

KORELASI

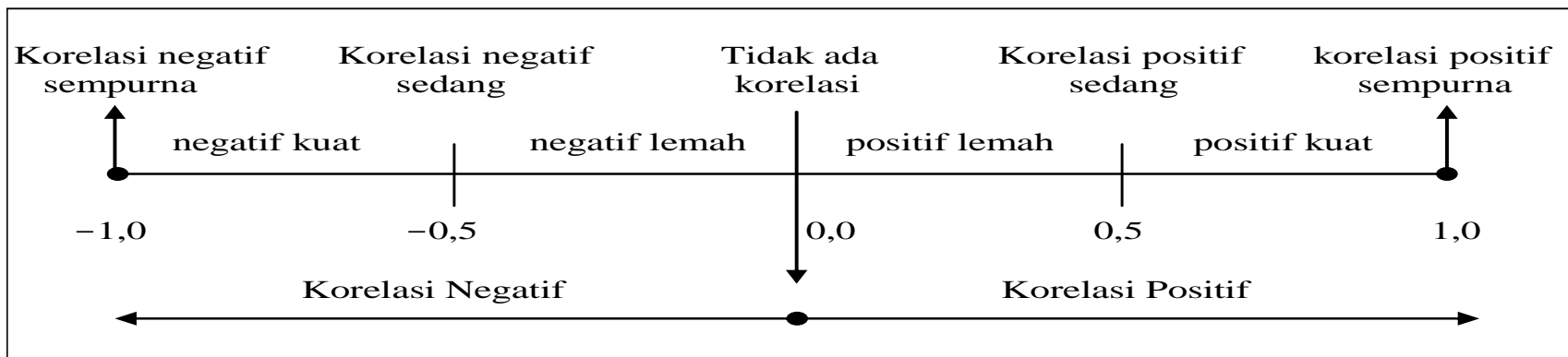
Tim Pengajar Statistika Dasar

Pengertian Korelasi

- Analisis bivariat yang mengukur kekuatan hubungan antara dua variabel dan arah hubungannya
- Teknik statistik yang dapat menunjukkan jika terdapat dan seberapa kuat pasangan variabel terkait

Apa fungsi dari koefisien korelasi?

- Nilai r selalu antara -1 dan $+1$ yaitu :
 $-1 \leq r \leq 1$.
- Besarnya korelasi r menunjukkan kekuatan hubungan linear antara x dan y . Nilai r mendekati -1 atau $+1$ menunjukkan hubungan linier yang lebih kuat antara x dan y .
- Jika $r = 0$ sama sekali tidak ada hubungan linier antara x dan y (tidak ada korelasi linier).
- Jika $r = 1$, ada korelasi positif sempurna. Jika $r = -1$, ada korelasi negatif sempurna. Dalam kedua kasus ini, semua titik data asli terletak pada garis lurus. Tentu saja, di dunia nyata, hal ini umumnya tidak akan terjadi.



Apa fungsi dari koefisien korelasi?

- Nilai r positif berarti ketika x meningkat, y cenderung meningkat dan ketika x menurun, y cenderung menurun (korelasi positif).
- Nilai r negatif berarti ketika x meningkat, y cenderung menurun dan ketika x menurun, y cenderung meningkat (korelasi negatif).

Catatan penting :

- Korelasi yang kuat **tidak menunjukkan** bahwa x menyebabkan y atau y menyebabkan x , "korelasi tidak menyiratkan sebab-akibat."

Menguji Signifikansi Koefisien Korelasi

1. Kekuatan koefisien korelasi pengukuran tergantung pada berapa banyak titik data yang diamati dalam sampel. Jadi, kita perlu melihat nilai koefisien korelasi r dan ukuran sampel n secara bersamaan
2. Uji hipotesis tentang "signifikansi koefisien korelasi" harus dilakukan untuk memutuskan apakah hubungan linier dalam data sampel cukup kuat untuk digunakan untuk memodelkan hubungan dalam populasi.

