



# **MODEL REGRESI DENGAN VARIABEL BEBAS DUMMY**

# PENDAHULUAN

- Regresi yang telah dipelajari → data kuantitatif
- Analisis → membutuhkan analisis kualitatif.

Contoh:

- Pengaruh jenis Kelamin terhadap gaji.
  - Pengaruh kualitas produk terhadap omset.
  - Pengaruh harga terhadap kepuasan pelayanan.
  - Pengaruh pendidikan terhadap umur perkawinan pertama.
- 
- Contoh (1) & (2) → variabel bebas kualitatif dan variabel terikat kuantitatif.
  - Contoh (3) → variabel bebas kuantitatif dan variabel terikat kualitatif.
  - Contoh (4) → variabel bebas kualitatif dan variabel terikat kualitatif.
- 
- (1) dan (2) → Regresi dengan Dummy Variabel
  - (3) dan (4) → Model Logistik atau Multinomial

# PENDAHULUAN

- Data Kualitatif harus berbentuk data kategorik → Belum bisa dibuat regresi secara langsung → Variabel Dummy.
- Variabel dummy disebut juga variabel indikator, biner, kategorik, kualitatif, boneka, atau variabel dikotomi.
- Variabel Dummy → pada prinsipnya merupakan perbandingan karakteristik. Misalnya:
  - Perbandingan kondisi (besaran/jumlah) konsumen yang merasa puas terhadap suatu produk dengan konsumen yang tidak puas.
  - Perbandingan besarnya gaji antara laki-laki dan perempuan.

# Tekhnik pembentukan Variabel Dummy dan Estimasi

- Dummy bernilai 1 atau 0. Kenapa?

Perhatikan data kategorik berikut:

1. Konsumen puas
2. Konsumen tidak puas

Bisakah kita membuat regresi dengan 'kode kategorik' diatas, yaitu 1 dan 2?

Bila digunakan kode kategorik tersebut, berarti kita sudah memberi nilai pada 'konsumen yang tidak puas' dua kali 'konsumen yang puas'.

Bila dibuat dummy, misalnya:

1. Konsumen puas = 1
2. Konsumen tidak puas = 0.

# Tekhnik pembentukan Variabel Dummy dan Estimasi

- Regresi yang dibuat menunjukkan kondisi dimana konsumen merasa puas (Dummy berharga 1  $\rightarrow$  Dummy ada dalam model), dan kondisi sebaliknya (Dummy berharga 0  $\rightarrow$  Dummy 'hilang' dari model). Jadi modelnya akan menunjukkan kondisi 'ada' atau 'tidak ada' Dummy.
- Untuk jelasnya perhatikan contoh berikut:  
Penelitian mengenai pengaruh daerah tempat, yaitu kota atau desa, terhadap harga berbagai macam produk.

$$\text{Model: } Y = \alpha + \beta D + u$$

Y = Harga produk

D = Daerah tempat tinggal

D = 1 ; Kota

D = 0 ; Desa

u = kesalahan random.

Catatan: Dummy yang bernilai 0 disebut dengan kategorik pembanding atau dasar atau reference.

# ILUSTRASI

- Dari model di atas, rata-rata harga produk :  
Kota :  $E(Y \mid D = 1) = \alpha + \beta$   
Desa :  $E(Y \mid D = 0) = \alpha$
- Jika  $\beta = 0 \rightarrow$  tidak terdapat perbedaan harga antara daerah perkotaan dengan pedesaan.
- Jika  $\beta \neq 0 \rightarrow$  terdapat perbedaan harga antara daerah perkotaan dengan pedesaan.
- Model diatas  $\rightarrow$  merupakan model Regresi  $\rightarrow$  OLS

# ILUSTRASI

- Misal hasil estimasi dengan OLS untuk model di atas didapat:

$$Y = 9,4 + 16 D$$
$$t \quad (53,22) \quad (6,245)$$
$$R^2 = 96,54\%$$

- $\alpha \neq 0$  dan  $\beta \neq 0$ ; yaitu :  $\alpha = 9,4$  dan  $\beta = 16$ .
- Artinya, harga rata-rata produk didaerah perkotaan adalah:  $9,4 + 16 = 25,4$  ribu rupiah, dan pedesaan sebesar 9,4 ribu rupiah. Dengan demikian dapat disimpulkan, harga produk daerah perkotaan lebih mahal dibanding pedesaan.

