

ANALISIS KORELASI DAN REGRESI

Achmad Samsudin, M.Pd.
Jurusan Pendidikan Fisika FPMIPA
Universitas Pendidikan Indonesia

ANALISIS KORELASI

- Analisis yang digunakan untuk menelaah tingkat keeratan hubungan 2 variabel (x dan y)
- X : Variabel bebas
- Y : Variabel tidak bebas (Peubah Respon)

• Bentuk Hubungan x dan y :

Hubungan + : Jika kenaikan / penurunan x pada umumnya diikuti oleh kenaikan/ penurunan y

Hubungan - : Jika kenaikan / penurunan x pada umumnya diikuti oleh penurunan / kenaikan y

Konsep Korelasi

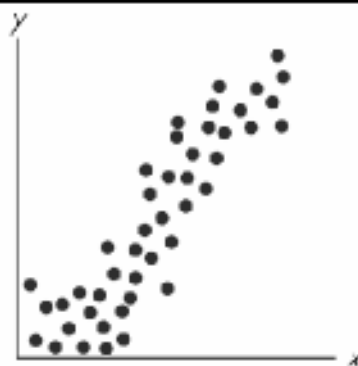
- **Korelasi**, juga disebut **koefisien korelasi**, adalah ukuran numerik yang menunjukkan kekuatan dan arah hubungan linier antara dua peubah acak (*random variable*).
- Nilai korelasi berkisar antara:

$$-1 \leq \rho \leq 1$$

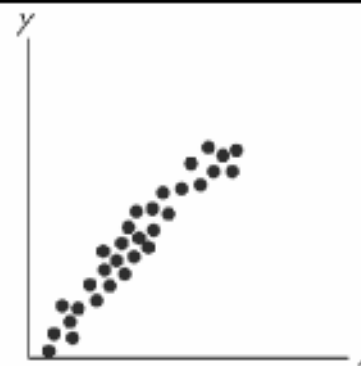
Korelasi tinggi	Tak ada korelasi (acak)	Korelasi tinggi
-1	0	+1

Konsep Korelasi

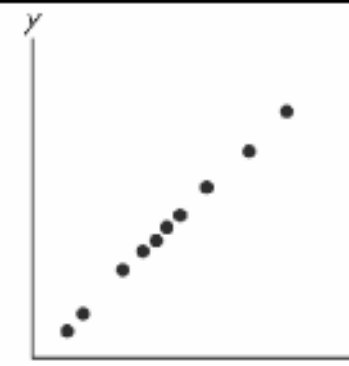
- Terdapat banyak metode korelasi antara lain korelasi **pearson product moment**, korelasi biserial, korelasi point biserial, korelasi spearman, korelasi tau kendall.
- Hal yang membedakan metode-metode tersebut adalah skala pengukuran data variabel-variabel yang akan dicari hubungannya, serta jumlah variabel yang akan dianalisis.
- **Korelasi pearson** mempunyai tujuan untuk mencari korelasi antara dua variabel yang berskala interval atau rasio.
- Korelasi spearman dan tau kendall digunakan untuk mencari korelasi antara dua variabel yang berskala nominal atau ordinal.
- Korelasi Biserial & Korelasi point biserial akan dijelaskan di bagian selanjutnya



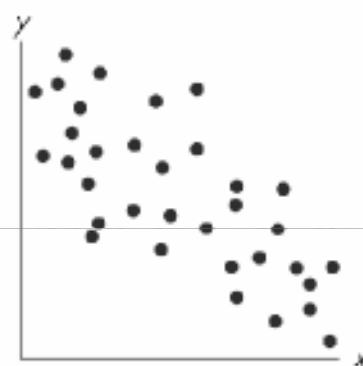
(a) Positive correlation between x and y



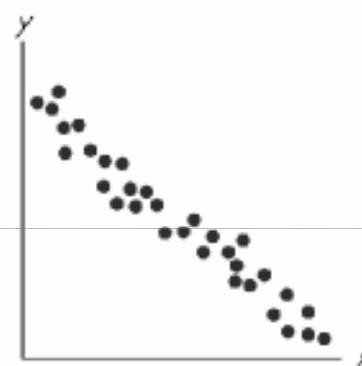
(b) Strong positive correlation between x and y



