

ANALISIS REGRESI SEDERHANA

Ir. T. Parulian, MSi

Pokok Bahasan

- **Pengertian**
- **Bentuk & Jenis Persamaan Regresi**
- **Ordinary Least Square**
- **Sum Square**
- **Koef. Determinasi**
- **Koef. Korelasi**
- **Standar Error**
- **F hitung**
- **t hitung**
- **Uji koefisien regresi**
- **Tabel ANOVA**

ANALISIS REGRESI

Tools yang sangat terkenal dalam statistika terapan untuk penelitian bidang ekonomi, sdm, ekonometrika, pemasaran dan bidang-bidang lain.

Diperkenalkan tahun 1886 oleh Francis Galton.

Analisis Regresi

Analisis ketergantungan dari satu atau lebih variabel bebas terhadap satu variabel terikat, dengan tujuan untuk memprediksi nilai rata-rata populasi berdasar nilai-nilai variabel bebasnya.

Sering digunakan sebagai salah satu alat analisis untuk membuat proyeksi melalui koefisien regresinya.

Analisis regresi sederhana :

Memprediksi satu variabel terikat berdasar pada satu variabel bebas.

Analisis regresi berganda:

Memprediksi satu variabel terikat berdasar pada dua atau lebih variabel bebas.

Perbedaan Analisis Korelasi dan Regresi

Analisis Korelasi hanya bertujuan mengukur kekuatan hubungan linier antar dua variabel (tidak membedakan antara variabel bebas dan variabel terikat).

Analisis Regresi, selain mengukur kekuatan hubungan antara variabel, juga untuk menunjukkan arah hubungan antara variabel bebas dengan variabel terikatnya.

Perbedaan antara “hubungan kausalitas dan Identitas”

Model regresi dibangun berdasarkan hubungan sebab akibat (kausalitas). Penentuan mana variabel bebas dan variabel terikat didasarkan kepada teori, penelitian sebelumnya atau berdasar pada penjas logis.

Hubungan identitas merupakan bentuk hubungan yang bukan disebabkan adanya fenomena sebab akibat, tetapi disebabkan oleh persamaan yang telah dibentuk.

Contoh Model regresi

- Hubungan antara biaya promosi dengan volume penjualan.
- Hubungan antar upah dengan produktivitas.
- Hubungan antara harga barang dengan jumlah yang diterima.

Contoh hubungan identitas

- Hubungan antara likuiditas dengan aktiva lancar.
Likuiditas = aktiva lancar / hutang lancar
- Hubungan antar produktivitas dengan hasil produksi.
Produktivitas = hasil produksi / jam kerja

Bentuk – bentuk Estimasi :

$Y = a + bX + e$: regresi sederhana

$Y = a + b_1X_1 + b_2X_2 + e$: regresi berganda

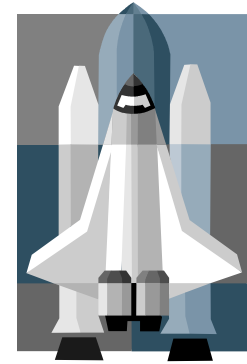
$Y = a + bX_1 + dD + e$: regresi berganda

- Y: variabel terikat (dependent) atau yang dipengaruhi.
- X,D: variabel bebas (independent) atau explanatory/penjelas/ yang mempengaruhi.
- D : variabel boneka (dummy).
- a : intercept (konstanta).
- b : koefisien regresi.
- e : error term (faktor pengganggu) atau residu.

Jenis-jenis Persamaan Regresi

- Regresi Linear Sederhana.
- Regresi Linier Berganda.
- Regresi Dummy : Variabel bebasnya berupa data nominal atau ordinal.
- Regresi Logistik : Variabel terikatnya berupa data nominal.
- Regresi Multinomial : Variabel terikat berupa data nominal atau ordinal yang lebih dari dua.

Model Regresi Sederhana

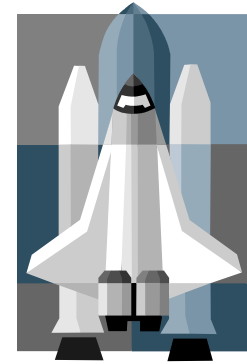


Bentuk sebenarnya dari persamaan populasi :

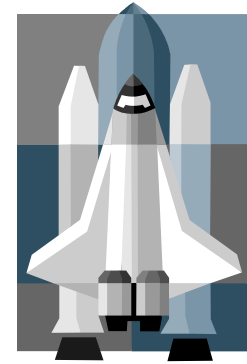
$$Y = A + BX + e$$

dapat ditaksir dengan persamaan dugaan sampel

$$Y' = a + bX.$$



Untuk menentukan koefisien regresi a dan b dilakukan analisis regresi dengan metode **OLS** (Ordinary Least Square).