# Model: variabel bebas merupakan variabel kuantitatif dan variabel kualitatif.

- Contoh: Analisis mengenai gaji dosen di sebuah perguruan tinggi swasta di Jakarta, berdasarkan jenis kelamin dan lamanya mengajar.

Didefinisikan :

$Y$  = gaji seorang dosen

$X$  = lamanya mengajar (tahun)

$G = 1$  ; dosen laki-laki

$0$  ; dosen perempuan

Model :

$$Y = \alpha_1 + \alpha_2 G + \beta X + u$$

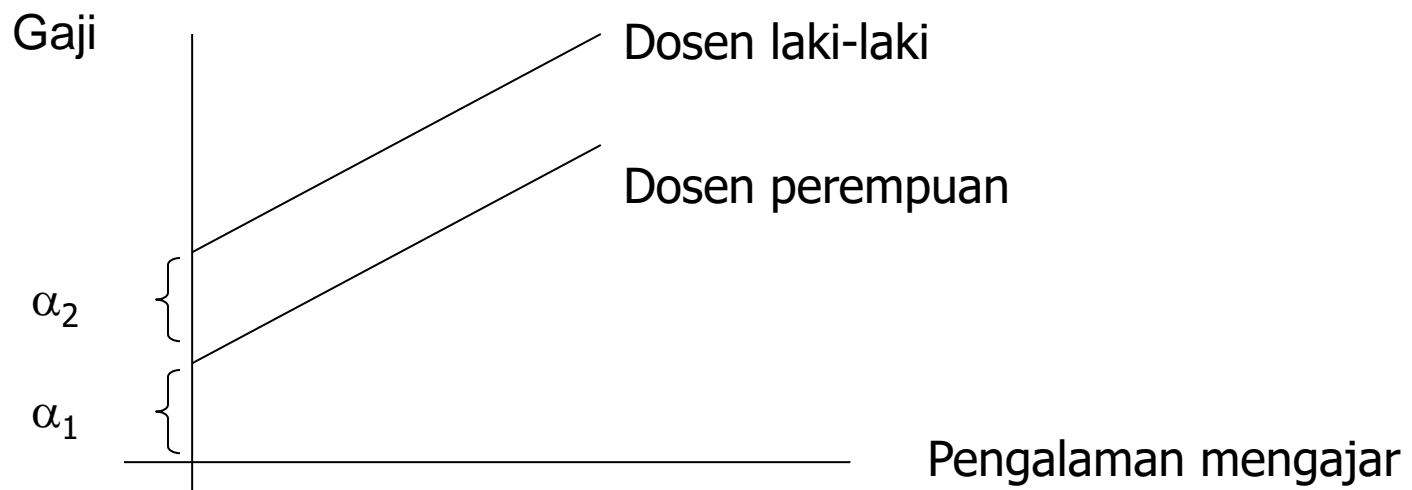
Dari model ini dapat dilihat bahwa :

- Rata-rata gaji dosen perempuan =  $\alpha_1 + \beta X$
- Rata-rata gaji dosen laki-laki =  $\alpha_1 + \alpha_2 + \beta X$

# Model: variabel bebas merupakan variabel kuantitatif dan variabel kualitatif.

- Jika  $\alpha_2 = 0 \rightarrow$  tidak ada diskriminasi gaji antara dosen laki-laki dan perempuan
- Jika  $\alpha_2 \neq 0 \rightarrow$  ada diskriminasi gaji antara dosen laki-laki dan perempuan

Misal: gaji dosen laki-laki  $>$  perempuan, maka secara geometris, model dapat digambarkan sebagai berikut :



