Modul Praktikum: Regresi Linier dan Berganda

Berikut adalah modul praktikum untuk memahami regresi linier dan berganda.

Setiap soal diikuti langkah-langkah pengerjaan menggunakan perhitungan man- ual (dengan rumus awal) dan Microsoft Excel. Gunakan rumus dalam format persamaan untuk semua perhitungan, visualisasikan hasilnya, dan tentukan penerimaan atau penolakan hipotesis nol (*H*0) menggunakan uji t dan F.

# Soal 1: Regresi Linier

Seorang peneliti ingin mengetahui hubungan antara jumlah jam belajar per minggu

(*X*) dan nilai ujian matematika (*Y* ) pada 5 mahasiswa. Data yang diperoleh adalah:

Jam Belajar (*X*) Nilai Ujian (*Y* )

|  |  |
| --- | --- |
| 10 | 70 |
| 12 | 75 |
| 15 | 80 |
| 18 | 85 |
| 20 | 90 |

1. Tentukan hipotesis nol (*H*0) dan hipotesis alternatif (*H*1) untuk menguji hubungan antara jam belajar dan nilai ujian.
2. Hitung koefisien korelasi Pearson (*r*) menggunakan rumus awal dan Excel, lalu interpretasikan nilainya.
3. Tentukan persamaan regresi linier (*Y* = *a* + *bX*) menggunakan rumus awal dan Excel, jelaskan artinya, dan buat visualisasi scatter plot dengan garis regresi.

## Langkah-Langkah Pengerjaan Soal 1

### Hipotesis

* + *H*0: Tidak ada hubungan linier antara jam belajar dan nilai ujian (*ρ* = 0 atau

*β* = 0).

* + *H*1: Ada hubungan linier antara jam belajar dan nilai ujian (*ρ ̸*= 0 atau *β ̸*= 0).

### Koefisien Korelasi Pearson (*r*)

Rumus awal:

[*n* ∑ *X*2 *−* (∑ *X*)2][*n* ∑ *Y* 2 *−* (∑ *Y* )2]

*r* = √

*n* ∑ *XY −* ∑ *X* ∑ *Y*

(1)

Perhitungan:

- *n* = 5, *X* = 10 + 12 + 15 + 18 + 20 = 75, *Y* = 70 + 75 + 80 + 85 + 90 = 400.

∑ ∑ ∑

*Y* 2 = 702 + 752 + 802 + 852 + 902 = 32250.

- ∑ *XY* = (10 *·* 70) + (12 *·* 75) + (15 *·* 80) + (∑18 *·* 85) + (20 *·* 90) = 6130.

* + Substitusi:

-

*X*2 = 102 + 122 + 152 + 182 + 202 = 1193,

5 *·* 6130 *−* 75 *·* 400

√ √

30650 *−* 30000

650

*r* = [5 *·* 1193 *−* 752][5 *·* 32250 *−* 4002] = [5965 *−* 5625][161250 *−* 160000] = *√*340 *·* 1250 *≈* 0*.*997

### Langkah Excel:

1. Masukkan data:

- Kolom A (A2:A6): 10, 12, 15, 18, 20 (label A1: “Jam Belajar”).

- Kolom B (B2:B6): 70, 75, 80, 85, 90 (label B1: “Nilai Ujian”).

1. Pada sel kosong (misalnya, C1), ketik =CORREL(A2:A6,B2:B6) untuk menda- patkan *r* 0*.*997.

*≈*

**Interpretasi**: Nilai *r* = 0*.*997 menunjukkan korelasi positif sangat kuat. Se- makin banyak jam belajar, semakin tinggi nilai ujian.

### Persamaan Regresi Linier

Rumus awal:

∑ *−* ∑ ∑

Perhitungan:

*n XY X Y*

*b* = *n* ∑ *X*2 *−* (∑ *X*)2 *, a* =

∑ *Y − b* ∑ *X*

*n*

(2)

- *b* = 5*·*6130*−*75*·*400 = 30650*−*30000 = 650 *≈* 1*.*912.

5*·*1193*−*752

5965*−*5625

340

- *a* = 400*−*1*.*912*·*75 = 400*−*143*.*4 *≈* 51*.*32.

5

5

- Persamaan: *Y* = 51*.*32 + 1*.*912*X*.

### Langkah Excel:

1. Masukkan data seperti pada langkah (b).
2. Buka Data > Data Analysis > Regression (aktifkan Analysis ToolPak

di File > Options > Add-ins jika belum ada).

1. Pilih:
   * Input Y Range: B2:B6
   * Input X Range: A2:A6
   * Output Range: D2
2. Hasil: Intercept (*a* 51*.*32), X Variable 1 (*b* 1*.*912).

