

Responsi 3

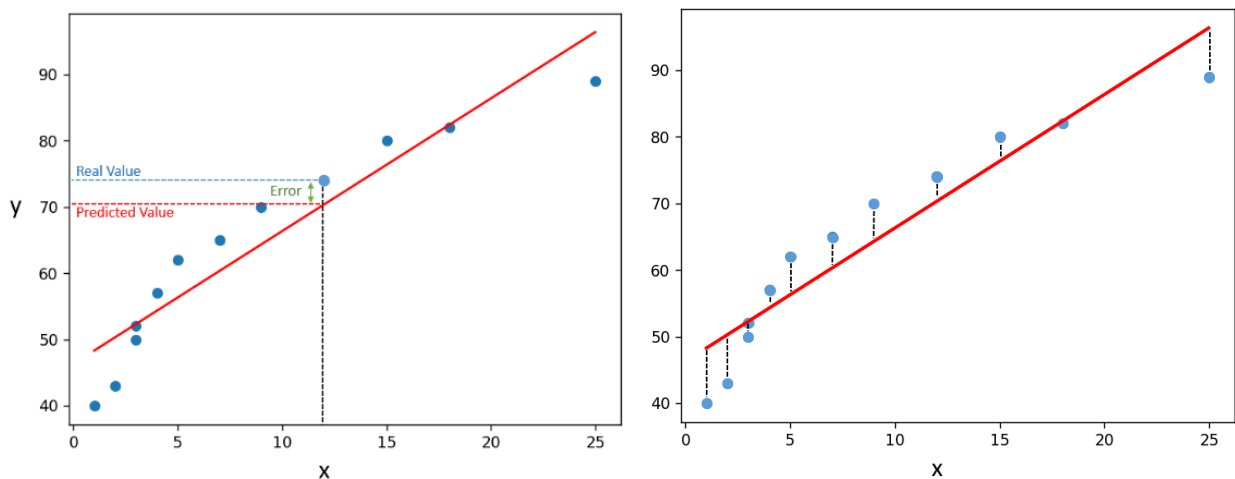
ANALISIS REGRESI

Pokok Bahasan:

1. Penguraian keragaman total
2. Uji hipotesis parameter regresi
3. Penduga selang kepaercayaan parameter model

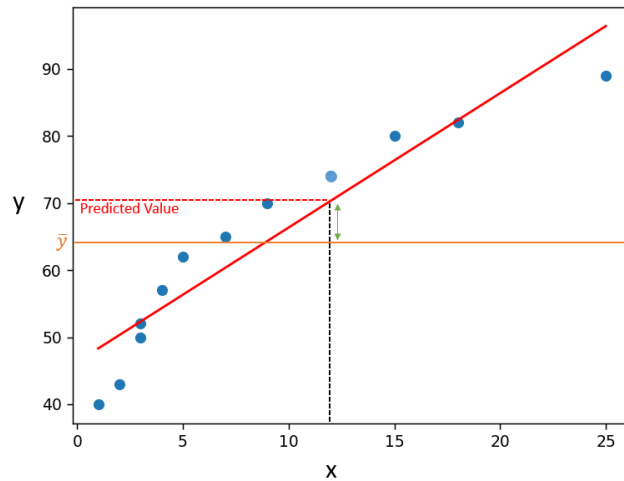
1. Penguraian keragaman total

Nilai Y_i yang menyebar secara acak dan bersifat stokastik menyebabkan titik amatannya tidak pasti, pada x tertentu. Akibatnya, terdapat keragaman data yang disebabkan oleh adanya **error** atau **sisaan** dan menyimpangnya **model regresi**. Perhatikan grafik berikut.



Grafik tersebut menunjukkan bahwa setiap amatan memiliki error yang besarnya berbeda-beda tiap amatan. Error tersebut merupakan $y_i - \hat{y}_i$ atau selisih antara nilai amatan sebenarnya dengan nilai y dugaan yang berada tepat di garis regresi.

Selain karena error, keragaman nilai y_i juga disebabkan oleh model regresi yang dibuat. Menyimpangnya suatu dugaan garis regresi terhadap rataannya menyebabkan beragamnya data.