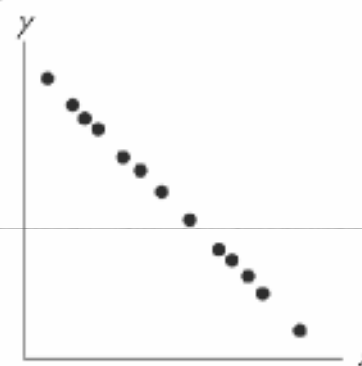
(c) Perfect positive correlation between x and y



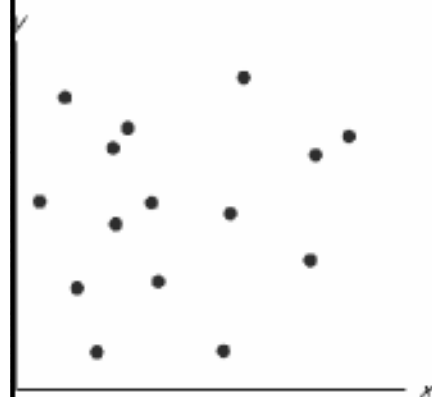
(d) Negative correlation between x and y



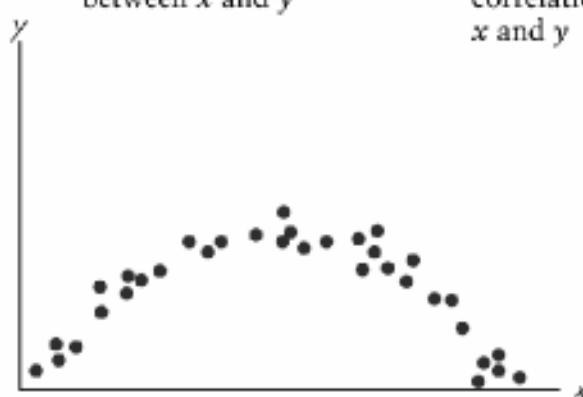
(e) Strong negative correlation between x and y



(f) Perfect negative correlation between x and y



(g) No correlation between x and y



(h) Nonlinear relationship between x and y

Korelasi Pearson Product-Moment

Kasus :

- Manajer personalia perusahaan *infotainment*, tertarik untuk mengetahui apakah terdapat hubungan antara skor tes bakat dengan performans kerja. Diambil sampel 10 orang karyawan baru dari berbagai divisi di perusahaan tersebut. Dari ke-10 orang tersebut diambil data skor tes bakat beserta performansi kerjanya, ditampilkan sebagai berikut :

Karyawan	1	2	3	4	5	6	7	8	9	10
Tes bakat	65	60	62	59	58	53	50	48	45	40
Performansi	90	95	82	87	80	75	60	69	60	72

- Menggunakan $\alpha = 5\%$, apakah terdapat hubungan antara kedua variabel tersebut ?

• Korelasi Spearman

• Kasus:

Dari kasus sebelumnya. Asisten Prof. Santi juga ingin mengetahui hubungan antara kelulusan post test dengan jenis kelamin. Data yang dimiliki dapat dilihat pada tabel di bawah ini. Lakukan analisa dengan menggunakan $\alpha = 5\%$, buatlah kesimpulan apakah terdapat hubungan yang signifikan antara nilai tes awal dalam mata kuliah Statistik dengan jenis kelamin.