*≈ ≈*

### Visualisasi Scatter Plot:

1. Pilih data A2:B6, lalu Insert > Scatter.
2. Klik kanan pada titik data, pilih Add Trendline, pilih Linear, dan centang

Display Equation on chart.

1. Hasil: Scatter plot dengan garis regresi *Y* = 1*.*912*X* + 51*.*32.

### Interpretasi:

* + Konstanta 51.32: Jika tidak belajar (*X* = 0), nilai ujian diperkirakan 51.32.
  + Koefisien 1.912: Setiap tambahan 1 jam belajar meningkatkan nilai ujian sebe- sar 1.912 poin.

### Penerimaan/Penolakan *H*0 (Uji t Manual):

Untuk menguji apakah koefisien regresi *b* signifikan (artinya hubungan linier

ada), kita hitung statistik *t* secara manual:

* + \*\*Rumus statistik *t*\*\*:

*b*

*t* =

*SE*(*b*)

(3)

di mana *b* adalah koefisien regresi (1.912), dan *SE*(*b*) adalah standar error koe- fisien.

- \*\*Menghitung Standar Error (*SE*(*b*))\*\*: Standar error dihitung dari residual dan variansi:

√ ∑(*Yi − Y*ˆ*i*)2

(4)

Langkah-langkah:

*SE*(*b*) =

(*n −* 2) *·* ∑(*Xi − X*¯ )2

1. Hitung nilai prediksi *Y*ˆ*i* menggunakan *Y* = 51*.*32 + 1*.*912*X*:

- *X* = 10: *Y*ˆ

- *X* = 12: *Y*ˆ

- *X* = 15: *Y*ˆ

- *X* = 18: *Y*ˆ

- *X* = 20: *Y*ˆ

= 51*.*32 + 1*.*912 *·* 10 = 70*.*52.

= 51*.*32 + 1*.*912 *·* 12 = 74*.*16.

= 51*.*32 + 1*.*912 *·* 15 = 80*.*00.

= 51*.*32 + 1*.*912 *·* 18 = 85*.*68.

= 51*.*32 + 1*.*912 *·* 20 = 89*.*52.

1. Hitung residual (*Yi − Y*ˆ*i*):

- *Y*1 = 70, *Y*ˆ1 = 70*.*52, residual = 70 *−* 70*.*52 = *−*0*.*52.

- *Y*2 = 75, *Y*ˆ2 = 74*.*16, residual = 75 *−* 74*.*16 = 0*.*84.

- *Y*3 = 80, *Y*ˆ3 = 80*.*00, residual = 80 *−* 80*.*00 = 0*.*00.

- *Y*4 = 85, *Y*ˆ4 = 85*.*68, residual = 85 *−* 85*.*68 = *−*0*.*68.

- *Y*5 = 90, *Y*ˆ5 = 89*.*52, residual = 90 *−* 89*.*52 = 0*.*48.

3. Hitung ∑(*Y Y*ˆ )2:

*—*

*i* *i*

- ( 0*.*52)2 +(0*.*84)2 +(0*.*00)2 +( 0*.*68)2 +(0*.*48)2 = 0*.*2704+0*.*7056+0+0*.*4624+0*.*2304 =

*− −*

1*.*6688.

∑

1. Hitung (*Xi − X*¯ )2, dengan *X*¯ = 75/5 = 15:

- (10 *−* 15)2 + (12 *−* 15)2 + (15 *−* 15)2 + (18 *−* 15)2 + (20 *−* 15)2 = 25 + 9 + 0 + 9 + 25 = 68.

√ √ √ *√*

1. *SE*(*b*) = 1*.*6688 = 1*.*6688 = 1*.*6688 *≈* 0*.*008178 *≈* 0*.*0904.

(5*−*2)*·*68

3*·*68

204

* + \*\*Hitung statistik *t*\*\*:

1*.*912

*t* = *≈* 21*.*15 0*.*0904

* + \*\*Membandingkan dengan Tabel Distribusi t\*\*:

1. Tentukan derajat kebebasan (*df* ): *df* = *n* 2 = 5 2 = 3.

*− −*

1. Pilih tingkat signifikansi *α* = 0*.*05 untuk uji dua sisi (karena *H*1 : *ρ* = 0), maka

*̸*

*α*/2 = 0*.*025.

1. Dari tabel distribusi t (biasanya tersedia di buku statistik), cari nilai kritis *t*

untuk *df* = 3 dan *α*/2 = 0*.*025. Nilai kritisnya adalah sekitar 3.182.

1. Bandingkan: *t* = 21*.*15 *>* 3*.*182. Karena nilai *t* yang dihitung lebih besar dari nilai kritis, \*\*tolak *H*0\*\*.

*| |*

1. Alternatif: Nilai *t* = 21*.*15 sangat besar, menunjukkan p-value sangat kecil (< 0.05), yang juga mendukung penolakan *H*0.