Dari metode OLS, diperoleh koefisien persamaan regresi sebagai berikut :

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

$$a = \bar{Y} - b \bar{X}$$

Sehingga diperoleh persamaan regresi : $Y' = a + bX$.

Dan selanjutnya dapat dituliskan: $Y = Y' + e$.

Y = nilai Y sebenarnya

Y' = nilai Y yang diduga

\bar{Y} = nilai rata-rata Y

Sum Square (Jumlah Kuadrat)

$$SSE = \sum (Y - Y')^2$$

SSE = Sum Square Error (Keragaman yang tidak terjelaskan)

$$SSR = \sum (Y' - \bar{Y})^2$$

SSR = Sum Square Regression (Keragaman yang terjelaskan)

$$SST = \sum (Y - \bar{Y})^2$$

SST = Sum Square Total

$$SST = SSE + SSR$$

Koefisien Determinasi (R^2)

Koefisien determinasi menunjukkan besarnya keragaman variabel terikat yang dapat dijelaskan oleh variabel bebas.

$$R^2 = \frac{SST - SSE}{SST} = \frac{SSR}{SST}$$

$$\text{atau : } R^2 = 1 - \frac{\sum (Y - Y')^2}{\sum (Y - \bar{Y})^2}$$

$(Y - Y')^2$ Kuadrat selisih nilai Y riil dengan nilai Y prediksi

$(Y - \bar{Y})^2$ Kuadrat selisih nilai Y riil dengan nilai Y rata-rata

Adjustment Koefisien Determinasi (R^2_{adj})

Koefisien determinasi bias terhadap jumlah variabel bebas dalam model, dimana setiap penambahan satu variabel bebas dan pengamatan, akan meningkatkan nilai R^2 , meskipun variabel yang dimasukkan tidak memiliki pengaruh yang signifikan terhadap variabel terikatnya.

Untuk mengurangi kelemahan tersebut, maka digunakan koefisien determinasi yang dikoreksi atau yang disesuaikan.

Sehingga dengan menggunakan nilai koefisien determinasi adjustment, maka nilai koefisien determinasi dapat naik turun akibat adanya penambahan variabel baru dalam model.

Formula Koefisien Determinasi yang dikoreksi (R^2_{adj}) :

$$R^2_{adj} = R^2 - \frac{P(1 - R^2)}{n - P - 1}$$

P = jumlah variabel bebas

n = ukuran sampel

R^2 = koefisien determinasi

Koefisien Korelasi (r)

Koefisien korelasi adalah ukuran keeratan antara dua variabel.

$$r = \sqrt{R^2}$$

Standar Error Estimasi (Kesalahan baku pendugaan)

Ukuran terhadap pencaran nilai-nilai pengamatan di sekitar garis regresi. Merupakan ukuran ketepatan pendugaan.

$$se_{Y.X} = \sqrt{\frac{\sum (Y - Y')^2}{n - k}} = \sqrt{\frac{SSE}{n - k}}$$

k = (jumlah variabel bebas + 1)

n = jumlah data

Rumus lain menentukan kesalahan baku persamaan regresi .

$$Se_{Y.X} = \sqrt{\frac{\sum Y^2 - a(\sum Y) - b(\sum XY)}{n - k}}$$

Standar Error Koefisien Regresi

Untuk mengukur besarnya penyimpangan dari masing-masing koefisien regresi.

Semakin rendah kesalahan bakunya, semakin berperan variabel tersebut dalam model dan sebaliknya.

Jika model estimasi : $Y = a + bx + e$, kesalahan baku koefisien regresi b adalah :

$$Sb = \frac{Se_{Y.X}}{\sqrt{\sum (X^2) - \frac{(\sum X)^2}{n}}}$$

atau :

$$Sb = \frac{Se_{Y.X}}{\sqrt{\sum (X - \bar{X})^2}}$$

Sb = kesalahan baku koefisien regresi.

Se = kesalahan baku estimasi .