# Bagaimana jika pendefinisian laki-laki dan perempuan dibalik?

- Misalkan :

$S = 1$ ; dosen perempuan  
 $= 0$ ; dosen laki-laki

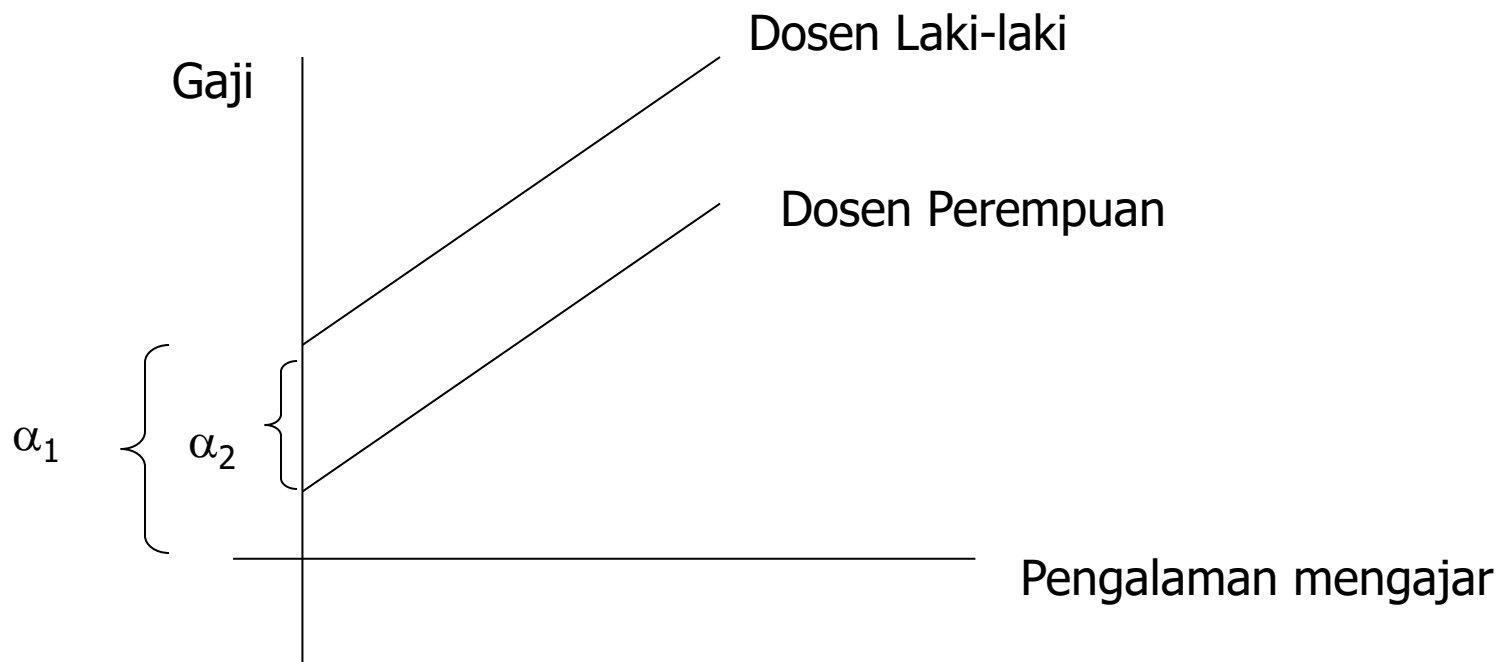
- Modelnya menjadi :

$$Y = \alpha_1 + \alpha_2 S + \beta X + u$$

- Jika  $\alpha_2 = 0 \rightarrow$  tidak ada diskriminasi gaji antara dosen laki-laki dan perempuan
- Jika  $\alpha_2 \neq 0 \rightarrow$  ada diskriminasi gaji antara dosen laki-laki dan perempuan

# Pembalikan Definisi

- Misal: gaji dosen laki-laki > perempuan  $\rightarrow \alpha_2$  akan bertanda negatif, maka secara geometris, model dapat digambarkan sebagai berikut :