• Keterangan nilai post test:

1 : tidak lulus

2 : lulus

• Keterangan nilai jenis kelamin:

1 : wanita

2 : pria

Post	sex
1	1
1	1
1	2
1	1
1	1
1	2
1	2
1	2
2	1
1	2
1	2
1	2
1	2
2	1
2	1
2	1
1	1
1	1
1	1
2	2
1	2

Korelasi Biserial & Point Biserial

Korelasi Biserial digunakan untuk mencari korelasi antara dua variabel berskala interval atau rasio, namun disebabkan oleh satu & lain hal salah satu variabel tidak dapat dicari datanya dalam bentuk asli, melainkan data tersebut telah diubah menjadi dua kategori saja.

$$r_b = \frac{(\bar{X}_p - \bar{X}_q) \times p \times q}{S_y}$$

Korelasi point biserial merupakan metode yang digunakan untuk mencari derajat keeratn hubungan antara dua variabel, dimana salah satu variabel ber-skala interval atau rasio sedangkan variabel yang lainnya berskala nominal atau ordinal dengan dua kategori saja.

$$r_{pb} = \frac{(\bar{X}_p - \bar{X}_q) \times \sqrt{p \times q}}{S_y}$$

Korelasi Biserial

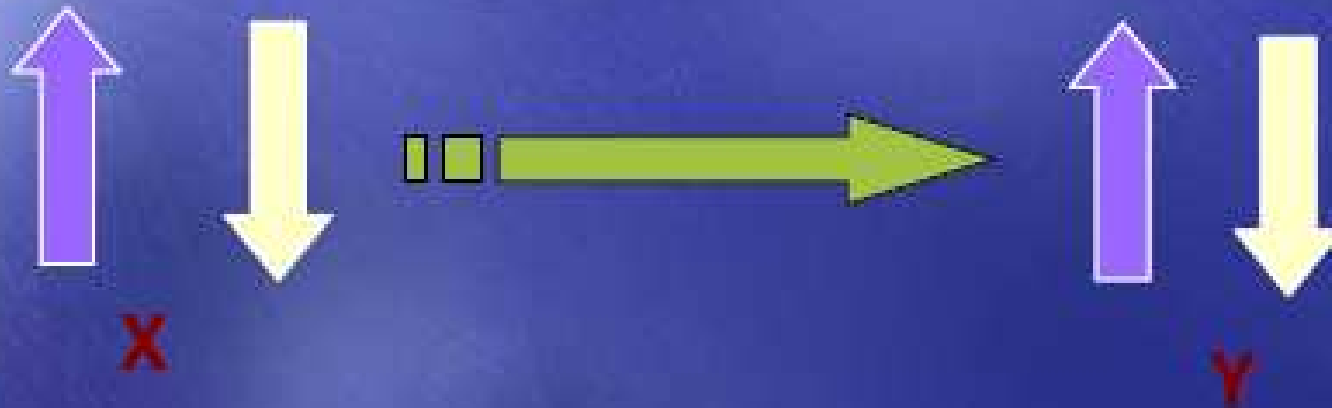
- **Kasus:**

Seorang mahasiswa fakultas psikologi menjadi asis-en untuk mata kuliah statistik yang diasuh oleh Prof. Santi. Diawal semester sebelum menerima materi, para mahasiswa diberi tes kecil, untuk mengukur kemampuan matematika. Sebagai seorang asisten ia diminta untuk memeriksa hasil tes tersebut. Dengan demikian ia memiliki nilai tes awal dari setiap mahasiswa di kelas tersebut. Sebagai seorang mahasiswa sang asisten tertarik untuk mengetahui hubungan antara nilai tes awal dengan nilai tes akhir. Akan tetapi ia menemui kendala yaitu ia tidak bisa memperoleh data nilai ujian akhir, karena ujian akhir diperiksa sendiri oleh Prof. Santi. Akan tetapi dari hasil wawancara ia mengetahui siapa saja yang lulus dan siapa saja yang tidak lulus dari mata kuliah Statistik. Data yang didapat dari 200 orang mahasiswa adalah sebagai berikut:

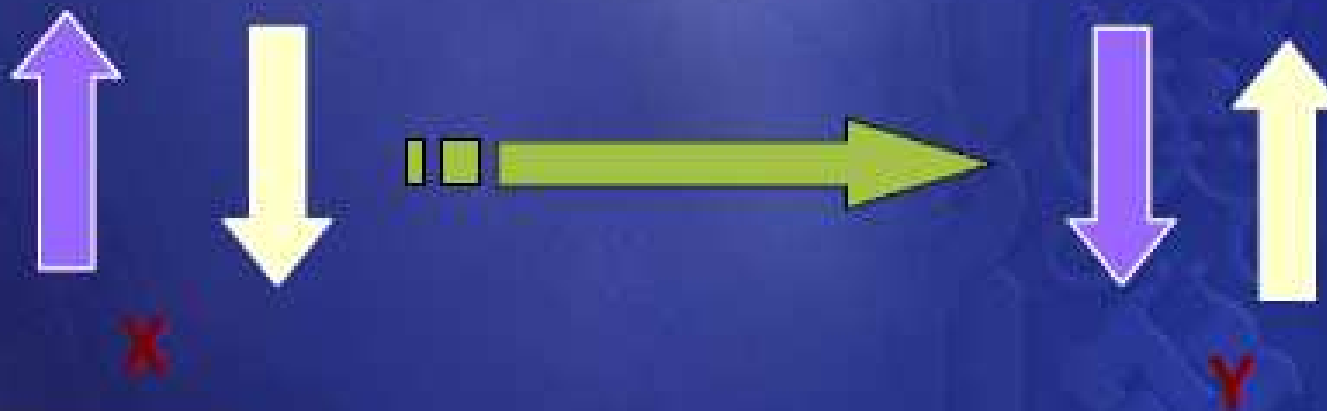
Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
47	1	73	1	79	1	84	1	87	2
51	1	74	1	79	1	84	2	87	2
57	1	76	1	80	1	84	2	87	2
62	1	76	1	80	1	84	2	89	2
62	1	75	1	80	1	84	2	89	2
63	1	76	1	80	1	86	2	89	2
65	1	76	1	80	1	85	2	89	2
66	1	76	1	81	1	86	2	89	2
67	2	76	2	81	1	85	2	89	2
67	1	76	2	81	1	86	2	89	2
67	1	76	2	81	1	86	2	89	1
69	1	77	2	81	1	86	2	89	1
69	1	77	2	82	1	86	2	89	1
69	2	77	2	82	1	86	2	89	1
70	2	78	2	82	1	86	2	90	1
70	2	78	2	82	1	87	2	90	1
71	1	78	2	82	1	87	2	90	1
71	1	79	2	83	1	87	2	90	1
72	1	79	1	83	1	87	2	90	1
119	2	119	2	119	2	119	2	124	2
91	1	95	2	98	1	102	2	106	2
91	1	96	2	99	1	102	2	106	2
91	1	96	2	99	1	103	2	108	2
91	1	96	2	99	1	103	2	108	2
92	1	96	2	99	1	103	2	108	2
92	1	96	2	99	1	103	2	108	2
92	1	96	2	99	2	104	2	110	2
92	2	96	2	99	2	104	2	111	2
92	2	96	2	99	2	104	2	112	2
92	2	96	2	99	2	104	2	112	2
92	2	97	2	99	2	104	2	112	2
93	2	97	2	99	2	105	2	112	2
93	2	97	2	99	2	105	2	114	2
93	2	97	2	100	2	105	1	114	2
94	2	98	2	100	2	105	1	115	2
94	2	98	2	100	2	105	1	115	2
94	2	98	2	101	2	105	2	115	2
95	2	98	2	101	2	106	2	116	2
95	2	98	1	102	2	106	2	116	2
125	2	125	2	125	2	129	2	132	2

Data Pre test adalah skor tes awal. Data Post test adalah kelulusan dimana angka 1 menunjukkan bahwa mahasiswa tersebut tidak lulus dari mata kuliah Statistik, sedangkan angka 2 menunjukkan bahwa mahasiswa yang bersangkutan lulus dari mata kuliah Statistik. Dengan menggunakan $\alpha = 5\%$, apakah terdapat hubungan yang signifikan antara nilai tes awal dengan kelulusan dalam mata kuliah Statistik?