**Kesimpulan**: Ada hubungan linier signifikan antara jam belajar dan nilai ujian (*β ̸*= 0).

### Langkah Excel untuk Verifikasi:

1. Jalankan regresi seperti di langkah (c).
2. Dari output, periksa kolom “t Stat” untuk X Variable 1 (misalnya, 22.5166605) dan “P-value” (misalnya, 0.000191815).
3. Jika P-value < 0.05, tolak *H*0, sesuai dengan perhitungan manual.

# Soal 2: Regresi Linier

Seorang manajer toko ingin mengetahui hubungan antara pengeluaran iklan (dalam juta rupiah, *X*) dan penjualan bulanan (dalam juta rupiah, *Y* ). Data dari 6 bulan adalah:

Pengeluaran Iklan (*X*) Penjualan (*Y* )

|  |  |
| --- | --- |
| 2 | 20 |
| 3 | 25 |
| 4 | 30 |
| 5 | 35 |
| 6 | 40 |
| 7 | 45 |

* 1. Tulis hipotesis nol (*H*0) dan hipotesis alternatif (*H*1) untuk hubungan ini.
  2. Hitung koefisien korelasi Pearson (*r*) menggunakan rumus awal dan Excel, lalu jelaskan maknanya.
  3. Tentukan persamaan regresi linier (*Y* = *a* + *bX*) menggunakan rumus awal dan Excel, interpretasikan, dan buat visualisasi scatter plot dengan garis regresi.

## Langkah-Langkah Pengerjaan Soal 2

### Hipotesis

* *H*0: Tidak ada hubungan linier antara pengeluaran iklan dan penjualan (*ρ* = 0

atau *β* = 0).

* *H*1: Ada hubungan linier antara pengeluaran iklan dan penjualan (*ρ* = 0 atau

*̸*

*β* = 0).

*̸*

### Koefisien Korelasi Pearson (*r*)

Rumus awal:

[*n* ∑ *X*2 *−* (∑ *X*)2][*n* ∑ *Y* 2 *−* (∑ *Y* )2]

*r* = √

*n* ∑ *XY −* ∑ *X* ∑ *Y*

(5)

- *n*∑= 6, ∑ *X* = 2 + 3 + 4 + 5 + 6 + 7 = 27, ∑ *Y* = 20 + 25 + 30 + 35 + 40 + 45 = 195.

Perhitungan:

-

*XY* = (2 *·* 20) + (3 *·* 25) + (4 *·* 30) + (5 *·* 35) + (6 *·* 40) + (7 *·* 45) = 965.

- *X*2 = 22 +32 +42 +52 +62 +72 = 139, *Y* 2 = 202 +252 +302 +352 +402 +452 = 6775.

∑ ∑

* Substitusi:

6 *·* 965 *−* 27 *·* 195

√ √

5790 *−* 5265

525

*r* = [6 *·* 139 *−* 272][6 *·* 6775 *−* 1952] = [834 *−* 729][40650 *−* 38025] = *√*105 *·* 2625 *≈* 0*.*999

### Langkah Excel:

1. Masukkan data:

- Kolom A (A2:A7): 2, 3, 4, 5, 6, 7 (label A1: “Iklan”).

- Kolom B (B2:B7): 20, 25, 30, 35, 40, 45 (label B1: “Penjualan”).

1. Pada sel C1, ketik =CORREL(A2:A7,B2:B7) untuk mendapatkan *r* 0*.*999.

*≈*

**Interpretasi**: Nilai *r* = 0*.*999 menunjukkan korelasi positif sangat kuat. Pen- geluaran iklan yang lebih besar sangat erat dengan peningkatan penjualan.

### Persamaan Regresi Linier

Rumus awal:

∑ *−* ∑ ∑

Perhitungan:

*n XY X Y*

*b* = *n* ∑ *X*2 *−* (∑ *X*)2 *, a* =

∑ *Y − b* ∑ *X*

*n*

(6)

- *b* = 6*·*965*−*27*·*195 = 5790*−*5265 = 525 = 5.

6*·*139*−*272

834*−*729

105

- *a* = 195*−*5*·*27 = 195*−*135 = 10.

6 6

- Persamaan: *Y* = 10 + 5*X*.

### Langkah Excel:

1. Masukkan data seperti pada langkah (b).
2. Buka Data > Data Analysis > Regression.
3. Pilih:
   * Input Y Range: B2:B7
   * Input X Range: A2:A7
   * Output Range: D2
4. Hasil: Intercept (*a* = 10), X Variable 1 (*b* = 5).

### Visualisasi Scatter Plot:

1. Pilih data A2:B7, lalu Insert > Scatter.
2. Klik kanan pada titik data, pilih Add Trendline, pilih Linear, dan centang

Display Equation on chart.