# PENDEFINISIAN

- Perlu diperhatikan sekarang bahwa berdasarkan pendefinisian baru:
  - Rata-rata gaji dosen perempuan =  $\alpha_1 - \alpha_2 + \beta X$
  - Rata-rata gaji dosen laki-laki =  $\alpha_1 + \beta X$
- Jadi, apapun kategorik pembanding akan menghasilkan kesimpulan yang sama, sekalipun taksiran nilai koefisien regresi berbeda.
- Bagaimana kalau definisi:
  - $D_2 = 1$ ; dosen laki-laki  
0; dosen perempuan
  - $D_3 = 1$ ; dosen perempuan  
0; dosen laki-laki

# PENDEFINISIAN

- Sehingga modelnya menjadi :

$$Y = \alpha_1 + \alpha_2 D_2 + \alpha_3 D_3 + \beta X + u$$

- Apa yang akan terjadi bila model ini diestimasi dengan OLS ?
- Perhatikan: ada hubungan linear antara  $D_2$  dan  $D_3$  yakni  $D_2 = 1 - D_3$  atau  $D_3 = 1 - D_2 \rightarrow$  *perfect colinearity* antara  $D_2$  dan  $D_3$  sehingga OLS tidak dapat digunakan.
- Dalam membuat Dummy: Jika data mempunyai kategori sebanyak  $r$ , maka kita hanya memerlukan  $r-1$  variabel dummy. Dalam contoh di atas, kategorinya hanya dua, yaitu laki-laki dan perempuan. Oleh sebab itu, hanya satu variabel dummy yang dibutuhkan.

# Varibel dengan Kategori Lebih dari Dua

- Misalkan:

Pendidikan mempunyai 3 kategori:

1. Tidak tamat SMU
2. Tamat SMU
3. Tamat Perguruan tinggi.

- Dibutuhkan variabel dummy sebanyak  $(3-1) = 2$ .
- Dua variabel dummy tersebut yaitu  $D_2$  dan  $D_3$  didefinisikan sebagai berikut:

$D_2 =$      1 ; pendidikan terakhir SMU  
              0 ; lainnya

$D_3 =$      1 ; pendidikan terakhir perguruan tinggi  
              0 ; lainnya

- Manakah kategorik pembandingnya?

# ILUSTRASI

- Perhatikan model berikut :

