

REGRESI LINEAR

Identitas Bahan Ajar

Mata Pelajaran	: Matematika
Materi	: Analisis Regresi
Kelas / Fase	: XI / F
Semester	: Genap

Capaian Pembelajaran

Di akhir fase F, peserta didik dapat melakukan proses penyelidikan statistika untuk data bivariat. Mereka dapat mengidentifikasi dan menjelaskan asosiasi antara dua variabel kategorikal dan antara dua variabel numerikal. Mereka dapat memperkirakan model linear terbaik (best fit) pada data numerikal. Mereka dapat membedakan hubungan asosiasi dan sebab-akibat. Peserta didik memahami konsep peluang bersyarat dan kejadian yang saling bebas menggunakan konsep permutasi dan kombinasi.

Tujuan Pembelajaran

- Peserta didik dapat menggambar persamaan garis regresi linear.
- Peserta didik dapat menentukan persamaan garis regresi linear.
- Peserta didik dapat menginterpretasikan persamaan garis regresi linear.
- Peserta didik dapat menerapkan interpolasi dan ekstrapolasi data berdasarkan suatu persamaan garis regresi linear.

Petunjuk Penggunaan

- Pelajari tujuan pembelajaran yang akan dicapai dalam pembelajaran.
- Pelajari uraian materi secara sistematis dan komprehensif dalam setiap kegiatan pembelajaran.
- Sesuaikan latihan soal dengan kebutuhan peserta didik.

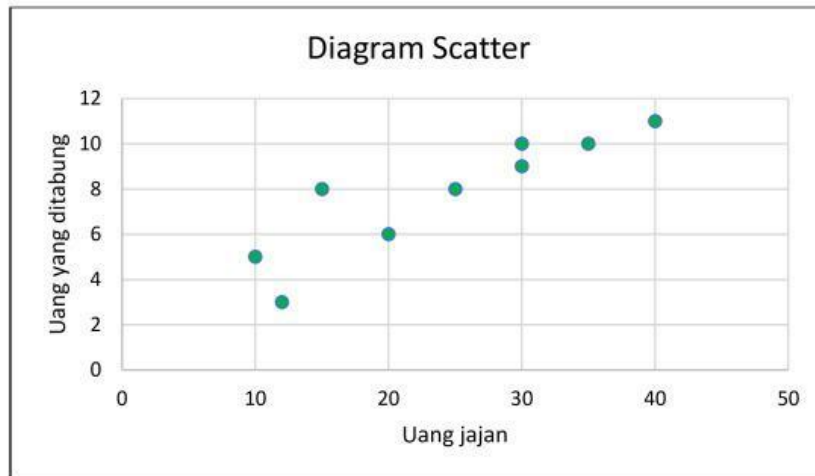
Materi Pokok

- Pengertian Regresi Linear Sederhana.
- Menyusun persamaan garis regresi linear sederhana dengan metode kuadrat terkecil.
- Menggambar garis regresi linear sederhana.
- Interpretasi garis regresi.
- Internolasi dan ekstranolasi data

Materi Pokok

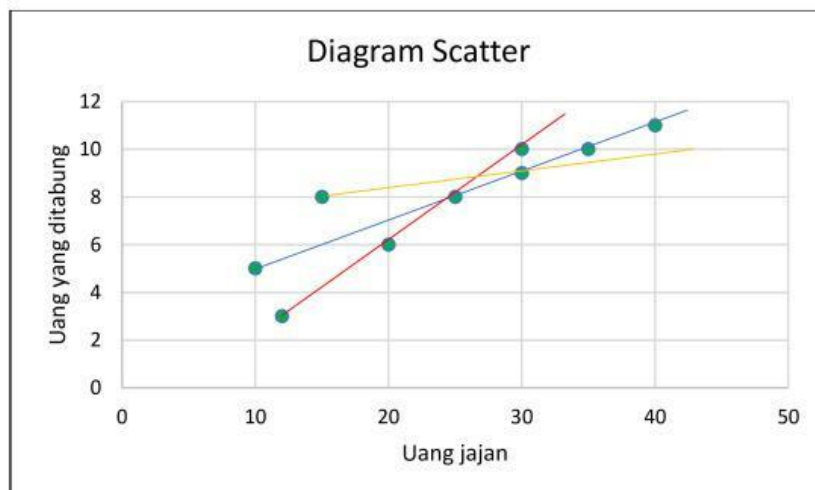
A. Pengertian Regresi Linear Sederhana

Perhatikan diagram scatter pada gambar 1.1 berikut.