Karena adanya model regresi maka $\bar{y} = \widehat{y}_l$. Sehingga keragaman data karena model regresi yaitu $\widehat{y}_l - \bar{y}$ atau selisih antara nilai y dugaan pada x tertentu dengan y rata-rata.

Ukuran Keragaman

1. Jumlah Kuadrat Total (JKT)

JKT merupakan jumlah kuadrat penyimpangan nilai amatan sebenarnya terhadap y rata-rata.

$$JKT = \sum (y_i - \bar{y})^2$$

2. Jumlah Kuadrat Regresi (JKR)

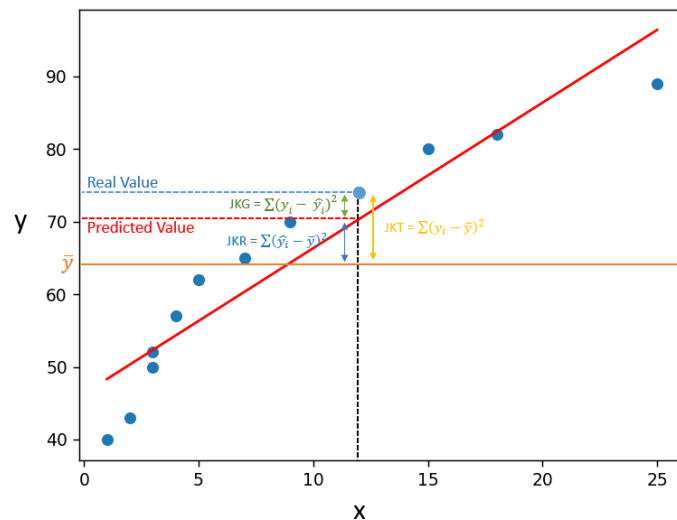
JKR merupakan jumlah kuadrat karena penyimpangan oleh regresi. Nilai JKR didapatkan dari penjumlahan kuadrat selisih antara nilai y dugaan dengan y rata-rata.

$$JKR = \sum (\widehat{y}_i - \bar{y})^2$$

3. Jumlah Kuadrat Galat (JKG)

Nilai JKG didapat dari penjumlahan kuadrat dari *error* (galat) tiap amatan.

$$JKG = \sum (y_i - \widehat{y}_i)^2$$



Berdasarkan gambar diatas, maka dapat juga dituliskan bahwa **JKT = JKR + JKG**
Adapun tabel sidik ragam dari model regresi linier sederhana adalah sebagai berikut.

Sumber Keragaman (SK)	Derajat Bebas (db)	Jumlah Kuadrat (JK)	Kuadrat Tengah (KT)
Regresi	1	$\sum (\hat{y}_i - \bar{y})^2$	$\frac{JKR}{1}$
Sisaan	$n - 2$	$\sum (y_i - \bar{y})^2$	$\frac{JKG}{n - 2}$
Total	$n - 1$	$\sum (y_i - \hat{y}_i)^2$	

2. Uji Hipotesis Parameter Regresi

Uji Hipotesis ini menggunakan uji t untuk menguji parameter β_0 (intersep) dan β_1 (koefisien kemiringan garis) pada model regresi. Uji ini memerlukan asumsi bahwa ε_i menyebar normal.

- Uji t untuk koefisien kemiringan garis regresi populasi (β_1)

Hipotesis uji:

$H_0: \beta_1 = 0$ (Tidak ada hubungan linier antara x dan y)

$H_1: \beta_1 \neq 0$ (Ada hubungan linier antara x dan y)

Statistik uji:

$$t = \frac{b_1 - \beta_1}{S_{b_1}}$$

$$S^2_{b_1} = \frac{S^2_e}{\sum (x_i - \bar{x})^2} = \frac{S^2_e}{(n - 1)S^2_x}$$

dengan:

dengan:

b_1 = koefisien kemiringan regresi

S_{b_1} = simpangan baku kemiringan

β_1 = kemiringan yang dihipotesiskan

S^2_x = dugaan ragam x

S_{b_1} = simpangan baku kemiringan garis regresi

$$S_e = \sqrt{\frac{JKG}{n-2}} = \sqrt{KTG}$$

Kaidah keputusan: Tolak H_0 jika $t > t_{(db, \alpha/2)}$; db = n-2

- **Uji t untuk intersep garis regresi populasi (β_0)**

Hipotesis uji:

