0027 **
outlet       0.08618    0.07683    1.122  0.3437
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.002 on 3 degrees of freedom
Multiple R-squared:  0.9676,    Adjusted R-squared:  0.9459 
F-statistic: 44.74 on 2 and 3 DF, p-value: 0.005842
```

Gambar 2b2

Pada Gambar 2b2 diperoleh nilai Multiple R-square (koefisien determinasi) adalah 0,9679 atau 96, 79%. Nilai ini menandakan bahwa: dari persamaan regresi dengan 2 variabel bebas yang diperoleh ($\hat{Y} = -17.4776 + 7.1005X_1 + 0.0862X_2$) mampu menjelaskan 96,79% informasi dari Y. Sementara informasi Y sisanya, yaitu 3,21% (100% - 96,79%) dijelaskan oleh variable yang lain.

Praktik 3 (uji signifikansi koefisien korelasi gabungan dan uji signifikansi koefisien korelasi parsial)

1. Koefisien korelasi gabungan adalah korelasi bersama-sama antara seluruh variable X dengan variabel Y. Korelasi gabungan dihitung dengan cara sebagai berikut:

$$r_{gab} = \sqrt{\text{koefisien determinasi}} = \sqrt{0,9679} = 0,9838 .$$

Dari nilai $r_{gab} = 0,9838$ selanjutnya akan dilihat apakah nilai tersebut signifikan di populasi atau tidak.

Langkah-langkah:

H_0 : Promosi dan luas outlet secara bersama-sama **tidak** berhubungan secara signifikan terhadap keuntungan.

H_1 : Promosi dan luas outlet secara bersama-sama berhubungan secara signifikan terhadap keuntungan.

Syarat: Jika $Prob > 0.05$ maka H_0 diterima. Jika $Prob < 0.05$ maka H_0 ditolak.

Nilai p-value = 0,005842 < 0.05 maka H_0 ditolak.

Kesimpulan: Promosi dan luas outlet secara bersama-sama berhubungan secara signifikan terhadap keuntungan.

2. Koefisien korelasi parsial adalah korelasi antara setiap variable X dengan variabel Y. Nilai korelasi parsial ini sudah dihitung pada Praktik 1.

$$r_{1y} = 0,976 \text{ dan } r_{2y} = 0,223$$

Dari nilai korelasi parsial yang diperoleh, selanjutnya akan dilihat apakah nilai tersebut signifikan di populasi atau tidak.

Langkah-langkah:

H_0 : promosi **tidak** berhubungan secara signifikan terhadap keuntungan.

H_1 : promosi berhubungan secara signifikan terhadap keuntungan.

Syarat: Jika Prob > 0.05 maka H_0 diterima. Jika Prob < 0.05 maka H_0 ditolak.

Nilai p-value = 0.0027 < 0.05 maka H_0 ditolak.

Kesimpulan: promosi berhubungan secara signifikan terhadap keuntungan.

H_0 : luas outlet **tidak** berhubungan secara signifikan terhadap keuntungan.

H_1 : luas outlet berhubungan secara signifikan terhadap keuntungan.

Syarat: Jika Prob > 0.05 maka H_0 diterima. Jika Prob < 0.05 maka H_0 ditolak.

Nilai p-value = 0.3437 > 0.05 maka H_0 diterima.

Kesimpulan: luas outlet **tidak** berhubungan secara signifikan terhadap keuntungan.

Dapat dilihat pada Gambar 2b2 terdapat tanda dua bintang pada baris **promosi**, berarti promosi berhubungan secara signifikan terhadap keuntungan. Sedangkan pada baris **luas outlet** tidak terdapat tanda bintang, berarti luas outlet **tidak** berhubungan secara signifikan terhadap keuntungan.



LATIHAN

Berikut ini adalah data keuntungan penjualan (Y) suatu produk yang dipengaruhi oleh jenis promosinya (iklan koran, iklan TV, iklan radio, dan jumlah outlet).