1. Hasil: Scatter plot dengan garis regresi *Y* = 5*X* + 10.

### Interpretasi:

* + Konstanta 10: Jika tidak ada pengeluaran iklan (*X* = 0), penjualan diperkirakan 10 juta rupiah.
  + Koefisien 5: Setiap tambahan 1 juta rupiah iklan meningkatkan penjualan sebe- sar 5 juta rupiah.

### Penerimaan/Penolakan *H*0 (Uji t):

1. Dari output regresi Excel, periksa kolom “t Stat” dan “P-value” untuk koefisien

*b*. Misalnya, *t* 71*.*18, p-value 0*.*00001.

*≈ ≈*

1. Dengan *α* = 0*.*05, jika p-value *<* 0*.*05, tolak *H*0.
2. Alternatif (menggunakan tabel t):
   * Derajat kebebasan: *df* = *n* 2 = 6 2 = 4.

*− −*

* + Nilai kritis *t* untuk *α*/2 = 0*.*025, *df* = 4 adalah 2*.*776 (lihat tabel distribusi t).

*≈*

- Karena *t* = 71*.*18 *>* 2*.*776, tolak *H*0.

*| |*

**Kesimpulan**: Ada hubungan linier signifikan antara pengeluaran iklan dan pen- jualan (*β ̸*= 0).

# Soal 3: Regresi Berganda

Seorang peneliti ingin memprediksi nilai ujian akhir (*Y* ) berdasarkan jam bela- jar per minggu (*X*1) dan jumlah kehadiran kuliah (*X*2). Data dari 5 mahasiswa adalah:

Jam Belajar (*X*1) Kehadiran (*X*2) Nilai Ujian (*Y* )

|  |  |  |
| --- | --- | --- |
| 10 | 12 | 70 |
| 12 | 14 | 75 |
| 15 | 16 | 80 |
| 18 | 18 | 85 |
| 20 | 20 | 90 |

1. Tulis hipotesis nol (*H*0) dan hipotesis alternatif (*H*1) untuk hubungan ini.
2. Hitung koefisien korelasi berganda (*R*), korelasi antar-variabel (*X*1 dengan *Y* , *X*2 dengan *Y* , *X*1 dengan *X*2), dan VIF untuk memeriksa multikolinearitas menggunakan rumus awal dan Excel, lalu interpretasikan.
3. Tentukan persamaan regresi berganda (*Y* = *a* + *b*1*X*1 + *b*2*X*2) menggunakan rumus awal dan Excel, jelaskan artinya, dan buat visualisasi plot prediksi.

## Langkah-Langkah Pengerjaan Soal 3

### Hipotesis

* + *H*0: Tidak ada hubungan linier antara jam belajar dan kehadiran dengan nilai ujian (*β*1 = *β*2 = 0).
  + *H*1: Ada hubungan linier antara jam belajar dan/atau kehadiran dengan nilai ujian (setidaknya salah satu *β*1 = 0 atau *β*2 = 0).

*̸ ̸*

### Koefisien Korelasi dan VIF Korelasi Pearson (*r*)

Rumus awal:

*r* = √

*n* ∑ *XY −* ∑ *X* ∑ *Y*

(7)

- Untuk *X*1 dan *Y* : ∑ *X*1 = 75, ∑ *Y* = 400, ∑ *X*1*Y* = 6130, ∑ *X*2 = 1193, ∑ *Y* 2 =

1

Perhitungan:

[*n* ∑ *X*2 *−* (∑ *X*)2][*n* ∑ *Y* 2 *−* (∑ *Y* )2]

32250.

5 6130 75 400

*· − ·*

*rX*1*,Y* = √[5 *·* 1193 *−* 752][5 *·* 32250 *−* 4002] *≈* 0*.*997

- Untuk *X*2 dan *Y* : ∑ *X*2 = 80, ∑ *X*2*Y* = 6550, ∑ *X*2 = 1320.

2

*rX*2*,Y* = √[5 *·* 1320 *−* 802][5 *·* 32250 *−* 4002] *≈* 0*.*997

5 *·* 6550 *−* 80 *·* 400

* + Untuk *X*1 dan *X*2: ∑ *X*1*X*2 = 1236.