Hubungan +

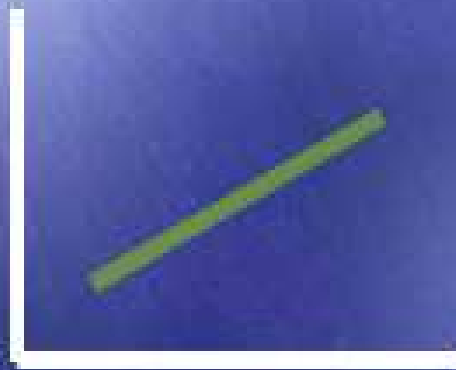


Hubungan -

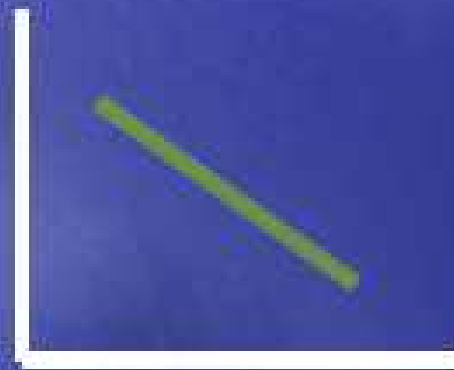


CONTOH.....

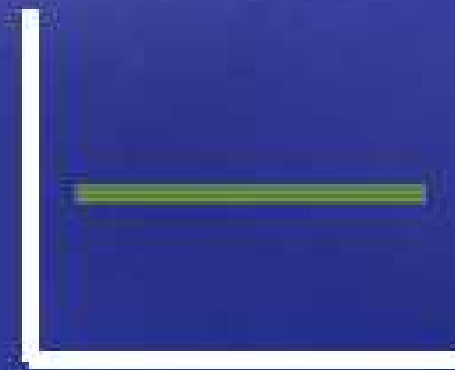
Pola Hubungan X dan Y



Hub +



Hub -



Tidak ada hub

Koefisien korelasi (r)

• Adalah nilai yang menunjukkan kuat atau lemah hubungan antara x dan y

Nilai ini (r) terletak antara -1 dan 1

$$-1 \leq r \leq 1$$

$$-1 \leq r \leq 1$$

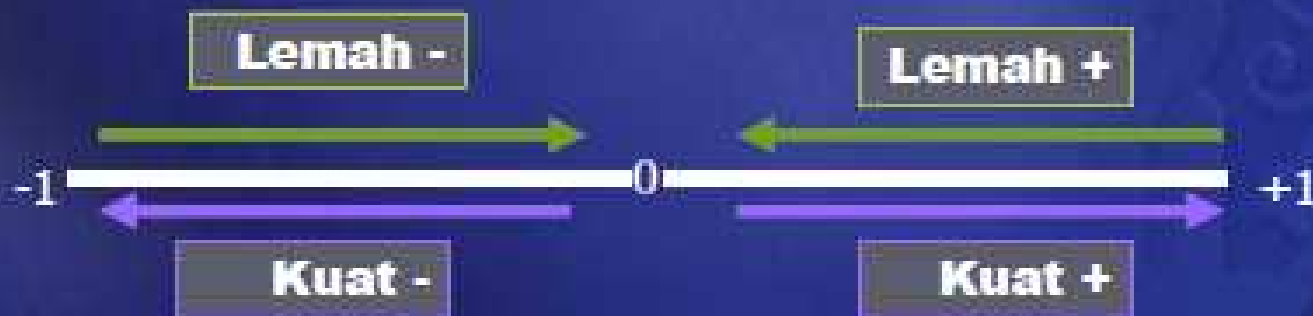
$r = 1$: Hubungan x dan y sempurna dan +

$r = -1$: Hubungan x dan Y sempurna dan -

$r = 0$: tidak ada hubungan antara x dan y

$r \sim \pm 1$: Hubungan x dan y sangat kuat dan positif atau negatif

$r \sim 0$: Hubungan x dan y sangat lemah dan positif atau negatif



Nilai r dihitung dengan persamaan :

$$r = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}}$$

- n : Banyaknya Data
- Xi : Peubah bebas ke-i
- Yi : Peubah respon ke-i

Koefisien Determinasi

Besarnya Kontribusi x terhadap naik turunnya y

$$R = r^2$$

ANALISIS REGRESI

Adalah analisis yang menelaah hubungan antara peubah respon Y dengan satu atau lebih peubah bebas X_i , yang dinyatakan dalam bentuk

