



# MODUL STATISTIK INFERENS (MIK 411)

## Materi 6 Analisis Korelasi

Disusun Oleh  
Mieke Nurmalasari, M.Si, M.Sc

UNIVERSITAS ESA UNGGUL  
2018

# Analisis Korelasi

## A. Pendahuluan

Pada pertemuan sebelumnya, kita telah melakukan analisis uji perbandingan atau komparatif terhadap nilai tengah (nilai rata-ratanya). Pada bagian ini kita akan mempelajari analisis hubungan, jadi bukan komparatif. Analisis korelasi atau hubungan ini diperlukan untuk mendapatkan gambaran awal mengenai keeratan hubungan antara dua variabel atau lebih.

## B. Kompetensi Dasar

Mahasiswa mengetahui konsep tentang hubungan suatu variabel satu dengan variabel lainnya. Mahasiswa mampu mengerti pentingnya mempelajari hubungan antar variabel yang menjadi pengamatan mereka.

## C. Kemampuan Akhir yang Diharapkan

1. Memahami pengertian analisis korelasi
2. Menghitung koefisien korelasi
3. Mampu melakukan pengujian hipotesis terhadap koefisien korelasi.

## D. Kegiatan Belajar

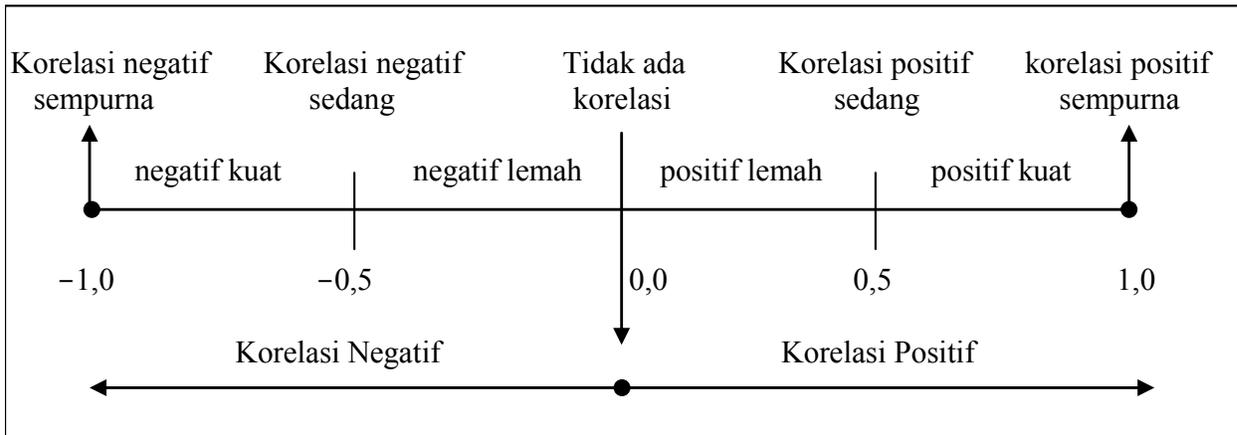
### Analisis Korelasi

#### 1. Uraian dan contoh

Korelasi adalah suatu ukuran yang digunakan untuk menyatakan derajat hubungan linier antar dua variabel. Hubungan antar kedua variabel ini tidak harus memperhatikan arah hubungan dari kedua variabel tersebut, atau tidak dibutuhkan adanya hubungan sebab akibat.

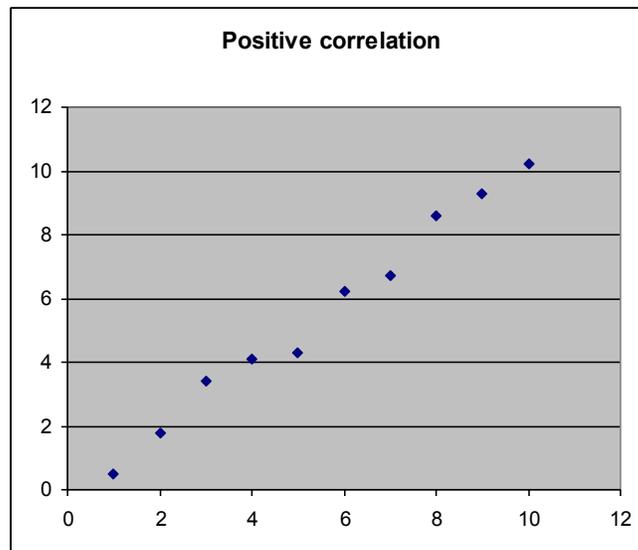
Korelasi yang akan kita pelajari adalah korelasi (*Moment Pearson*). Korelasi Pearson ini adalah korelasi yang paling banyak digunakan untuk melihat hubungan linier antar dua variabel. Dua variabel yang diukur syaratnya berskala interval dan rasio. Korelasi Pearson dapat diinterpretasikan ada hubungan linier atau tidak antara variabel X dengan variabel Y.