*rX*1*,X*2 = [5 *·* 1193 *−* 752][5 *·* 1320 *−* 802] *≈* 0*.*986

5 *·* 1236 *−* 75 *·* 80

√

### Koefisien Korelasi Berganda (*R*)

Rumus awal:

∑ *i −*

2 (*Y*ˆ *Y*¯ )2

*√*

*R* = ∑(*Y − Y*¯ )2

*i*

*, R* = *R*2

(8)

Catatan: *R*2 dihitung **setelah model regresi** dibuat (lihat langkah (c)). Sete- lah perhitungan, *R*2 *≈* 0*.*994, sehingga:

*R* = *√*0*.*994 *≈* 0*.*997

### VIF untuk Multikolinearitas

Rumus awal:

VIF*i*

1

= 1 *− R*2

*i*

(9)

- Untuk *X*1 (regresi *X*1 terhadap *X*2): *R*2

*X*1*,X*2

= (0*.*986)2 *≈* 0*.*972.

VIF*X*

1

= 1 *−* 0*.*972 *≈* 35*.*71

1

* + Untuk *X*2 (sama, karena simetris): VIF*X*2 35*.*71.

*≈*

### Langkah Excel:

1. Masukkan data:

- Kolom A (A2:A6): 10, 12, 15, 18, 20 (label A1: “Jam Belajar”).

- Kolom B (B2:B6): 12, 14, 16, 18, 20 (label B1: “Kehadiran”).

- Kolom C (C2:C6): 70, 75, 80, 85, 90 (label C1: “Nilai Ujian”).

1. Korelasi:
   * Pada sel D1, ketik =CORREL(A2:A6,C2:C6) untuk *rX*1*,Y* 0*.*997.

*≈*

* + Pada sel E1, ketik =CORREL(B2:B6,C2:C6) untuk *rX*2*,Y* 0*.*997.

*≈*

* + Pada sel F1, ketik =CORREL(A2:A6,B2:B6) untuk *rX*1*,X*2 0*.*986.

*≈*

1. VIF:
   * Regresi *X*1 terhadap *X*2: Buka Data > Data Analysis > Regression, pilih Y Range (A2:A6), X Range (B2:B6), Output Range (G2).
   * Ambil *R*2 0*.*972 dari output.

*≈*

* + Pada sel I1, ketik =1/(1-0.972) untuk VIF*X*1 35*.*71.

*≈*

* + Ulangi untuk *X*2 terhadap *X*1, hasil sama.

### Interpretasi:

* + *rX ,Y* = 0*.*997 dan *rX ,Y* = 0*.*997 menunjukkan korelasi positif sangat kuat antara

1 2

masing-masing variabel independen dan nilai ujian.

* + *rX ,X*

1

2

= 0*.*986 menunjukkan korelasi kuat antar-variabel independen, mengindikasikan

potensi multikolinearitas.

* + VIF 35*.*71 (jauh di atas 10) mengkonfirmasi multikolinearitas serius. Dalam praktik, salah satu variabel (misalnya, kehadiran) dapat dihapus jika kurang rel- evan secara teoretis, tetapi keduanya tetap digunakan di sini karena relevansi konteks.

*≈*

* + *R* = 0*.*997 menunjukkan bahwa kombinasi jam belajar dan kehadiran sangat kuat memprediksi nilai ujian.

### Persamaan Regresi Berganda

Rumus awal (metode matriks):

**b** = (**X***T* **X**)*−*1**X***T* **Y** (10)

Matriks **X** (termasuk kolom 1 untuk konstanta):

1 10 12

1 12 14

70

 

 

 

75

**X** = 1 15 16 *,* **Y** = 80

   



 

85

90



1 18 18

1 20 20

Perhitungan (disederhanakan): *b*1 0*.*956, *b*2 0*.*956, *a* 51*.*32.

*≈ ≈ ≈*

Persamaan: *Y* = 51*.*32 + 0*.*956*X*1 + 0*.*956*X*2.

### Langkah Excel:

1. Masukkan data seperti pada langkah (b).
2. Buka Data > Data Analysis > Regression.
3. Pilih:
   * Input Y Range: C2:C6
   * Input X Range: A2:B6
   * Output Range: E2
4. Hasil: Intercept ( ), X Variable 1 ( ), X Variable 2 ( ).

*a ≈* 51*.*32 *b*1 *≈* 0*.*956 *b*2 *≈* 0*.*956

1. Ambil *R*2 0*.*994 dari output untuk menghitung *R* = *√*0*.*994 0*.*997.

*≈ ≈*

### Visualisasi Plot Prediksi:

1. Hitung nilai prediksi: Pada kolom D (D2:D6), ketik =51.32+0.956\*A2+0.956\*B2

untuk D2, lalu salin ke bawah.

1. Pilih data C2:C6 (aktual) dan D2:D6 (prediksi), lalu Insert > Scatter.
2. Tambahkan trendline untuk data aktual dan prediksi untuk melihat kesesua- ian.

