

# PERTEMUAN 2-3

**SIFAT DASAR ANALISIS REGRESI**

# Penafsiran Regresi

- Analisis regresi berkaitan dengan studi ketergantungan satu variabel- variabel tak bebas terhadap satu atau lebih variabel lain-variabel yang menjelaskan, dengan maksud menaksir dan atau meramalkan mean atau rata-rata populasi variabel tak bebas.

# Regresi dan Penyebab (kausalitas)

- Analisis regresi berkaitan dengan ketergantungan satu variabel dengan variabel lain, namun hal tersebut tidak berarti terjadi hubungan sebab akibat (kausalitas)
- Dalam hubungan kausalitas semua variabel menjadi variabel dependen

# Regresi dan Korelasi

- Dalam analisis korelasi semua variabel diperlakukan secara simetri. Semua variabel diasumsikan random.
- Dalam analisis regresi variabel tak bebas diasumsikan bersifat stokastik, sedangkan variabel bebas bersifat tetap atau nonstokastik

# Regresi dan Korelasi

- Dalam analisis regresi ada asimetri dalam memperlakukan variabel tak bebas dan variabel bebas. Variabel tak bebas diasumsikan bersifat statistik, random, atau stokastik, yaitu mempunyai distribusi probabilitas. Variabel bebas diasumsikan mempunyai nilai yang tetap

# Istilah dan Notasi

- Dependent Variable
- Explained variable
- Predictand
- Regressand
- Response
- Explanatory variable
- Independent variable
- Predictor
- Regressor
- Stimulus or control variable

# Istilah dan Notasi

- Biasanya  $Y$  menyatakan variabel tak bebas dan  $X$  variabel bebas
- $X_k$  artinya variabel bebas yang ke  $k$ . Indeks bawah (subscript)  $i$  atau  $t$  menggambarkan pengamatan atau nilai ke  $i$  atau ke  $t$
- Secara umum indeks bawah  $i$  digunakan untuk data *cross-section*, yaitu data yang dikumpulkan pada satu titik waktu
- Indeks bawah  $t$  digunakan untuk data *time series*

# LINEARITAS

- Linearitas dalam variabel dapat diartikan bahwa harapan bersyarat (conditional expectation) dari Y adalah fungsi linear dari  $X_i$ .

$$E(Y|X_i) = \beta_0 + \beta_1 X_i$$

Kurva regresi dalam kasus ini berbentuk garis lurus.



# LINEARITAS

- Linearitas dalam parameter

Penafsiran linearitas dalam parameter adalah bahwa harapan bersyarat dari Y merupakan fungsi linear dari parameter. Dalam fungsi ini variabel X mungkin bersifat linear atau nonlinear.

$$E(Y|X_i) = \beta_0 + \beta_1 X_i^2$$

# Population Regression Function

$$E(Y|X_i) = f(X_i) \dots\dots\dots (1)$$

- Dimana  $f(X_i)$  menggambarkan suatu fungsi dari variable yang menjelaskan  $X_i$ .
- $E(Y|X_i)$  merupakan fungsi linear dari  $X_i$ .
- Persamaan (1) dikenal sebagai fungsi regresi populasi (PRF) atau seringkali disingkat regresi populasi (PR).

# PRF

- Fungsi tersebut menyatakan bahwa rata-rata (populasi) dari distribusi  $Y$  untuk  $X_i$  tertentu berhubungan secara fungsional dengan  $X_i$ . Dengan kata lain fungsi tersebut menyatakan bagaimana nilai rata-rata (populasi) bervariasi bersama dengan  $X$ .

# PRF

- Asumsikan bahwa PRF  $E(Y|X_i)$  merupakan fungsi linear dari  $X_i$  :

$$E(Y|X_i) = \beta_0 + \beta_1 X_i \dots\dots\dots (2)$$

- Dimana  $\beta_0$  dan  $\beta_1$  parameter yang tidak diketahui besarnya tetapi tetap (fixed) dikenal dengan istilah koefisien regresi.

# PRF

- $\beta_0$  dan  $\beta_1$  secara berturut-turut dikenal sebagai intersep dan koefisien kemiringan (slope coefficient).
- Persamaan (2) dikenal sebagai fungsi regresi linear atau regresi populasi linear.

# SPEKIFIKASI STOKASTIK PRF

$$u_i = Y_i - E(Y|X_i) \dots\dots\dots (3)$$

$$Y_i = E(Y|X_i) + u_i \dots\dots\dots (4)$$

- Dimana penyimpangan  $u_i$  merupakan variable random yang tidak bisa diamati yang bisa mengambil nilai positif atau negative. Secara teknis  $u_i$  dikenal sebagai gangguan stokastik (*stochastic disturbance*), atau factor kesalahan stokastik (*stochastic error term*).

# SPEKIFIKASI STOKASTIK PRF

- Kalau  $E(Y|X_i)$  diasumsikan linear dalam  $X_i$ , persamaan (4) dapat ditulis sebagai berikut:

$$\begin{aligned} Y_i &= E(Y|X_i) + u_i \\ &= \beta_0 + \beta_1 X_i + u_i \dots \dots \dots (5) \end{aligned}$$

$$\begin{aligned} E(Y|X_i) &= E[E(Y|X_i)] + E(u|X_i) \\ &= E(Y|X_i) + E(u|X_i) \dots \dots \dots (6) \end{aligned}$$

-

# SPEKIFIKASI STOKASTIK PRF

- Persamaan (6) mengandung arti bahwa :

$$E(u|X_i) = 0$$

- Jadi asumsi bahwa garis regresi melewati rata-rata bersyarat dari Y berarti bahwa nilai rata-rata bersyarat dari  $u_i$  (tergantung pada X tertentu) adalah nol.



# SAMPLE REGRESSION FUNCTION

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_i$$

Dimana :

- $\wedge$  : dibaca sebagai hat atau cap (topi)

$\hat{Y}$  : penaksir (estimator)

$\hat{\beta}_0$  : penaksir  $\beta_0$

$\hat{\beta}_1$  : penaksir  $\beta_1$

# SRF

- Suatu penaksir , juga dikenal dengan statistic (sampel) hanyalah suatu formula atau metode yang menyatakan bagaimana untuk menaksir parameter populasi dari informasi yang diberikan oleh sampel yang dimiliki.
- Suatu nilai angka khusus yang diperoleh oleh penaksir dalam suatu analisis dikenal sebagai taksiran (estimate).

# SPESIFIKASI STOKASTIF SRF

$$Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + e_i$$

- Dimana  $e_i$  menyatakan factor residual/ sisa (sampel).

Kesimpulan :

- Tujuan utama kita dalam analisis regresi adalah untuk menaksir PRF :

$$Y = \beta_0 + \beta_1 X_i + u_i$$

Atas dasar SRF

$$Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + e_i$$

- Untuk  $X = X_i$ , kita mempunyai pengamatan (sampel)  $Y = Y_i$ . Dalam pengertian SRF,  $Y_i$  yang teramati dapat dinyatakan sebagai :

$$Y = \bar{Y}_i + e_i$$

- Dan dalam pengertian PRF dapat dinyatakan sebagai :

$$Y_i = E(Y|X_i) + u_i$$