



Universitas Gadjah Mada  
Fakultas Teknik  
Departemen Teknik Sipil dan Lingkungan

# KORELASI

Statistika dan Probabilitas

# Koefisien Korelasi

2

$$r^2 = \frac{S_t - S_r}{S_t} = 1 - \frac{S_r}{S_t} \quad \rightarrow \quad r = \sqrt{\frac{S_t - S_r}{S_t}} = \sqrt{1 - \frac{S_r}{S_t}} \quad \text{koefisien korelasi}$$

$$S_t = \sum (y_i - \bar{y})^2$$

$$S_r = \sum (y_i - a_0 - a_1 x_i)^2$$

$$r = \sqrt{1 - \frac{\sum (y_i - a_0 - a_1 x_i)^2}{\sum (y_i - \bar{y})^2}}$$

# Koefisien Korelasi

3

$$r = \sqrt{1 - \frac{\sum (y_i - a_0 - a_1 x_i)^2}{\sum (y_i - \bar{y})^2}} \quad \rightarrow \quad r = \frac{n \sum x_i y_i - (\sum x_i)(\sum y_i)}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}}$$

# Koefisien Korelasi

4

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

$$s_{X,Y} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{n - 1}$$



kovarian X dan Y

$$s_X = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n - 1}}$$



simpangan baku X

$$s_Y = \sqrt{\frac{\sum (Y_i - \bar{Y})^2}{n - 1}}$$



simpangan baku Y

# Koefisien Korelasi

5

$$r = r_{X,Y} = \frac{S_{X,Y}}{S_X S_Y} \rightarrow r_{X,Y} \leftarrow \frac{= \text{COVARIANCE.S}(X,Y)}{= \text{STDEV.S}(X) \times \text{STDEV.S}(y)}$$
$$r_{X,Y} \leftarrow = \text{CORREL}(X,Y)$$

MS Excel

- koefisien korelasi antara variabel random X dan Y

# Koefisien Korelasi

6

- ❑ Pengertian koefisien korelasi
  - ❑ Koefisien korelasi menunjukkan tingkat **keeratan hubungan linear** antara suatu variabel random  $Y$  dan suatu variabel kedua yang merupakan fungsi linear dari satu atau lebih variabel(-variabel)  $X$ 
    - Setiap variabel  $X$  dapat berupa variabel random atau bukan variabel random

# Koefisien Korelasi

7

- ❑ Nilai koefisien korelasi adalah  $-1 < r_{X,Y} < 1$ 
  - ❑  $r_{X,Y} = \pm 1$  menunjukkan hubungan linear sempurna antara  $X$  dan  $Y$
  - ❑  $r_{X,Y} = 0$  menunjukkan independensi (ketidak-bergantungan) linear, namun dapat saja keduanya memiliki hubungan (kebergantungan) yang lain, yang tidak linear
  - ❑ Jika  $X$  dan  $Y$  tidak saling bergantung (*independent*), maka  $r_{X,Y} = 0$
- ❑ Koefisien korelasi sampel dan populasi
  - ❑  $r_{X,Y}$  koefisien korelasi sampel
  - ❑  $\rho_{X,Y}$  koefisien korelasi populasi

# Inferensi terhadap Koefisien Korelasi Populasi

8

- ❑ Dua variabel random
  - ❑ tak berkorelasi,  $\rho_{X,Y} = 0$
  - ❑ berkorelasi,  $\rho_{X,Y} \neq 0$
- ❑ Situasi
  - ❑ Sampel yang diperoleh dari variabel random yang tidak berkorelasi
    - jarang menunjukkan nilai  $r_{X,Y} = 0$
    - koefisien korelasi sering  $r_{X,Y} \neq 0$ , karena kebetulan
  - ❑ Oleh karena itu perlu pengujian
    - untuk mengetahui penyimpangan koefisien korelasi dari nol tersebut disebabkan oleh kebetulan, atau
    - penyimpangan tersebut terlalu besar untuk dikatakan sebagai akibat kebetulan



# Inferensi terhadap $\rho$

9

## □ Uji hipotesis

□  $H_0: \rho_{X,Y} = 0$

□  $H_1: \rho_{X,Y} \neq 0$

statistik uji  $T = r \left[ \frac{n-2}{1-r^2} \right]^{1/2} \rightarrow |T| > t_{1-\alpha/2, n-2} \quad H_0 \text{ ditolak}$