### Interpretasi:

* + Konstanta 51.32: Nilai ujian dasar jika *X*1 = 0 dan *X*2 = 0.
  + Koefisien 0.956 (*X*1): Setiap tambahan 1 jam belajar meningkatkan nilai 0.956 poin (dengan *X*2 konstan).
  + Koefisien 0.956 (*X*2): Setiap tambahan 1 kehadiran meningkatkan nilai 0.956 poin (dengan *X*1 konstan).
  + Catatan: *R*2 0*.*994 dihitung setelah model regresi, menunjukkan 99.4% variasi nilai ujian dijelaskan oleh jam belajar dan kehadiran.

*≈*

### Penerimaan/Penolakan *H*0 (Uji t dan Uji F):

1. **Uji t (untuk** *b*1 **dan** *b*2**)**:
   * Dari output regresi Excel, periksa kolom “t Stat” dan “P-value” untuk *b*1 dan *b*2. Misalnya, untuk *b*1: *t* 2*.*11, p-value 0*.*169; untuk *b*2: *t* 2*.*11, p-value 0*.*169.

*≈ ≈ ≈ ≈*

* + Dengan *α* = 0*.*05, p-value *>* 0*.*05, sehingga gagal tolak *H*0 untuk masing-masing koefisien (jam belajar dan kehadiran tidak signifikan secara individu, kemungk- inan karena multikolinearitas).
  + Alternatif (tabel t):
  + Derajat kebebasan: *df* = *n − k −* 1 = 5 *−* 2 *−* 1 = 2.
  + Nilai kritis *t* untuk *α*/2 = 0*.*025, *df* = 2 adalah *≈* 4*.*303.
  + Karena *|t|* = 2*.*11 *<* 4*.*303, gagal tolak *H*0 untuk *b*1 dan *b*2.

### Uji F (untuk model keseluruhan):

* + Dari tabel ANOVA di output regresi, periksa kolom “F” dan “Significance F”. Misalnya, *F* 74*.*66, p-value 0*.*013.

*≈ ≈*

* + Dengan *α* = 0*.*05, p-value *<* 0*.*05, tolak *H*0.
  + Alternatif (tabel F):
  + Derajat kebebasan: *df*1 = *k* = 2, *df*2 = *n k* 1 = 2.

*− −*

* + Nilai kritis *F* untuk *α* = 0*.*05, *df*1 = 2, *df*2 = 2 adalah 19*.*0 (lihat tabel distribusi F).

*≈*

- Karena *F* = 74*.*66 *>* 19*.*0, tolak *H*0.

**Kesimpulan**: Model regresi secara keseluruhan signifikan (setidaknya salah satu variabel berpengaruh), tetapi secara individu, jam belajar dan kehadiran tidak signifikan karena multikolinearitas tinggi (VIF *≈* 35*.*71).

# Soal 4: Regresi Berganda

Seorang pengusaha ingin memprediksi penjualan (*Y* , dalam juta rupiah) berdasarkan pengeluaran iklan (*X*1, dalam juta rupiah) dan jumlah tenaga penjual (*X*2). Data

dari 5 bulan adalah:

|  |  |  |
| --- | --- | --- |
| Iklan (*X*1) | Tenaga Penjual (*X*2) | Penjualan (*Y* ) |
| 2 | 3 | 20 |
| 3 | 4 | 25 |
| 4 | 5 | 30 |
| 5 | 6 | 35 |
| 6 | 7 | 40 |

1. Tulis hipotesis nol (*H*0) dan hipotesis alternatif (*H*1).
2. Hitung koefisien korelasi berganda (*R*), korelasi antar-variabel (*X*1 dengan *Y* , *X*2 dengan *Y* , *X*1 dengan *X*2), dan VIF untuk memeriksa multikolinearitas menggunakan rumus awal dan Excel, lalu interpretasikan.
3. Tentukan persamaan regresi berganda (*Y* = *a* + *b*1*X*1 + *b*2*X*2) menggunakan rumus awal and Excel, jelaskan artinya, dan buat visualisasi plot prediksi.

## Langkah-Langkah Pengerjaan Soal 4

### Hipotesis

* + *H*0: Tidak ada hubungan linier antara pengeluaran iklan dan jumlah tenaga penjual dengan penjualan (*β*1 = *β*2 = 0).
  + *H*1: Ada hubungan linier antara pengeluaran iklan dan/atau jumlah tenaga pen- jual dengan penjualan (setidaknya salah satu *β*1 = 0 atau *β*2 = 0).

*̸ ̸*

### Koefisien Korelasi dan VIF Korelasi Pearson (*r*)

Rumus awal:

*r* = √

*n* ∑ *XY −* ∑ *X* ∑ *Y*

(11)

[*n* ∑ *X*2 *−* (∑ *X*)2][*n* ∑ *Y* 2 *−* (∑ *Y* )2]

Perhitungan:

- Untuk *X*1 dan *Y* : ∑ *X*1 = 20, ∑ *Y* = 150, ∑ *X*1*Y* = 680, ∑ *X*2 = 90, ∑ *Y* 2 = 4950.