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \epsilon_i$$

- Y_i : Peubah respon ke- i
- X_i : Peubah bebas ke- i
- β_i : Konstanta ke- i
- ϵ_i : Sisaan ke- i

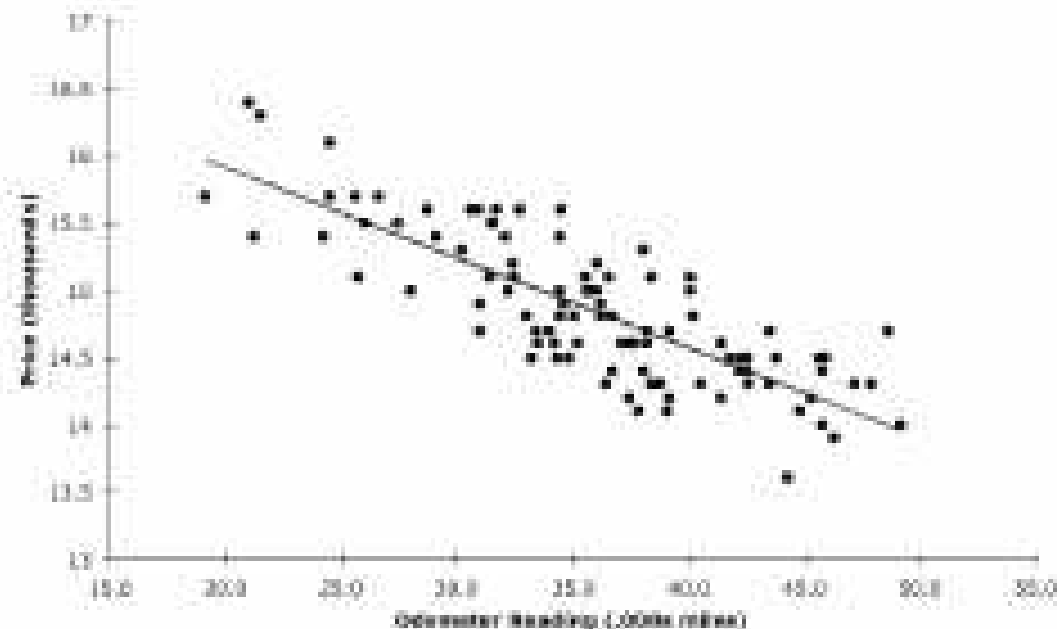
Analisis Regresi

Analisis regresi adalah metode statistik yang digunakan untuk menggambarkan pola hubungan antara sebuah variabel dependen dengan satu atau beberapa variabel independen (prediktor)

Regresi linier sederhana melibatkan hanya satu variabel independen

Regresi linier berganda melibatkan lebih dari satu variabel independen

Ex. 17.2 - Odometer Line Fit Plot



ANALISIS REGRESI LINEAR SEDERHANA

- ▶ Analisis yang membahas hubungan antara 2 variabel (y dengan satu peubah bebas x) yang biasanya terletak dalam suatu garis lurus
- ▶ Pola hubungan x dan y dapat dilihat dari diagram pencar. Manfaat dari diagram ini:
 1. Melihat apakah ada hubungan yang bermanfaat antara 2 variabel
 2. Membantu menetapkan tipe persamaan yang menunjukkan hubungna kedua variable tersebut

Persamaan regresi linear sederhana :

$$y = \alpha + \beta x + \varepsilon$$

$$\hat{y} = a + bx$$



Interpretasi persamaan

Setiap perubahan x sebesar 1 satuan maka y akan berubah sebesar b

dimana :

$$b = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2}$$

dan

$$a = \bar{y} - b\bar{x}$$

$$a = \frac{\sum y_i}{n} - b \frac{\sum x_i}{n}$$

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

Penggunaan persamaan regresi dalam peramalan

- ▶ Dengan menggunakan persamaan regresi

$$\hat{y} = a + bx$$

Kita bisa memprediksi nilai Y (variabel tidak bebas) pada nilai x tertentu (x : Variabel bebas)