$\sum X$ = jumlah nilai variabel bebas.

n = ukuran sampel

Uji Partial

Nilai t hitung digunakan untuk menguji apakah variabel bebas berpengaruh secara signifikan terhadap variabel terikat.

$$t_{hitung} = \frac{b}{Sb}$$

t_{tabel} diperoleh dari tabel t yaitu $t_{\alpha, (n-k)}$

b = koefisien regresi b .

Sb = kesalahan baku koefisien regresi.

Uji Koefisien Regresi

Ho : $b=0$

H1 : $b \neq 0$

Jika t hitung $<$ t tabel : Terima Ho.

Jika t hitung $>$ t tabel : Terima H1.

Terima Ho (t hitung $<$ t tabel) : variabel bebas tidak memiliki pengaruh yang signifikan (berarti) terhadap variabel terikat pada α tertentu.

Terima H1 (t hitung $>$ t tabel) : variabel bebas mempengaruhi variabel terikat secara signifikan pada α tertentu.

Uji Global

Nilai F hitung digunakan untuk menguji apakah model fit (tepat) atau tidak, sehingga nilai F disebut ukuran goodness of fit.

Uji F digunakan untuk menguji apakah variabel bebas dalam model, mampu menjelaskan perubahan nilai variabel terikat (uji simultan).

Jika nilai F hitung $>$ F tabel, maka disimpulkan bahwa model persamaan regresi yang terbentuk, masuk kriteria cocok atau fit.

$$F_{hitung} = \frac{R^2 / (k - 1)}{1 - R^2 / (n - k)}$$

F_{tabel} diperoleh dari tabel F yaitu $F_{df} = F_{(k-1), (n-k)}$

k = jumlah variabel bebas + 1

n = ukuran sampel

Uji Global (Serentak)

Jika $F_{hitung} > F_{tabel}$

Variabel bebas secara bersama-sama berpengaruh secara signifikan terhadap variabel terikat pada α tertentu.

Jika $F_{hitung} < F_{tabel}$

Variabel bebas secara bersama-sama tidak berpengaruh secara signifikan terhadap variabel terikat pada α tertentu.

Tabel ANOVA

Merupakan alat atau piranti yang dapat menggambarkan hubungan antara koefisien korelasi, koefisien determinasi dan standard error pendugaan.

Contoh tabel ANOVA dengan 1 variabel bebas

| Source | df | SS | MS | F |
|------------|---------|-----|-----------------|---------------|
| Regression | 1 | SSR | $MSR = SSR/1$ | $F = MSR/MSE$ |
| Error | $n-k$ | SSE | $MSE = SSE/n-k$ | |
| Total | $n-k+1$ | SST | | |

Dari tabel diperoleh nilai : R^2 , r , Se , F hitung

Latihan

Besarnya biaya sales promotion dan besarnya volume penjualan barang yang dipromosikan dari suatu perusahaan selama 5 semester, ditunjukkan sebagai berikut :

| Semester | Biaya Sales Promotion | Jumlah Penjualan |
|----------|-----------------------|------------------|
| 1 | 2 | 4 |
| 2 | 3 | 4 |
| 3 | 4 | 5 |
| 4 | 5 | 7 |
| 5 | 6 | 5 |

Ditanya:

- Tentukan persamaan regresi pengaruh biaya sales promotion terhadap jumlah penjualan.
- Hitung standar error estimasi.
- Uji koefisien regresi dari biaya sales promotion pada α 5%.

| Sem | X | Y | XY | X ² | Y' | (Y-Y') ² |
|-----|----|----|-----|----------------|-----|---------------------|
| 1 | 2 | 4 | 8 | 4 | 4,0 | 0 |
| 2 | 3 | 4 | 12 | 9 | 4,5 | 0,25 |
| 3 | 4 | 5 | 20 | 16 | 5,0 | 0 |
| 4 | 5 | 7 | 35 | 25 | 5,5 | 2,25 |
| 5 | 6 | 5 | 30 | 36 | 6,0 | 1,0 |
| Σ | 20 | 25 | 105 | 90 | 25 | 3,50 |

$$b = \frac{n \sum XY - \sum X \sum Y}{n \sum X^2 - (\sum X)^2} = 0,5$$

$$a = \bar{Y} - b \bar{X} = 3$$

Diperoleh model estimasi : $Y' = 3 + 0,5X$

Ketelitian pendugaan (Standar error estimasi) atau kesalahan baku pendugaan

$$Se_{Y.X} = \sqrt{\frac{\sum (Y - Y')^2}{n - k}} = \sqrt{\frac{3,5}{5 - 2}} = 1,08$$