Menguji Signifikansi Koefisien Korelasi

3. Hipotesis menyatakan apakah nilai koefisien korelasi populasi ρ "mendekati nol" atau "berbeda secara signifikan dari nol"
4. Jika pengujian menyimpulkan bahwa koefisien korelasi berbeda secara signifikan dari nol, kita katakan bahwa koefisien korelasi "signifikan"
5. Jika pengujian menyimpulkan bahwa koefisien korelasi tidak berbeda nyata dari nol (mendekati nol), maka koefisien korelasi dikatakan "tidak signifikan"
6. Untuk menguji signifikansi, kita melakukan "Uji Hipotesis"

Hipotesis Nol:

$$H_0 : \rho = 0$$

Artinya : Koefisien korelasi populasi TIDAK berbeda nyata dari nol. TIDAK ADA hubungan (korelasi) linier yang signifikan antara x dan y dalam populasi.

Hipotesis Alternatif :

$$H_a : \rho \neq 0$$

Artinya : Koefisien korelasi populasi BERBEDA nyata DARI nol. ADA HUBUNGAN LINEAR (korelasi) yang SIGNIFIKAN antara x dan y dalam populasi.

Jenis Perhitungan Koefisien Korelasi

Jenis Koefisien Korelasi	Penggunaan
Rank Spearman	Tipe Hubungan: Monoton (tetapi tidak harus linier) Tipe Data: Ordinal, atau data interval/rasio yang tidak normal atau memiliki outlier Distribusi Data: Tidak normal atau data ordinal
Pearson (Product Moment)	Tipe Hubungan: Linier Tipe Data: Interval atau rasio Distribusi Data: Normal atau mendekati normal
Ganda	Mengukur kekuatan hubungan antara satu variabel dependen dan beberapa variabel independen

Koefisien korelasi ini mengukur kedekatan hubungan antara dua variabel ordinal. Koefisien korelasi ini dinamakan koefisien korelasi pangkat atau koefisien korelasi Spearman, yang disimbolkan dengan r .

Pasangan data hasil pengamatan (X_i, Y_i) kita susun menurut urutan besar nilainya dalam tiap variabel. Kemudian kita bentuk selisih atau beda peringkat X_i dan peringkat Y_i yang data aslinya berpasangan. Beda ini disimbolkan dengan b_i , maka koefisien korelasi peringkat r dihitung dengan rumus:

$$r = 1 - \frac{6 \sum b_i^2}{n(n^2 - 1)}$$

Percobaan 1

Data berikut adalah penilaian 2 orang juri terhadap 8 orang peserta perlombaan.

Peserta	Juri I	Juri II
A	70	80
B	85	75
C	65	55
D	50	60
E	90	85
F	80	70
G	75	90
H	60	65

Tentukan koefisien korelasi rank dengan menggunakan semi otomatis di Ms. Excel !

Jawab

PESERTA	Juri I (X _i)	Juri II (Y _i)	Rank X _i	Rank Y _i	BEDA (b _i)	b _i ²
A	70	80	5	3	2	4
B	85	75	2	4	-2	4
C	65	55	6	8	-2	4
D	50	60	8	7	1	1
E	90	85	1	2	-1	1
F	80	70	3	5	-2	4
G	75	90	4	1	3	9
H	60	65	7	6	1	1
Jumlah (Σ b _i ²)						28

$$r = 1 - \frac{6 \sum b_i^2}{n(n^2 - 1)} = 1 - \frac{6(28)}{8(64-1)}$$

$$r = 1 - \frac{168}{504}$$

$$r = 1 - 0,333$$

$$r = 0,66667$$

Kesimpulan : Terdapat hasil korelasi **positif kuat** antara penilaian Juri 1 dan Juri 2

Koefisien Korelasi Pearson (Product Moment)

Simbol :

r atau ρ

ρ = koefisien korelasi populasi (tidak diketahui)

r = koefisien korelasi sampel (diketahui; dihitung dari data sampel)

Termasuk perhitungan numerik dan memberikan ukuran kekuatan dan arah hubungan linier antara dua variabel.

Koefisien Korelasi Pearson (Product Moment)

dihitung sebagai:

di mana n = jumlah titik data

$$r = \frac{n \sum X_i Y_i - (\sum X_i)(\sum Y_i)}{\sqrt{\{n (\sum X_i^2) - (\sum X_i)^2\} \{n (\sum Y_i^2) - (\sum Y_i)^2\}}}$$

Percobaan 2

Diketahui data jumlah SKS dan IPK mahasiswa sbb.

Jumlah SKS (X)	IPK (Y)
10	3,00
10	2,50
15	2,00
10	1,50
5	1,00

Tentukan nilai koefisien korelasi dengan metode product moment dan jelaskan artinya!