Gambar 1. Diagram scatter jumlah uang jajan dengan jumlah uang yang ditabung

Diagram scatter tersebut menyajikan informasi mengenai hubungan jumlah uang jajan dengan jumlah uang yang ditabung. Terlihat bahwa diagram mengarah pada korelasi positif dengan bentuk tren data linear. Jika kamu ingin mengetahui jumlah uang yang ditabung, kita dapat memprediksinya dengan menggunakan persamaan garis yang terbentuk melalui pola sebaran data pada gambar 1.1. Dari beberapa garis yang terbentuk, terdapat satu garis yang paling tepat untuk memperkirakan hubungan dua variabel kuantitatif. Garis tersebut dinamakan garis regresi linear. Berikut adalah prediksi beberapa garis regresi linear yang terbentuk dari diagram scatter hubungan antara jumlah uang jajan dengan jumlah uang yang ditabung.



Gambar 2 Prediksi garis regresi linear

Hubungan linear antara dua variabel tersebut dapat dinyatakan dalam persamaan matematika atau model matematika yang dapat digunakan untuk melakukan prediksi yang disebut sebagai model regresi linear sederhana. Karena bersifat estimasi, maka garis regresi linear sederhana ini tidak harus melalui semua titik-titik koordinat.

B. Menyusun persamaan garis regresi linear sederhana dengan metode kuadrat terkecil

Sebagai langkah awal untuk mengenal metode koadrat terkecil dalam menyusun persamaan garis regresi linear sederhana, maka berdasarkan contoh pada 2.1 mengenai prediksi garis linear hubungan antara jumlah uang jajan dengan jumlah uang yang akan ditabung selesaikan aktivitas berikut:

1. Carilah persamaan garisnya dengan mencari gradien terlebih dahulu dan titik potong sumbu y, kemudian lakukan substitusi kedalam persamaan garis lurus $\hat{Y} = mx + c$, dengan m = gradien dan c = titik potong sumbu y.
2. Lengkapi tabel berikut ini untuk menghitung jumlah kuadrat residu.

x	y	$\hat{y}$	$y - \hat{y}$	$(y - \hat{y})^2$
1	3			
3	8			
5	12			
7	20			
10	27			
14	35			
21	55			
Total jumlah kuadrat residu				

Berdasarkan diagram scatter, seseorang dapat menentukan model matematika yang berkaitan dengan diagram scatter tersebut. Apabila titik-titik tersebut atau pola sebaran datanya berada di sekitar garis lurus, maka model matematikanya linear. Model hubungan linear antara dua variable kuantitatif pada populasi adalah:

$$Y_i = \alpha + \beta X_i + \varepsilon_i$$

Untuk setiap pasangan (X_i, Y_i) dengan :

Y_i = nilai ke-i variabel Y

α = suku tetap

β = suku tetap

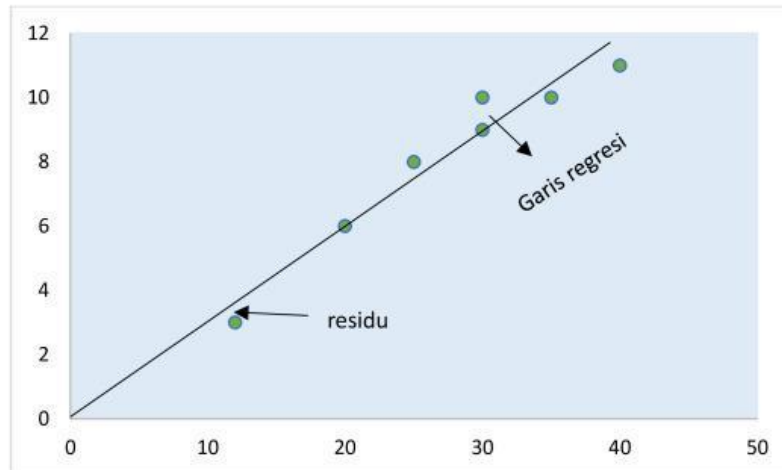
ε_i = error

Koefisien regresi (β) memiliki arti bahwa setiap perubahan satu unit variabel independen diharapkan terjadi perubahan β unit variabel dependen. Karena tidak dapat melihat nilai α dan β , maka harus dicari estimator untuk α , β , dan ε yang diperoleh dari sampel. Estimator untuk

α disajikan dengan a , Estimator untuk β disajikan dengan $\hat{\beta}$, Estimator untuk ε disajikan dengan e . Dengan demikian, model hubungan linear antara dua variabel pada sampel adalah :

$$Y_i = a + bX_i + e_i$$

Perhatikan gambar berikut:



Gambar 3. Ilustrasi garis regresi linear

Garis regresi adalah garis yang dipakai untuk memprediksi nilai Y jika diketahui nilai X tertentu. Sesuai formula matematika persamaan linear, model persamaan garis regresi linear adalah:

$$\hat{Y} = a + bX$$

Keterangan:

$\hat{Y}$: nilai variabel dependen yang diprediksi

a : titik potong sumbu y

b : gradien garis regresi

X : nilai variabel independen

Notasi $\hat{Y}$ (dibaca Y topi) merupakan nilai Y prediktif jika diketahui nilai X tertentu. Umumnya, nilai Y prediktif tidak sama dengan nilai Y sesungguhnya pada data. Selisih nilai Y sesungguhnya dengan Y prediktif disebut residu dan dilambangkan dengan “ e ” pada model hubungan linear pada sampel. Oleh karena itu didapat:

$$e_1 = Y_1 - \hat{Y}$$

Garis regresi harus berada disekitar titik-titik (X_i, Y_i) pada diagram pencar. Kita tidak dapat menarik garis lurus yang melewati semua titik (X_i, Y_i) . Oleh karena itu, kita mencari garis

regresi yang paling dekat ke titik-titik tersebut apabila jumlah kuadrat residu paling kecil. Cara menyusun persamaan garis regresi seperti ini disebut dengan metode kuadrat terkecil yang pada dasarnya meminimumkan:

$$D = \sum_{i=1}^n (Y_i - a - bX_i)^2$$

D akan mencapai minimum jika nilai a dan b adalah:

$$a = \frac{(\sum Y)(\sum X^2) - (\sum X)(\sum XY)}{n \sum X^2 - (\sum X)^2}$$

$$b = \frac{n(\sum XY) - (\sum X)(\sum Y)}{n \sum X^2 - (\sum X)^2}$$

Jadi, persamaan regresinya adalah:

$$\hat{Y} = a + Bx$$

$$Y = \frac{(\sum Y)(\sum X^2) - (\sum X)(\sum XY)}{n \sum X^2 - (\sum X)^2} + \frac{n(\sum XY) - (\sum X)(\sum Y)}{n \sum X^2 - (\sum X)^2}$$

C. Menggambar Garis Regresi Linear Sederhana

Setelah didapat garis regresi linear sederhana, kita dapat menggambar garis tersebut pada diagram scatter dengan cara substitusi x dan y ke persamaan garis regresi linear sederhana $\hat{Y} = a + Bx$ sebagai berikut:

1. Jika $Y = 0$ sehingga didapat titik $(X, 0)$
2. Jika $X = 0$ sehingga didapat titik $(Y, 0)$
3. Dikarenakan persamaan garis regresi linear sederhana melewati rata-rata x dan y, maka substitusikanlah $x = \bar{x}$ dan $y = \bar{y}$ sehingga didapatkan nilai $(\bar{x}, \bar{y})$

Setelah didapatkan ketiga titik pada langkah 1, 2, dan 3 kemudian buatlah garis yang melewati ketiga titik tersebut.

D. Interpretasi Garis Regresi

Suatu perusahaan makanan akan meluncurkan produk makanan terbarunya. Tiap bulan, Perusahaan tersebut memiliki target penjualan. Agar target terpenuhi, Perusahaan tersebut melakukan iklan di berbagai sosial media serta mencari sponsor untuk membantu untuk mengiklankan produk terbarunya.

- a. Hubungan antara hasil penjualan dan biaya pengeluaran iklan diperoleh persamaan regresi linear $\hat{Y} = 2,94 + 0,95 x$.

Interpretasi:

- Hasil penjualan sebesar 2,94 pada saat pengeluaran iklan bernilai 0 satuan.
- Hasil penjualan akan bertambah sebesar 0,95 pada setiap kenaikan pengeluaran iklan sebanyak 1 satuan.

b. Hubungan antara biaya iklan dengan banyaknya sponsor diperoleh persamaan regresi linear $\hat{Y} = 49,73 - 3,27x$.

- Biaya iklan sebesar 49,73 pada saat sponsor bernilai 0 satuan.
- Biaya iklan akan berkurang sebesar 3,27 pada setiap penambahan sponsor sebanyak 1 satuan.

E. Interpolasi dan Ekstrapolasi Data

Regresi linear sederhana bertujuan untuk memprediksi nilai variabel dependen (Y) apabila nilai variabel independen (X) diketahui. Pada umumnya, nilai X ini terdapat batasan atau domain. Berdasarkan posisi nilai X, kita dapat memprediksi nilai Y dengan cara:

- a. Interpolasi, adalah penggunaan hubungan antar variabel untuk memprediksi nilai Y jika nilai X berada di dalam domain.
- b. Ekstrapolasi, adalah penggunaan hubungan antar variabel untuk memprediksi nilai Y jika nilai X berada di luar domain.

Untuk memprediksi nilai Y menggunakan persamaan garis regresi linear sederhana, hanya dapat dilakukan dengan interpolasi.