Tentukan standar error koefisien regresi

$$Sb = \frac{Se_{Y.X}}{\sqrt{\sum (X^2) - \frac{(\sum X)^2}{n}}}$$

$$Sb = \frac{1,08}{\sqrt{90 - 80}} = 0,342$$

Uji koefisien regresi

Hipotesis

$$H_0 : b = 0$$

$$H_1 : b \neq 0$$

Untuk sampel kecil ($n < 30$) , gunakan t statistik.

$$t \text{ tabel} = t_{5\%/2 ; df} = t_{5\%/2 ; 5-2} = 3,18$$

Menentukan t hitung

$$t_{hitung} = \frac{b}{Sb} = \frac{0,5}{0,342} = 1,46$$

t hitung < t tabel : Ho diterima

Kesimpulan :

Tidak ada hubungan yang signifikan antara sales promotion dengan jumlah penjualan pada taraf signifikan 5%.

Contoh No. 1

Data hipotetik pengaruh pupuk terhadap produksi padi adalah sebagai berikut:

| Y | X |
|----|---|
| 20 | 2 |
| 30 | 3 |
| 30 | 5 |
| 10 | 1 |
| 60 | 9 |

Y = produksi (Ton)

X = pupuk (kg)

Ditanya:

- Buat model estimasi pengaruh pupuk terhadap produksi padi.
- Hitung standar error model estimasi.
- Berapa produksi padi meningkat jika pemberian pupuk dinaikkan 10 kg.
- Berikan interpretasi dari model estimasi yang diperoleh.

Contoh No. 2

Persamaan regresi berikut diperoleh berdasarkan sampel 20 pengamatan.

$$Y' = 15 - 5X$$

Diperoleh juga nilai $SSE = 100$ dan $SST = 400$.

- Tentukan kesalahan baku pendugaannya.
- Tentukan koefisien determinasinya.
- Tentukan koefisien korelasinya.

Contoh No. 3

Direktur pemasaran sebuah perusahaan mempelajari penjualan bulanan dari 6 wilayah perusahaannya. Tiga variabel bebas dipilih sebagai peramal jumlah penjualan yaitu populasi daerah itu, pendapatan per kapita dan tingkat pengangguran daerah itu.

Persamaan regresi setelah dihitung (dalam dolar) adalah:

$$Y' = 64.100 + 0,394 X1 + 9,6 X2 - 11600 X3$$

- Terangkan, apakah arti angka 64.100
- Berapa nilai pendugaan penjualan total bulanan untuk wilayah tersebut, jika wilayah tersebut berpenduduk 796.000, pendapatan perkapita \$6.940 dan tingkat pengangguran 6,0%.

Contoh No. 4

Sejumlah sampel janda manula dipelajari untuk mengetahui derajat kepuasan terhadap kehidupan mereka. Indeks khusus yang disebut indeks kepuasan, digunakan untuk mengukur kepuasan. Enam faktor dipelajari, yaitu umur mereka saat pertama menikah (X1), pendapatan per tahun (X2), jumlah anak (X3), nilai seluruh aset (X4), status kesehatan dalam bentuk indeks (X5), dan jumlah rata-rata aktivitas sosial yang diikuti tiap minggu, seperti bowling, dansa (X6). Persamaan regresi yang diperoleh adalah :

$$Y' = -16,24 + 0,017X1 + 0,0028 X2 + 42X3 + 0,0012 X4 + 0,19 X5 + 26,8 X6$$

- a. Tentukan pendugaan indeks kepuasan orang yang menikah pertama sekali pada umur 18 tahun, pendapatan per tahun \$26.500, mempunyai 3 anak, asetnya \$156.000, indeks status kesehatan 141 dan aktivitas sosial rata-rata 2,5 per minggu.
- b. Yang mana memberi tambahan lebih banyak terhadap kepuasan, apakah tambahan pendapatan \$10.000 per tahun atau tambahan 2 aktivitas sosial per minggu?