Jawab

No	X _i	Y _i	X _i Y _i	X _i ²	Y _i ²
1	10	3	30	100	9
2	10	2,5	25	100	6,25
3	15	2	30	225	4
4	10	1,5	15	100	2,25
5	5	1	5	25	1
Ket	ΣX _i	ΣY _i	ΣX _i Y _i	ΣX _i ²	ΣY _i ²
Jumlah	50	10	105	550	22,5

$$r = \frac{n \sum X_i Y_i - (\sum X_i)(\sum Y_i)}{\sqrt{\{n(\sum X_i^2) - (\sum X_i)^2\} \{n(\sum Y_i^2) - (\sum Y_i)^2\}}}$$

$$r = \frac{5(105) - (50)(10)}{\sqrt{\{5(550) - (50)^2\} \{5(22,5) - (10)^2\}}} = \frac{25}{\sqrt{(250)(12,5)}} = 0,447$$

Kesimpulan : Terdapat hasil korelasi **positif lemah** antara jumlah SKS dan IPK yang didapat

Untuk 2 variabel bebas (X_1 dan X_2) maka r dihitung dengan rumus:

$$r_{y x_1 x_2} = \sqrt{\frac{r_{y x_1}^2 + r_{y x_2}^2 - 2 r_{y x_1} r_{y x_2} r_{x_1 x_2}}{1 - r_{x_1 x_2}^2}}$$

dimana : $r_{y x_1 x_2}$ = Koefisien korelasi ganda antara variable X_1 dan X_2 secara bersama-sama dengan variable Y

$r_{y x_1}$ = Koefisien korelasi X_1 dengan Y

$r_{y x_2}$ = Koefisien korelasi X_2 dengan Y

$r_{x_1 x_2}$ = Koefisien korelasi X_1 dengan X_2

Percobaan 3

Misalkan kita melakukan pengamatan terhadap 10 keluarga mengenai:

X_1 = pendapatan dalam ribuan rupiah

X_2 = jumlah keluarga dalam satuan jiwa

Y = pengeluaran untuk membeli barang A dalam ratusan rupiah

X_1	10	2	4	6	8	7	4	6	7	6
X_2	7	3	2	4	6	5	3	3	4	3
Y	23	7	15	17	23	22	10	14	20	19

Akan dibuktikan ada hubungan linier positif dan signifikan antara variabel X_1 dan X_2 secara bersama-sama dengan variabel Y .

Jawab

No	X_1	X_2	Y	X_1Y	X_2Y	X_1X_2	X_1^2	X_2^2	Y^2
1	10	7	23	230	161	70	100	49	529
2	2	3	7	14	21	6	4	9	49
3	4	2	15	60	30	8	16	4	225
4	6	4	17	102	68	24	36	16	289
5	8	6	23	184	138	48	64	36	529
6	7	5	22	154	110	35	49	25	484
7	4	3	10	40	30	12	16	9	100
8	6	3	14	84	42	18	36	9	196
9	7	4	20	140	80	28	49	16	400
10	6	3	19	114	57	18	36	9	361
Ket.	ΣX_1	ΣX_2	ΣY	ΣX_1Y	ΣX_2Y	ΣX_1X_2	ΣX_1^2	ΣX_2^2	ΣY^2
Jumlah	60	40	170	1122	737	267	406	182	3162

- Koefisien Korelasi X_1 dengan Y

$$r_{y x_1} = \frac{n \sum X_1 Y - (\sum X_1)(\sum Y)}{\sqrt{\{n(\sum X_1^2) - (\sum X_1)^2\} \{n(\sum Y^2) - (\sum Y)^2\}}} =$$

$$r_{y x_1} = \frac{10(1122) - (60)(170)}{\sqrt{\{10(406) - (60)^2\} \{10(3162) - (170)^2\}}}$$

$$r_{y x_1} = \frac{1020}{\sqrt{460 \times 2720}} = \frac{1020}{1118,57}$$

$$r_{y x_1} = 0,912$$

- Koefisien Korelasi X_2 dengan Y

$$r_{y x_2} = \frac{n \sum X_2 Y - (\sum X_2)(\sum Y)}{\sqrt{\{n(\sum X_2^2) - (\sum X_2)^2\} \{n(\sum Y^2) - (\sum Y)^2\}}} =$$

$$r_{y x_2} = \frac{10(737) - (40)(170)}{\sqrt{\{10(182) - (40)^2\} \{10(3162) - (170)^2\}}}$$