# Inferensi terhadap $\rho$

10

## Uji hipotesis

□  $H_0: \rho_{X,Y} = \rho^*$  ( $\rho^*$  konstanta)

ukuran sampel  $n > 25$

□  $H_1: \rho_{X,Y} \neq \rho^*$

statistik uji  $Z = (W - \omega)(n - 3)^{1/3} \rightarrow |Z| > z_{1-\alpha/2}$   $H_0$  ditolak

$$W = \frac{1}{2} \ln \left[ \frac{1+r}{1-r} \right] = \tanh^{-1} r$$

$$\omega = \frac{1}{2} \ln \left[ \frac{1+\rho^*}{1-\rho^*} \right] = \tanh^{-1} \rho^*$$

□ Rentang keyakinan  $\rho$ :  $\ell = \tanh \left[ W - \frac{z_{1-\alpha/2}}{(n-3)^{1/2}} \right]$   $u = \tanh \left[ W + \frac{z_{1-\alpha/2}}{(n-3)^{1/2}} \right]$

11

# Korelasi

## Korelasi Serial

# Korelasi Serial

12

- ❑ Korelasi serial (*serial correlation*)
  - ❑ dikenal pula sebagai autokorelasi (*autocorrelation*)
  - ❑ yaitu korelasi antara data hasil pengukuran pada suatu waktu dengan data hasil pengukuran pada waktu sebelumnya
  - ❑ elemen dalam sampel yang memiliki korelasi serial **bukan** elemen random (ingat definisi variabel random)

# Korelasi Serial

13

- ❑ Pada korelasi serial, dengan demikian
  - ❑ sampel berukuran  $n$  yang memiliki korelasi serial akan memberikan informasi yang lebih sedikit dibandingkan dengan informasi yang dimiliki oleh sampel random berukuran  $n$
  - ❑ sebagian informasi pada sampel yang memiliki korelasi serial dapat diperoleh dari atau telah diketahui dalam data hasil pengukuran pada waktu sebelumnya

# Korelasi Serial

14

- ❑ Korelasi serial (*serial correlation*)
  - ❑ Dapat pula dijumpai antara suatu pengukuran pada waktu tertentu dengan pengukuran pada waktu  $k$  periode waktu sebelumnya (terdahulu),  $k = 1, 2, \dots$
  - ❑ Asumsi
    - Selang waktu antar pengukuran adalah sama (seragam)
    - Sifat-sifat statistis proses atau peristiwa yang diukur tidak berubah terhadap waktu (bersifat permanen)
    - $\rho(k)$  koefisien korelasi serial populasi
    - $r(k)$  koefisien korelasi serial sampel

# Korelasi Serial

15

$$r(k) = \frac{\sum_{i=1}^{n-k} x_i x_{i+k} - \sum_{i=1}^{n-k} x_i \sum_{i=1}^{n-k} x_{i+k} / (n - k)}{\left[ \sum_{i=1}^{n-k} x_i^2 - \left( \sum_{i=1}^{n-k} x_i \right)^2 / (n - k) \right]^{1/2} \left[ \sum_{i=1}^{n-k} x_{i+k}^2 - \left( \sum_{i=1}^{n-k} x_{i+k} \right)^2 / (n - k) \right]^{1/2}}$$

$$r(k) = r_{X_i, X_{i+k}} = \frac{S_{X_i, X_{i+k}}}{S_{X_i} S_{X_{i+k}}} \quad r_{X,Y} \leftarrow \frac{= \text{COVARIANCE.S}(X,Y)}{= \text{STDEV.S}(X) \times \text{STDEV.S}(Y)}$$

$$r_{X,Y} \leftarrow = \text{CORREL}(X,Y)$$

# Korelasi Serial

16

- ❑  $r(0) = 1$  korelasi suatu elemen data dengan dirinya sendiri adalah sama dengan satu
- ❑ semakin besar  $k$ , jumlah pasangan data untuk menghitung  $r(k)$  semakin sedikit;  $r(k)$  adalah nilai estimasi  $\rho(k)$
- ❑ oleh karena itu,  $k \ll n$
- ❑ jika  $\rho(k) = 0$  untuk semua  $k$ , maka proses atau peristiwa atau populasi tersebut bersifat random murni



# Terima kasih