| No | Keuntungan Penjualan | Iklan Koran | Iklan TV | Iklan Radio | Jumlah Outlet |
|----|----------------------|-------------|----------|-------------|---------------|
| 1 | 215.36 | 20.98 | 27.90 | 13.23 | 7.00 |
| 2 | 295.15 | 22.41 | 32.28 | 13.44 | 5.00 |
| 3 | 254.26 | 22.98 | 29.49 | 15.26 | 10.00 |
| 4 | 452.62 | 23.21 | 39.17 | 18.45 | 5.00 |
| 5 | 330.92 | 23.25 | 34.25 | 19.58 | 8.00 |
| 6 | 320.14 | 23.45 | 33.63 | 12.03 | 8.00 |
| 7 | 254.25 | 24.86 | 29.38 | 13.87 | 6.00 |
| 8 | 235.26 | 24.88 | 29.19 | 15.69 | 9.00 |
| 9 | 302.21 | 25.00 | 32.82 | 16.35 | 9.00 |
| 10 | 312.25 | 25.12 | 33.44 | 12.88 | 8.00 |
| 11 | 222.32 | 25.87 | 29.14 | 18.97 | 8.00 |
| 12 | 265.99 | 25.89 | 32.09 | 12.05 | 11.00 |
| 13 | 300.12 | 26.23 | 32.33 | 12.23 | 7.00 |
| 14 | 265.21 | 26.23 | 30.22 | 15.87 | 5.00 |
| 15 | 354.25 | 26.25 | 35.42 | 13.67 | 6.00 |
| 16 | 323.45 | 28.94 | 33.72 | 18.29 | 9.00 |
| 17 | 362.02 | 29.80 | 35.84 | 15.26 | 8.00 |
| 18 | 423.00 | 32.26 | 37.12 | 13.56 | 5.00 |
| 19 | 400.23 | 32.79 | 36.10 | 18.78 | 9.00 |
| 20 | 412.60 | 33.45 | 36.85 | 13.02 | 6.00 |
| 21 | 423.22 | 33.98 | 37.44 | 16.59 | 7.00 |
| 22 | 400.25 | 34.55 | 36.15 | 14.23 | 9.00 |
| 23 | 366.25 | 34.76 | 35.92 | 15.26 | 9.00 |
| 24 | 435.23 | 35.99 | 38.20 | 15.78 | 8.00 |
| 25 | 430.22 | 36.21 | 37.91 | 13.33 | 10.00 |
| 26 | 352.16 | 36.25 | 34.79 | 12.89 | 9.00 |
| 27 | 365.21 | 36.87 | 35.91 | 12.45 | 8.00 |
| 28 | 415.25 | 36.99 | 36.96 | 19.25 | 8.00 |
| 29 | 451.29 | 40.12 | 38.98 | 14.32 | 8.00 |
| 30 | 512.33 | 44.98 | 39.33 | 13.45 | 8.00 |

1. Carilah koefisien korelasi antara keuntungan penjualan dengan masing-masing jenis promosinya!
2. Carilah persamaan regresi dari data tersebut!
3. Lakukan uji signifikansi koefisien korelasi gabungan dan parsial dari data tersebut!



TUGAS

Berikut ini adalah data berat badan (Y) dari sejumlah mahasiswa pada suatu kelas yang dipengaruhi oleh tinggi badan dan umur

| No. | Berat | Tinggi | Umur |
|-----|-------|--------|------|
| 1 | 64 | 57 | 8 |
| 2 | 71 | 59 | 10 |
| 3 | 53 | 49 | 6 |
| 4 | 67 | 62 | 11 |
| 5 | 55 | 51 | 8 |
| 6 | 58 | 50 | 7 |
| 7 | 77 | 55 | 10 |
| 8 | 57 | 48 | 9 |
| 9 | 56 | 52 | 10 |
| 10 | 51 | 42 | 6 |
| 11 | 76 | 61 | 12 |
| 12 | 68 | 57 | 9 |

1. Carilah koefisien korelasi antara keuntungan penjualan dengan masing-masing jenis promosinya!
2. Carilah persamaan regresi dari data tersebut!
3. Lakukan uji signifikansi koefisien korelasi gabungan dan parsial dari data tersebut!



REFERENSI

- [1] Gio, P.U., Effendie, A.R. 2017. Belajar Bahasan Pemrograman R (Dilengkapi Cara Membuat Aplikasi Olah Data Sederhana dengan R Shiny). Medan: USU Press.