Nilai  $r$  bergerak antara  $-1$  dan  $+1$  dengan tanda negatif menyatakan adanya korelasi tak langsung atau korelasi negatif dan tanda positif menyatakan adanya korelasi langsung atau korelasi positif.  $r=0$  menyatakan tidak ada hubungan linier antara variabel X dan Y. Kekuatan keeratan hubungan berdasarkan range nilai korelasi di atas secara garis besar dapat ditampilkan dalam gambar 1. berikut:



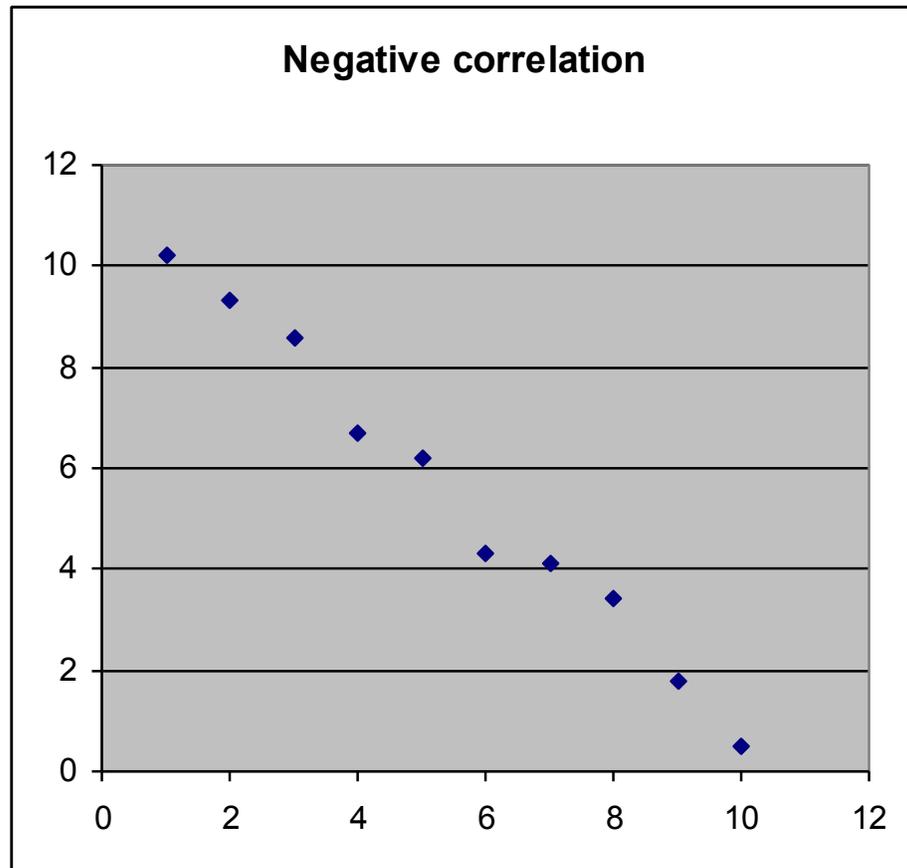
Gambar 1: Ilustrasi keeratan hubungan berdasarkan range nilai r

Keeratan hubungan antara dua variabel dapat kita eksplorasi terlebih dahulu dengan menggunakan scatter plot atau diagram pencar. Dari diagram pencar ini kita bisa mendapatkan gambaran awal tentang hubungan kedua variabel tersebut. Contoh pada Gambar 2, merupakan ilustrasi adanya hubungan yang erat dan positif. Perubahan satu variabel diikuti perubahan pada variabel lain dan arahnya positif (arah positif disini maksudnya adalah kenaikan pada variabel satu diikuti oleh kenaikan pada variabel lainnya).