$H_0: \beta_0 = 0$ (Semua nilai y dapat dijelaskan oleh x)

$H_1: \beta_0 \neq 0$ (Ada nilai y yang tidak dapat dijelaskan oleh x)

Statistik uji:

$$t = \frac{b_0 - \beta_0}{S_{b_0}}$$

dengan:

b_0 = intersep garis regresi

β_0 = intersep yang dihipotesiskan

S_{b_0} = simpangan baku intersep

Kaidah keputusan:

Tolak H_0 jika $t > t_{(db, \alpha/2)}$; db = n-2

$$S^2_{b_0} = \frac{S^2_e \sum x_i^2}{n \sum (x_i - \bar{x})^2}$$

dengan:

S_{b_0} = simpangan baku intersep

$$S_e = \sqrt{\frac{JKG}{n-2}} = \sqrt{KTG}$$

3. Penduga selang kepercayaan parameter model

Untuk membangun selang kepercayaan parameter regresi, kita perlu mengasumsikan bahwa **error atau galat/sisaan menyebar normal** sehingga **b_0 dan b_1 juga menyebar normal**.

Dugaan selang kepercayaan $(1 - \alpha) \times 100\%$ bagi β :

$$b_0 \pm t_{n-2, \frac{\alpha}{2}} \times S_{b_0}$$

Dugaan selang kepercayaan $(1 - \alpha) \times 100\%$ bagi β :

$$b_1 \pm t_{n-2, \frac{\alpha}{2}} \times S_{b_1}$$

Jika kita mengambil sampel berulang dengan ukuran yang sama pada nilai x yang sama dan membangun **selang kepercayaan 95%**, selang ini diharapkan **memuat nilai yang sebenarnya**. Misalkan **selang kepercayaan 95%** untuk parameter kemiringan garis (**β_1**), maka 95% ini diharapkan **memuat nilai sebenarnya dari kemiringan garis/slope**. Jika selang bagi β_1 **tidak memuat angka 0 (nol)**, dapat dikatakan bahwa peubah penjelas/bebas **x berpengaruh** pada peubah respon/peubah tak bebas **y**.

Ilustrasi Kasus

Diketahui data sebuah toko meliputi banyaknya promosi dalam satu bulan (x) dan rata-rata unit barang yang terjual per hari (y)

Promosi (kali/bulan) (x)	Penjualan (unit/hari) (y)
25	100
27	105
29	108
30	122
35	120
50	145
55	143
60	150
63	148
65	157
70	161
71	175
73	174
75	176
80	185

Dugaan persamaan garis regresi:

$$y = 69.60912 + 1.392157x$$

$$\text{Slope} = 1.392157$$

$$\text{Intersep} = 69.60912$$

1. Uji hipotesis untuk parameter regresi

Dengan nilai slope (b_1) sebesar 1.392, selanjutnya akan diuji apakah promosi (x) mempengaruhi penjualan (y) (berhubungan linier) atau tidak.

Hipotesis uji:

$H_0: \beta_1 = 0$ (Tidak ada hubungan linier antara banyaknya promosi dan rata-rata unit terjual)

$H_1: \beta_1 \neq 0$ (Ada hubungan linier antara banyaknya promosi dan rata-rata unit terjual)

Statistik uji:

$$t = \frac{b_1 - \beta_1}{S_{b_1}} = \frac{1.392 - 0}{0.07551} = 18.44$$

$$t_{(0.05,13)} = 2.16037$$

Kaidah keputusan:

Karena $t > t_{(0.05,13)}$ maka tolak H_0 yang berarti ada hubungan linier antara x dan y. Cukup bukti untuk menyatakan bahwa banyaknya promosi mempengaruhi rata-rata unit terjual per hari pada taraf nyata 5%.

Dengan data yang sama beserta nilai intersep (b_0) = 69.609, akan diuji apakah ada penjualan (y) yang tidak dapat dijelaskan oleh banyaknya promosi atau tidak.