$$r_{y x_2} = \frac{570}{\sqrt{220 \times 2720}} = \frac{570}{773,56}$$

$$r_{y x_2} = 0,74$$

- Koefisien Korelasi X_1 dengan X_2

$$r_{x_1 x_2} = \frac{n \sum X_1 X_2 - (\sum X_1)(\sum X_2)}{\sqrt{\{n(\sum X_1^2) - (\sum X_1)^2\} \{n(\sum X_2^2) - (\sum X_2)^2\}}} =$$

$$r_{x_1 x_2} = \frac{10(267) - (60)(40)}{\sqrt{\{10(406) - (60)^2\} \{10(182) - (40)^2\}}}$$

$$r_{x_1 x_2} = \frac{270}{\sqrt{460 \times 220}} = \frac{270}{318,12}$$

$$r_{x_1 x_2} = 0,85$$

$$\begin{aligned}
 r_{y x_1 x_2} &= \sqrt{\frac{r^2_{y x_1} + r^2_{y x_2} - 2 r_{y x_1} r_{y x_2} r_{x_1 x_2}}{1 - r^2_{x_1 x_2}}} \\
 &= \sqrt{\frac{(0,912)^2 + (0,74)^2 - 2 (0,912) (0,74) (0,85)}{1 - (0,85)^2}} \\
 &= \sqrt{0,8354} \\
 &= 0,914
 \end{aligned}$$

Kesimpulan :

Terdapat hubungan korelasi yang signifikan antara X_1 , X_2 , dan Y .

Terdapat **korelasi/hubungan yang signifikan positif** antara pendapatan, jumlah keluarga dengan pengeluaran untuk membeli barang A yaitu sebesar 0,914

Koefisien determinasi merupakan ukuran untuk mengetahui kesesuaian atau ketepatan antara nilai dugaan dengan data sampel. Koefisien determinasi didefinisikan sebagai berikut.

Koefisien determinasi adalah bagian dari keragaman total variabel tak bebas Y (variabel yang dipengaruhi atau dependen) yang dapat diterangkan atau diperhitungkan oleh keragaman variabel bebas X (variabel yang mempengaruhi, independen)

Jadi koefisien determinasi adalah kemampuan variabel X mempengaruhi variabel Y. Semakin besar koefisien determinasi menunjukkan semakin baik kemampuan X mempengaruhi Y.

$$\text{Koefisien Determinasi} = r^2 = \frac{[n \sum X_i Y_i - (\sum X_i)(\sum Y_i)]^2}{\sqrt{\{n(\sum X_i^2) - (\sum X_i)^2\} \{n(\sum Y_i^2) - (\sum Y_i)^2\}}}$$

Koefisien determinasi digunakan dalam analisis regresi untuk mengevaluasi seberapa baik model memprediksi atau menjelaskan variabilitas dalam data.

Latihan

1. Dua orang ibu rumah tangga diminta untuk mengemukakan tingkat preferensinya terhadap sabun mandi berbagai merk. Hasilnya adalah sebagai berikut :

Merk Sabun Mandi	Ny. Witono	Ny. Hartoko
A	3	5
B	5	6
C	8	4
D	12	9
E	10	8
F	7	12
G	9	11
H	1	3
I	4	1
J	6	2
K	2	10
L	11	7

Hitunglah nilai koefisien rank pada kasus di atas

2. Data di bawah ini menunjukkan jumlah pemakaian pupuk (X) dan hasil panen padi yang diperoleh (Y):

Pupuk (kg)	Hasil Panen (kw)
20	8
40	9
50	11
70	11
100	12
110	14
120	15
150	16

Hitung koefisien korelasi dengan metode product moment dan jelaskan artinya.

3. Tabel di bawah ini menunjukkan berat badan, tinggi badan, dan umur dari sampel random 12 anak laki-laki. Berat badan diukur dalam pound, tinggi badan diukur dalam inci, dan umur diukur dalam tahun.

Berat Badan (X1)	Tinggi Badan (X2)	Umur (Y)
64	57	8
71	59	10
53	49	6
67	62	11
55	51	8
58	50	7
77	55	10
57	48	9
56	52	10
51	42	6
76	61	12
68	57	9

Hitung koefisien korelasi antara variabel X_1 dan X_2 secara bersama-sama dengan variabel Y.