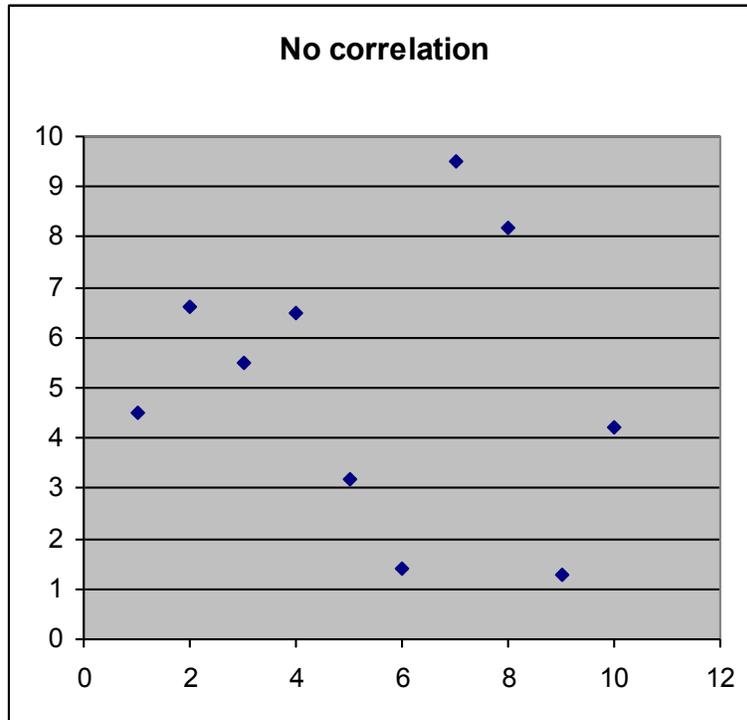
Gambar 2. Korelasi Positif

Pada Gambar 3. merupakan ilustrasi untuk keeratan hubungan dengan korelasi negatif. Dapat dilihat bahwa perubahan kenaikan pada variabel 1 membuat penurunan pada variabel lainnya. Sehingga membentuk pola trend menurun.



Gambar 3. Korelasi Negatif

Pola acak atau random yang ditampilkan pada Gambar 4, merupakan salah satu indikasi bahwa kedua variabel lemah hubungannya atau hampir tidak berhubungan antara variabel satu dengan variabel kedua.



Gambar 4. Tidak Ada Korelasi

Analisis Korelasi digunakan untuk mengetahui derajat/keeratan hubungan antara kedua variabel numerik tersebut. Analisis Korelasi Pearson adalah merupakan salah satu analisis uji statistik yang termasuk dalam statistik parametrik.

Syaratnya adalah:

- Data terdistribusi secara normal
- Mempunyai variansi yang sama

Jika syarat diatas tidak terpenuhi disarankan menggunakan analisis lain untuk menguji hubungan yaitu uji korelasi non parametrik. Skala data biasanya dalam bentuk interval atau rasio.

$$r = \frac{(n \sum XY) - (\sum X \sum Y)}{\sqrt{[(n \sum X^2) - (\sum X)^2] [n \sum Y^2 - (\sum Y)^2]}}$$

Nilai korelasi berkisar -1 sampai dengan 1 atau bisa ditulis  $-1 \leq r \leq 1$ .

$r = 0 \rightarrow$  tidak ada hubungan linier

$r = -1 \rightarrow$  hubungan linier negatif sempurna

$r = +1 \rightarrow$  hubungan linier positif sempurna

- Korelasi tidak selalu berarti hubungan sebab akibat (*causality*).
- Korelasi yang lemah tidak selalu berarti tidak adanya hubungan.
- Korelasi yang kuat tidak selalu berarti adanya garis lurus

Koefisien Determinasi ( $R^2$ )

- Koefisien Determinasi sering dituliskan dengan R-kuadrat ( $R^2$ ), dapat diartikan sebagai besarnya proporsi variasi variabel Y yang dapat dijelaskan oleh variabel X.
- Koefisien Determinasi sering ditampilkan dalam bentuk prosentase. Misalkan  $R^2 = 65\%$ , artinya variasi Y dapat dijelaskan 65% oleh variabel X, sedangkan sisanya 35% ( $100\% - 65\%$ ) diterangkan oleh faktor lain.

Subjek	$X_i$	$Y_i$	$X_i^2$	$Y_i^2$	$X_i Y_i$
1	$X_1$	$Y_1$	$X_1^2$	$Y_1^2$	$X_1 Y_1$
2	$X_2$	$Y_2$	$X_2^2$	$Y_2^2$	$X_2 Y_2$
...	...	...	...	...	...
n	$X_n$	$Y_n$	$X_n^2$	$Y_n^2$	$X_n Y_n$
	$\sum X = \dots$	$\sum Y \dots$	$\sum X^2 = \dots$	$\sum Y^2 = \dots$	$\sum XY = \dots$

### Contoh

Ingin diselidiki apakah skor kepuasan pasien berhubungan dengan umur pasien tersebut. Maka dikumpulkan data 10 pasien dan dicatat skor kepuasan dan umurnya. Data ditampilkan pada slide berikutnya:

Kepuasan Pasien (Y)	Umur (X)	Y <sup>2</sup>	X <sup>2</sup>	XY
48	50	2304	2500	2400
57	36	3249	1296	2052
66	40	4356	1600	2640
70	41	4900	1681	2870
89	28	7921	784	2492
36	49	1296	2401	1764
46	42	2116	1764	1932
54	45	2916	2025	2430
26	52	676	2704	1352
77	29	5929	841	2233
$\sum Y = 569$	$\sum X = 412$	$\sum Y^2 = 35663$	$\sum X^2 = 17596$	$\sum XY = 22165$

### Pengujian Koefisien Korelasi

- $H_0 : \rho = 0$  (Tidak terdapat hubungan yang signifikan antara variabel kepuasan pasien dengan umur pasien).
- $H_a : \rho \neq 0$  (Terdapat hubungan yang signifikan antara variabel kepuasan pasien dengan umur pasien).
- Statistik uji yang digunakan adalah statistik Uji-t.
- $t\text{-hitung} = (r \sqrt{(n-2)}) / (\sqrt{(1-r^2)})$ ,
- dibandingkan dengan  $t_{\text{tabel}} = t_{(n-2)}$
- Jika  $|t_{\text{hitung}}| > t_{(\alpha/2, n-2)}$  maka Tolak  $H_0$

### Analisis Korelasi

Ingin diselidiki apakah terdapat hubungan antara kepuasan pasien dengan tingkat keparahan dari si pasien. Data yang diperoleh ditampilkan dalam table di bawah ini:

Patient Satisfaction	Severity Illness
48	51
57	46
66	48
70	44
89	43
36	54
46	50
54	48
26	62
77	50

Hitung nilai koefisien korelasinya!

Y	X	Y <sup>2</sup>	X <sup>2</sup>	XY
48	51	2304	2601	2448
57	46	3249	2116	2622
66	48	4356	2304	3168
70	44	4900	1936	3080
89	43	7921	1849	3827
36	54	1296	2916	1944
46	50	2116	2500	2300
54	48	2916	2304	2592
26	62	676	3844	1612
77	50	5929	2500	3850
Σ Y= 569    ΣX= 496    ΣY <sup>2</sup> = 35663    ΣX <sup>2</sup> = 24870    ΣXY = 27443				

$$r = \frac{n(\sum XY) - (\sum X)(\sum Y)}{\sqrt{[(n \sum X^2) - (\sum X)^2][(n \sum Y^2) - (\sum Y)^2]}}$$

$$r = \frac{10(27443) - (496)(569)}{\sqrt{[(10(24870)) - (496)^2][(10(35663) - (569)^2]}}$$

$$r = \frac{274430 - 282224}{\sqrt{(248700 - 246016)(356630 - 323761)}}$$

$$r = \frac{-7794}{\sqrt{(2684)(32869)}}$$

$$r = 0,83 \sim 1$$

Dapat disimpulkan bahwa terdapat hubungan linear positif sempurna antara kepuasan pasien dengan tingkat keparahan dari si pasien

### Soal

Seorang pegawai bagian penjaminan mutu sebuah rumah sakit ingin melakukan studi tentang kepuasan pasien dilihat dari usia pasien. Pegawai tersebut memilih secara acak 10 orang pasien untuk dijadikan data sampel.

Patient Satisfaction	Age
48	50
57	36
66	40
70	41
89	28
36	49
46	42
54	45
26	52
77	29

Berdasarkan data di atas:

1. Definiskan variabel bebas dan variabel tidak bebasnya!
2. Hitung koefisien regresinya dan buat persamaannya!
3. Apa arti dari koefisien regresinya!
4. Jika seorang pasien baru masuk rumah sakit dengan umur 35, hitunglah prediksi tingkat kepuasan pasien baru tersebut!
5. Lakukan uji koefisien regresi untuk koefisien  $b_1$ !
6. Lakukan uji koefisien regresi secara simultan!