Hipotesis uji:

$H_0: \beta_0 = 0$ (Semua nilai penjualan dapat dijelaskan oleh banyaknya promosi)

$H_1: \beta_0 \neq 0$ (Ada nilai penjualan yang tidak dapat dijelaskan oleh banyaknya promosi)

Statistik uji:

$$t = \frac{b_0 - \beta_0}{S_{b_0}} = \frac{69.609 - 0}{4.31285} = 16.14$$

$$t_{(0.05,1)} = 12.7062$$

Kaidah keputusan:

Karena $t > t_{(0.05,1)}$ maka tolak H_0 . Cukup bukti untuk menyatakan bahwa ada nilai penjualan yang tidak dapat dijelaskan oleh banyaknya promosi pada taraf nyata 5%.

2. Penduga selang kepercayaan parameter model

Promosi (kali/bulan) (x)	Penjualan (unit/hari) (y)	\hat{y}	x_i^2	$(y_i - \hat{y}_i)^2$	$(x_i - \bar{x})^2$
25	100	104.409	625	19.439281	833.284444
27	105	107.193	729	4.809249	721.817778
29	108	109.977	841	3.908529	618.351111
30	122	111.369	900	113.01816	569.617778
35	120	118.329	1225	2.792241	355.951111
50	145	139.209	2500	33.535681	14.9511111
55	143	146.169	3025	10.042561	1.28444444
60	150	153.129	3600	9.790641	37.6177778
63	148	157.305	3969	86.583025	83.4177778
65	157	160.089	4225	9.541921	123.951111
70	161	167.049	4900	36.590401	260.284444
71	175	168.441	5041	43.020481	293.551111
73	174	171.225	5329	7.700625	366.084444
75	176	174.009	5625	3.964081	446.617778
80	185	180.969	6400	16.248961	682.951111
			48934	400.98584	5409.73333

$$S_e = \sqrt{\frac{JKG}{n-2}} = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n-2}} = \sqrt{\frac{400.98584}{15-2}} = 5.55383$$

$$S_{b_0} = \sqrt{\frac{S_e^2 \sum x_i^2}{n \sum (x_i - \bar{x})^2}} = \sqrt{\frac{(5.55383)^2 \times 48934}{15 \times 5409.7333}} = 4.31285$$

$$S_{b_1} = \sqrt{\frac{S_e^2}{\sum (x_i - \bar{x})^2}} = \sqrt{\frac{(5.55383)^2}{5409.7333}} = 0.07551$$

$$t_{(n-2, \frac{\alpha}{2})} = 2.160369$$

Penduga selang kepercayaan 95% bagi β_0

$$\begin{aligned} b_0 \pm t_{n-2, \frac{\alpha}{2}} \times S_{b_0} &= 69.60912 \pm (2.160369 \times 4.31285) \\ &= 69.60912 \pm 9.317347 = [\mathbf{60.2918, 78.926}] \end{aligned}$$

Penduga selang kepercayaan 95% bagi β_1

$$\begin{aligned} b_1 \pm t_{n-2, \frac{\alpha}{2}} \times S_{b_1} &= 1.392157 \pm (2.160369 \times 0.07551) \\ &= 1.392157 \pm 0.07551 = [\mathbf{1.229028, 1.555287}] \end{aligned}$$

3. Tabel Sidik Ragam

Sumber Keragaman (SK)	Derajat Bebas (db)	Jumlah Kuadrat (JK)	Kuadrat Tengah (KT)
Regresi	1	10484.62	10484.62
Sisaan	13	400.9846	30.84497
Total	14	10885.6	

$$JKR = \sum (\hat{y}_i - \bar{y})^2 = 10484.62$$

$$KTR = \frac{JKR}{dbr} = \frac{10484.62}{1} = 10484.62$$

$$JKG = \sum (y_i - \hat{y}_i)^2 = 400.9846$$

$$KTG = \frac{JKG}{dbg} = \frac{400.9846}{13} = 30.84497$$

$$JKT = \sum (y_i - \bar{y})^2 = 10885.6$$