1

5 *·* 680 *−* 20 *·* 150 [5 *·* 90 *−* 202][5 *·* 4950 *−* 1502]

*rX*1*,Y* = √ *≈* 0*.*993

- Untuk *X*2 dan *Y* : ∑ *X*2 = 25, ∑ *X*2*Y* = 875, ∑ *X*2 = 135.

2

*rX*2*,Y* = [5 *·* 135 *−* 252][5 *·* 4950 *−* 1502] *≈* 0*.*993

5 *·* 875 *−* 25 *·* 150

√

* + Untuk *X*1 dan *X*2: ∑ *X*1*X*2 = 110.

*rX*1*,X*2 = [5 *·* 90 *−* 202][5 *·* 135 *−* 252] *≈* 0*.*964

5 *·* 110 *−* 20 *·* 25

√

### Koefisien Korelasi Berganda (*R*)

Rumus awal:

∑ *i −*

2 (*Y*ˆ *Y*¯ )2

*√*

*R* = ∑(*Y − Y*¯ )2

*i*

*, R* = *R*2

(12)

Catatan: *R*2 dihitung **setelah model regresi** dibuat (lihat langkah (c)). Sete- lah perhitungan, *R*2 *≈* 0*.*998, sehingga:

*R* = *√*0*.*998 *≈* 0*.*999

### VIF untuk Multikolinearitas

Rumus awal:

VIF*i*

1

= 1 *− R*2

*i*

(13)

- Untuk *X*1 (regresi *X*1 terhadap *X*2): *R*2

*X*1*,X*2

= (0*.*964)2 *≈* 0*.*929.

VIF*X*

1

= 1 *−* 0*.*929 *≈* 14*.*08

1

* + Untuk *X*2 (sama, karena simetris): VIF*X*2 14*.*08.

*≈*

### Langkah Excel:

1. Masukkan data:

- Kolom A (A2:A6): 2, 3, 4, 5, 6 (label A1: “Iklan”).

* + Kolom B (B2:B6): 3, 4, 5, 6, 7 (label B1: “Tenaga Penjual”).

- Kolom C (C2:C6): 20, 25, 30, 35, 40 (label C1: “Penjualan”).

1. Korelasi:
   * Pada sel D1, ketik =CORREL(A2:A6,C2:C6) untuk *rX*1*,Y* 0*.*993.

*≈*

* + Pada sel E1, ketik =CORREL(B2:B6,C2:C6) untuk *rX*2*,Y* 0*.*993.

*≈*

* + Pada sel F1, ketik =CORREL(A2:A6,B2:B6) untuk *rX*1*,X*2 0*.*964.

*≈*

1. VIF:
   * Regresi *X*1 terhadap *X*2: Buka Data > Data Analysis > Regression, pilih Y Range (A2:A6), X Range (B2:B6), Output Range (G2).
   * Ambil *R*2 0*.*929 dari output.

*≈*

* + Pada sel I1, ketik =1/(1-0.929) untuk VIF*X*1 14*.*08.

*≈*

* + Ulangi untuk *X*2 terhadap *X*1, hasil sama.

### Interpretasi:

* + *rX ,Y* = 0*.*993 dan *rX ,Y* = 0*.*993 menunjukkan korelasi positif sangat kuat antara

1 2

masing-masing variabel independen dan penjualan.

* + *rX ,X*

1

2

= 0*.*964 menunjukkan korelasi kuat antar-variabel independen, mengindikasikan

potensi multikolinearitas.

* + VIF 14*.*08 (di atas 10) mengkonfirmasi multikolinearitas. Dalam praktik, salah satu variabel (misalnya, tenaga penjual) dapat dihapus jika kurang rele- van, tetapi keduanya tetap digunakan di sini karena relevansi konteks.

*≈*

* + *R* = 0*.*999 menunjukkan bahwa kombinasi pengeluaran iklan dan jumlah tenaga penjual sangat kuat memprediksi penjualan.

### Persamaan Regresi Berganda

Rumus awal:

**b** = (**X***T* **X**)*−*1**X***T* **Y** (14)

Matriks **X**:

1 2 3

 

1 3 4

   

**X** = 1 4 5 *,* **Y** = 30

20

 

25

1 5 6 35

 

 

1 6 7 40

Perhitungan: *b*1 2*.*5, *b*2 2*.*5, *a* 7*.*5.

*≈ ≈ ≈*

Persamaan: *Y* = 7*.*5 + 2*.*5*X*1 + 2*.*5*X*2.