Contoh No. 5

Direktorat Pajak Departemen Keuangan telah melakukan penelitian mengenai hubungan pajak dengan tingkat pendapatan. Dalam penelitian itu ditetapkan bahwa pajak merupakan fungsi daripada pendapatan. Hasil estimasi dengan menggunakan data Cross Section untuk 32 Propinsi di Indonesia, ditunjukkan oleh hasil regresi sebagai berikut :

$$\text{Tax} = 0.150 + 0.234 \text{ Income}$$

(0.079) (0.005)

$$R^2 = 0.997$$

dimana : Tax = Total pajak yang dibayar (dalam milyar)
Income = Total pendapatan (dalam milyar)

Angka dalam kurung adalah nilai p value.

Pertanyaan :

- a. Dari hasil regresi di atas apakah tanda-tanda untuk koefisien regresi sudah sesuai dengan intuisi saudara. Jelaskan.
- b. Apa interpretasi dari koefisien income tersebut di atas.
- c. Apakah koefisien signifikan pada $\alpha = 1\%$. Jelaskan jawaban Saudara.

Contoh No. 6

Ketahanan pangan memerlukan produksi pangan yang cukup. Untuk meningkatkan produksi pangan, setiap tahun pemerintah memberikan kredit untuk pangan sebagai tambahan modal kerja.

Data produksi pangan dan kredit dalam enam tahun terakhir adalah sebagai berikut :

$$\begin{array}{lll} \Sigma Y = 30 & \Sigma X = 24 & \Sigma XY = 146 \\ \Sigma Y^2 = 176 & \Sigma X^2 = 124 & \end{array}$$

Y = produksi pangan (juta ton)

X = jumlah kredit (triliun rupiah)

Dari data tersebut, tentukan :

- Persamaan regresi produksi pangan ($Y = a + b X$).
- Standard error estimasi.
- Uji koefisien regresi pada taraf signifikan 10 %.

Contoh No. 7

Suatu penelitian untuk mengetahui hubungan antara saving per-capita (S_t) dan personal income per-capita (Y_t) dalam rupiah diberikan seperti model ekonomi berikut ini :

$$S_t = a + bY_t + U_t$$

Data yang digunakan dalam penelitian ini adalah 25 tahun.
Hasil estimasi model adalah sebagai berikut :

$$\begin{aligned} S_t &= 241,415 + 0,125 Y_t \\ \text{se} &\quad (120.204) \quad (0.015) \\ R^2 &= 0.616 \end{aligned}$$

Angka dalam kurung adalah standard error

Pertanyaan :

- Apakah interpretasi ekonomi dari koefisien a.
- Tanda apa yang saudara harapkan untuk a dan b. Apakah tanda tersebut sudah sesuai dengan intuisi dan apa yang dapat saudara simpulkan dari hasil tersebut.
- Seandainya saving dan income diukur dalam ribuan rupiah, apa yang dapat saudara komentari

Contoh No. 8

Diberikan tabel ANOVA berikut:

| Sumber | Df | SS | MS | F |
|-----------|-----|-----|-----|-----|
| Regresi | 1 | 50 | --- | --- |
| Kesalahan | --- | --- | --- | |
| Total | 24 | 500 | | |

- Lengkapi tabel ANOVA tersebut.
- Berapa ukuran sampelnya?
- Tentukan kesalahan baku pendugaannya.
- Tentukan koefisien determinasinya.

Contoh No. 9

$$\begin{array}{cccc} \text{JPA} = 3,0 + 2,5 \text{ JAK} + 5,0 \text{ Income} + 2,0 \text{ LL} \\ (\text{t}) & (3,41) & (1,45) & (2,95) \end{array}$$

$$n = 20$$

JPA = Jumlah Pemakaian Air (liter)

JAK = Jumlah Anggota Keluarga

Income = Pendapatan Kepala Keluarga (Juta)

LL = Luas Lantai (m²)

Uji apakah JAK, Income dan Luas Lantai berpengaruh terhadap JPA pada α 5%

Contoh No. 10

| Pekerja | Masa Kerja (X) | Produksi /Minggu (Y) |
|---------|----------------|----------------------|
| A | 14 | 6 |
| B | 7 | 5 |
| C | 3 | 3 |
| D | 15 | 9 |
| E | 11 | 7 |

- Tentukan persamaan regresi dari $Y=a+bX$
- Tentukan besarnya SSR, SSE dan SST
- Tentukan nilai Koefisien Determinasi dan Korelasi
- Beri interpretasi