Regresi linier sederhana

Kasus :

Dekan fakultas Psikologi universitas X, sedang membuat perencanaan kurikulum. Ada keinginan untuk meningkatkan mahasiswa di mata kuliah utama agar nantinya setelah lulus, mereka mendapatkan pekerjaan dengan gaji yang cukup baik. Untuk mendukung usaha ini, ia ingin memperlihatkan dari data lulusan selama 7 tahun terakhir bahwa rata-rata nilai pada mata kuliah utama mempengaruhi rata-rata gaji awal yang diterima oleh para alumni itu. Berikut adalah datanya

Nilai	2.58	3.27	3.85	3.50	3.33	2.89	2.23
Gaji (\$ ribu)	16.5	18.8	19.5	19.2	18.5	16.6	15.6

Menggunakan $\alpha = 5\%$, lakukan analisis regresi untuk data diatas

Hipotesis serentak:

H_0 : regresi tidak signifikan

H_1 : regresi signifikan

Hipotesis individu :

$H_0 : \beta_0 = 0$

$H_1 : \beta_0 \neq 0$

$H_0 : \beta_1 = 0$

$H_1 : \beta_1 \neq 0$

Regresi linier berganda

Kasus:

- Dekan fakultas Psikologi universitas X (lihat kasus regresi linier sederhana) memberikan pelatihan organisasi dan kepemimpinan kepada mahasiswanya di tingkat akhir. Dekan memperoleh masukan bahwa rata-rata gaji pertama lulusan juga dipengaruhi oleh nilai pelatihan tersebut. Berdasarkan masukan ini, kembali akan dibuat analisis regresi yang juga melibatkan variabel rata-rata nilai pelatihan kepemimpinan dan organisasi

Nilai MK	2.58	3.27	3.85	3.50	3.33	2.89	2.23
Gaji (\$ ribu)	16.5	18.8	19.5	19.2	18.5	16.6	15.6
Nilai pelatihan	75	86	93	80	82	76	70

Menggunakan $\alpha = 5\%$, lakukan analisis regresi untuk data diatas

Hipotesis serentak:

H_0 : regresi tidak signifikan

H_1 : regresi signifikan

Hipotesis individu :

$$H_0 : \beta_0 = 0$$

$$H_1 : \beta_0 \neq 0$$

$$H_0 : \beta_1 = 0$$

$$H_1 : \beta_1 \neq 0$$

$$H_0 : \beta_2 = 0$$

$$H_1 : \beta_2 \neq 0$$

CONTOH

(Regresi Linear Sederhana)

- ▶ Sebuah penelitian dilakukan oleh seorang pedagang eceran untuk menentukan hubungan antara biaya pemasangan iklan perminggu dengan hasil penjualannya. Data yang diperoleh adalah sebagai berikut

Biaya Iklan (\$)	Penjualan (\$)
40	385
20	400
25	395
20	365
30	475
50	440
40	490
20	420

Pertanyaan

1. Periksalah apakah ada hubungan antara biaya iklan dan penjualan
2. Seberapa erat hubungan tersebut

LATIHAN

- Seorang dokter ingin mengetahui apakah ada hubungan antara berat badan seseorang dengan tinggi badan seseorang, untuk keperluan tsb dilakukan penelitian terhadap 10 orang dengan data sbb:

Tinggi (cm)	Berat Badan (kg)
140	42
132	39
136	52
143	47
156	41
158	68
161	46
171	48
166	57
160	62

Buat persamaan regresinya

ANALISIS KORELASI

- Untuk melihat tingkat hubungan antar dua variabel
- Dinyatakan oleh nilai Koefisien Korelasi (r)
→ kekuatan hubungan
- Kekuatan Determinasinya: $R = r^2$ (persen)
- Jenis-jenis korelasi:
 - a. Korelasi positif
 - b. Korelasi negatif
 - c. Tidak ada Korelasi

The background of the slide is a dark brown, textured surface with a repeating pattern of small, irregular shapes, resembling a marbled or embossed paper texture.

SELESAI & TERIMA KASIH

Syam 2009