### Langkah Excel:

1. Masukkan data seperti pada langkah (b).
2. Buka Data > Data Analysis > Regression.
3. Pilih:
   * Input Y Range: C2:C6
   * Input X Range: A2:B6
   * Output Range: E2
4. Hasil: Intercept ( ), X Variable 1 ( ), X Variable 2 ( ).

*a ≈* 7*.*5 *b*1 *≈* 2*.*5 *b*2 *≈* 2*.*5

1. Ambil *R*2 0*.*998 dari output untuk menghitung *R* = *√*0*.*998 0*.*999.

*≈ ≈*

### Visualisasi Plot Prediksi:

1. Hitung nilai prediksi: Pada kolom D (D2:D6), ketik =7.5+2.5\*A2+2.5\*B2 un- tuk D2, lalu salin ke bawah.
2. Pilih data C2:C6 (aktual) dan D2:D6 (prediksi), lalu Insert > Scatter.
3. Tambahkan trendline untuk data aktual dan prediksi untuk melihat kesesua- ian.

### Interpretasi:

* + Konstanta 7.5: Penjualan dasar jika *X*1 = 0 dan *X*2 = 0 adalah 7.5 juta rupiah.
  + Koefisien 2.5 (*X*1): Setiap tambahan 1 juta rupiah iklan meningkatkan pen- jualan 2.5 juta rupiah (dengan *X*2 konstan).
  + Koefisien 2.5 (*X*2): Setiap tambahan 1 tenaga penjual meningkatkan penjualan

2.5 juta rupiah (dengan *X*1 konstan).

- Catatan: *R*2 0*.*998 dihitung setelah model regresi, menunjukkan 99.8% variasi penjualan dijelaskan oleh pengeluaran iklan dan tenaga penjual.

*≈*

### Penerimaan/Penolakan *H*0 (Uji t dan Uji F):

1. **Uji t (untuk** *b*1 **dan** *b*2**)**:
   * Dari output regresi Excel, periksa kolom “t Stat” dan “P-value” untuk *b*1 dan *b*2. Misalnya, untuk *b*1: *t* 3*.*54, p-value 0*.*071; untuk *b*2: *t* 3*.*54, p-value 0*.*071.

*≈ ≈ ≈ ≈*

* + Dengan *α* = 0*.*05, p-value *>* 0*.*05, sehingga gagal tolak *H*0 untuk masing-masing

koefisien (iklan dan tenaga penjual tidak signifikan secara individu, kemungki- nan karena multikolinearitas).

* + Alternatif (tabel t):
  + Derajat kebebasan: *df* = *n k* 1 = 5 2 1 = 2.

*− − − −*

* + Nilai kritis *t* untuk *α*/2 = 0*.*025, *df* = 2 adalah 4*.*303.

*≈*

* + Karena *t* = 3*.*54 *<* 4*.*303, gagal tolak *H*0 untuk *b*1 dan *b*2.

*| |*

### Uji F (untuk model keseluruhan):

* + Dari tabel ANOVA di output regresi, periksa kolom “F” dan “Significance F”. Misalnya, *F* 250*.*0, p-value 0*.*004.

*≈ ≈*

* + Dengan *α* = 0*.*05, p-value *<* 0*.*05, tolak *H*0.
  + Alternatif (tabel F):
  + Derajat kebebasan: *df*1 = *k* = 2, *df*2 = *n k* 1 = 2.

*− −*

* + Nilai kritis *F* untuk *α* = 0*.*05, *df*1 = 2, *df*2 = 2 adalah 19*.*0.

*≈*

- Karena *F* = 250*.*0 *>* 19*.*0, tolak *H*0.

**Kesimpulan**: Model regresi secara keseluruhan signifikan (setidaknya salah satu variabel berpengaruh), tetapi secara individu, iklan dan tenaga penjual tidak signifikan karena multikolinearitas (VIF *≈* 14*.*08).

# Catatan untuk Mahasiswa

* + - Pastikan memahami konsep hipotesis, korelasi, regresi, multikolinearitas, serta uji t dan F sebelum mengerjakan praktikum.
    - Gunakan Excel untuk perhitungan dan visualisasi. Aktifkan Analysis ToolPak

jika belum tersedia.

* + - Perhatikan multikolinearitas (VIF *>* 10) dan diskusikan dengan dosen jika perlu menghapus variabel.
    - Untuk uji t dan F, fokus pada p-value dari Excel. Gunakan tabel distribusi t dan F (di buku statistik) untuk memahami konsep:
      * **Tabel t**: Cari nilai kritis berdasarkan *df* dan *α*/2 (misalnya, 0.025 untuk uji dua sisi).
      * **Tabel F**: Cari nilai kritis berdasarkan *df*1 (jumlah variabel indepen- den), *df*2 (sisa *df* ), dan *α* = 0*.*05.
    - Interpretasi hasil harus jelas dan sesuai konteks soal.