$$Y = \alpha_1 + \alpha_2 D_2 + \alpha_3 D_3 + \beta X + u$$

Y = pengeluaran untuk *health care* per tahun

X = pendapatan per tahun

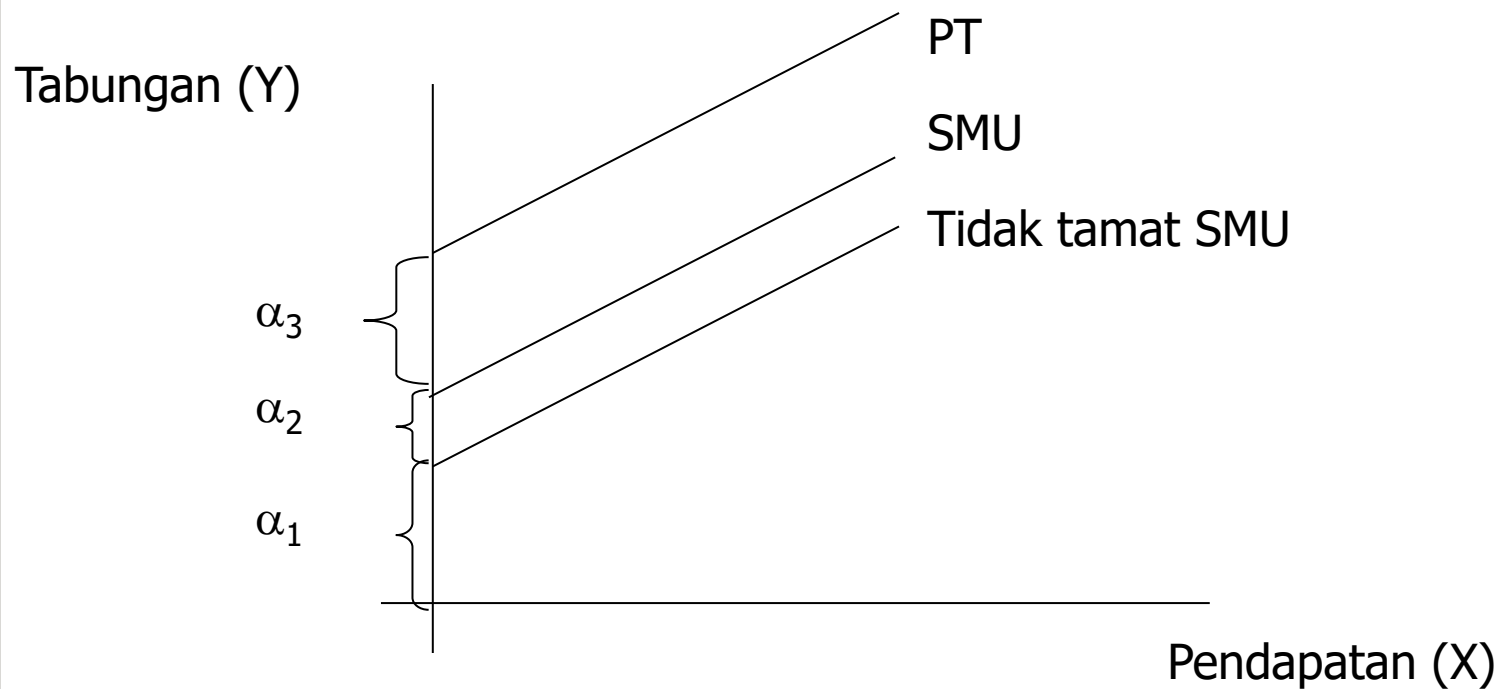
D<sub>2</sub> = 1 ; pendidikan tertinggi SMU  
0 ; lainnya

D<sub>3</sub> = 1 ; pendidikan tertinggi perguruan tinggi (SI)  
0 ; lainnya

- Berapa rata-rata pengeluaran seseorang berdasarkan pendidikannya?
  - Tidak tamat SMU :  $\alpha_1 + \beta X$
  - Tamat SMU:  $\alpha_1 + \alpha_2 + \beta X$
  - Berijazah SI :  $\alpha_1 + \alpha_3 + \beta X$

# ILUSTRASI

- Kalau dilihat secara geometris, pengeluaran untuk *health care* tersebut adalah sebagai berikut :



# Regresi Dengan Beberapa Variabel Kualitatif

- Contoh:

$$Y = \alpha_1 + \alpha_2 D_2 + \alpha_3 D_3 + \beta X + u$$

$Y$  = gaji

$D_2 = 1$  ; dosen laki-laki  
 $0$  ; dosen perempuan

$X$  = pengalaman (tahun)

$D_3 = 1$  ; Fakultas teknik  
 $0$  ; lainnya

Dari model didapatkan:

- Rata-rata gaji dosen perempuan yang mengajar diluar fakultas teknik =  $\alpha_1 + \beta X$
- Rata-rata gaji dosen laki-laki yang mengajar diluar fakultas teknik =  $\alpha_1 + \alpha_2 + \beta X$
- Rata-rata gaji dosen perempuan yang mengajar di fakultas teknik =  $\alpha_1 + \alpha_3 + \beta X$
- Rata-rata gaji dosen laki-laki yang mengajar di fakultas teknik =  $\alpha_1 + \alpha_2 + \alpha_3 + \beta X$