### Jawab:

1. Variabel Bebas (X) = Age (Umur)  
Variabel tidak bebas (Y) = Patients Satisfaction (Kepuasan Pasien)
2. Estimasi koefisien regresi dan persamaan Regresinya  
Untuk membuat persamaan regresi kita harus menghitung nilai  $b_0$  dan  $b_1$

Y	X	Y <sup>2</sup>	X <sup>2</sup>	XY
48	50	2304	2500	2400
57	36	3249	1296	2052
66	40	4356	1600	2640
70	41	4900	1681	2870
89	28	7921	784	2492
36	49	1296	2401	1764
46	42	2116	1764	1932
54	45	2916	2025	2430
26	52	676	2704	1352
77	29	5929	841	2233
$\Sigma Y = 569$	$\Sigma X = 412$	$\Sigma Y^2 = 35663$	$\Sigma X^2 = 17596$	$\Sigma XY = 22165$

Estimasi koefisien regresi  $b_1$ :

$$\begin{aligned} b_1 &= \frac{n(\sum XY) - (\sum X)(\sum Y)}{n \sum X^2 - (\sum X)^2} \\ &= \frac{10(22165) - (412)(569)}{10(17596) - (412)^2} \\ &= \frac{221650 - 234428}{175960 - 169744} \\ &= \frac{-12778}{6216} \\ &= -2,056 \end{aligned}$$

$$\begin{aligned} b_0 &= \frac{\sum Y}{n} - b_1 \frac{\sum X}{n} \\ &= \frac{569}{10} - (-2,056) \frac{412}{10} \\ &= 56,9 + 84,707 \\ &= 141,607 \end{aligned}$$

Jadi persamaan regresinya adalah:

$$\hat{Y} = b_0 + b_1 X_1 = 141,607 - 2,056 X_1$$

3. Arti Koefisien Regresi:  
 $b_1$  = setiap kenaikan umur pasien 1 tahun, maka skor kepuasan pasien menurun sebesar 2,056.
4. Jika umur=35, maka prediksi tingkat kepuasan pasiennya adalah:

$$\begin{aligned} \hat{Y} &= 141,607 - 2,056 X_1 \\ &= 141,607 - 71,96 \\ &= 69,647 \end{aligned}$$

Jadi pasien dengan umur 35 tahun, diprediksi skor kepuasannya sebesar 69,647

## 5. Uji koefisien regresi secara parsial (uji koefisien $b_1$ ):

### Langkah 1.

Rumuskan Hipotesis

$$H_0 : \beta_1 = 0$$

$$H_1 : \beta_k \neq 0$$

### Langkah 2:

$$t_{hitung} = \frac{b_1}{sb_1}$$

$$S_b = \sqrt{\frac{KTE}{(n-1)S_{xx}}}$$

$$S_{xx} = \frac{\sum X^2 - \frac{(\sum X)^2}{n}}{n-1} = \frac{(17596) - \frac{(412)^2}{10}}{10-1}$$

$$S_{yy} = \frac{\sum Y^2 - \frac{(\sum Y)^2}{n}}{n-1} = \frac{(35663) - \frac{(569)^2}{10}}{10-1}$$

$$KTE = \frac{(n-1)(S_{yy} - b^2 S_{xx})}{(n-2)}$$

### Uji koefisien regresi secara simultan:

Pada uji ini akan dilakukan uji koefisien regresi secara bersamaan, uji ini juga digunakan untuk menguji kelayakan persamaan regresinya.

### Langkah 1.

Menyusun Hipotesis

$H_0 : \beta_0 = \beta_1 = 0$  (tidak ada pengaruh umur terhadap kepuasan pasien)

$H_1 : \beta_k \neq 0$  (Umur berpengaruh secara signifikan terhadap skor kepuasan pasien).

**Langkah 2:**

Menghitung statistik uji – F

$$F_{hitung} = \frac{KTR}{KTE} = \frac{JKR/(1)}{JKE/((n-1))}$$

JKT = Jumlah Kuadrat Total = (n-1) S<sub>yy</sub> = (10-1) (....)

JKT bisa dihitung dengan rumus sebagai berikut:

$$JKT = \sum Y^2 - \frac{(\sum Y)^2}{n}$$

$$JKT = (35663) - \frac{(569)^2}{10}$$

= .....

JKR = (n-1)b<sup>2</sup>S<sub>xx</sub> = (10-1) (-2, 056)<sup>2</sup> (....)

JKE = JKT – JKR

**E. Daftar Pustaka**

1. Kurtner, MH, Nachtsheim CJ, Neter J, Li W. Applied Linear Statistical Models. Fifth Edition. McGraw-Hill. 2005.
2. Efendi, A., Pramoedyo, H., Rosner, B. Fundamentals of Biostatistics. Brooks Cole. 2015.
3. Biostatistika Dengan R dan MS Excel, UB Press, 2017
4. <http://www.stat.ufl.edu/~rrandles/sta4210/Rclassnotes/data/textdatasets/Chapter%20%206%20Data%20Sets.html>
5. <https://mysite.science.uottawa.ca/rkulik/mat3378/mat3378-textbook.pdf>  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2992018/>