# ILUSTRASI

- Seandainya didapat persamaan regresi sebagai berikut:

$$Y = 7,43 + 0,207 D_2 + 0,164 D_3 + 1,226 X$$
$$R^2 = 91,22\%$$

- Apa artinya jika uji-t menunjukkan  $D_2$  dan  $D_3$  signifikan?
- Berapa rata-rata gaji dosen perempuan yang mengajar diluar fakultas teknik dengan pengalaman 1 tahun?  
 $7,43 + 1,226 = \text{Rp.}8,656$  juta.
- Berapa rata-rata gaji dosen laki-laki yang mengajar diluar fakultas teknik dengan pengalaman 1 tahun?  
 $7,43 + 0,207 + 1,226 = \text{Rp.}8,863$  juta.
- Rata-rata gaji dosen perempuan yang mengajar di fakultas teknik dengan pengalaman 1 tahun?  
 $7,43 + 0,164 + 1,226 = \text{Rp.}8,820$  juta.

# Manfaat Lain Variabel Dummy

- Dalam analisis menggunakan data time series, variabel dummy bermanfaat untuk membandingkan suatu kurun waktu dengan kurun waktu tertentu.
- Misalnya:
  - Bagaimana produksi PT Astra antara sebelum terjadi krisis dan saat krisis ekonomi?
  - Bagaimana minat masyarakat untuk menabung di Bank Syariah setelah MUI mengeluarkan fatwa bahwa bunga haram?
  - Apakah benar setiap bulan Desember harga dolar cenderung naik?
  - Apakah benar setiap hari senin harga saham Indofood naik?
- Model diatas: Perbedaan hanya diakomodasi oleh intersep. Bagaimana jika slop juga berbeda → Membandingkan 2 regresi

# MEMBANDINGKAN DUA REGRESI

- Perhatikan persamaan berikut:

$$\text{Tabungan (Y)} = \alpha_1 + \alpha_2 \text{Pendapatan (X)} + u$$

- Apakah hubungannya selalu demikian (sama) pada saat sebelum krisis moneter dan ketika krisis moneter?
- Data dibagi dua berdasarkan kurun waktu, yaitu sebelum dan saat krisis, sehingga didapat dua model regresi, yaitu:
  - Periode I, sebelum krisis:  $Y_i = \alpha_1 + \alpha_2 X_i + u_i$  ;  
 $i = 1, 2, \dots, n$
  - Periode II, sesudah krisis:  $Y_i = \beta_1 + \beta_2 X_i + \varepsilon_i$  ;  
 $i = n+1, n+2, \dots, N$

# MEMBANDINGKAN DUA REGRESI

- Kemungkinan-kemungkinan yang akan didapat:
  - Kasus 1:  $\alpha_1 = \beta_1$  dan  $\alpha_2 = \beta_2$  (model sama)
  - Kasus 2:  $\alpha_1 \neq \beta_1$  dan  $\alpha_2 = \beta_2$
  - Kasus 3:  $\alpha_1 = \beta_1$  dan  $\alpha_2 \neq \beta_2$
  - Kasus 4:  $\alpha_1 \neq \beta_1$  dan  $\alpha_2 \neq \beta_2$  (pergeseran model)

# MEMBANDINGKAN DUA REGRESI

- Untuk menanggulangi permasalahan diatas → variabel dummy
- Model:

$$Y_i = \alpha_1 + \alpha_2 D + \beta_1 X_i + \beta_2 D X_i + u_i$$

D = 1 ; pengamatan pada periode I (Sebelum Krisis)

0 ; pengamatan pada periode II (Saat Krisis)

- Sehingga, rata-rata tabungan (Y) pada periode :

I :  $Y_i = (\alpha_1 + \alpha_2) + (\beta_1 + \beta_2) X_i$

II :  $Y_i = \alpha_1 + \beta_1 X_i$

# MEMBANDINGKAN DUA REGRESI

- Dengan demikian:
- Kasus 1: Bila  $\alpha_2 = 0$  dan  $\beta_2 = 0$   $\Rightarrow$  Model I = Model II
- Kasus 2: Bila  $\alpha_2 \neq 0$  dan  $\beta_2 = 0$   $\Rightarrow$  Slope sama, intercept beda
- Kasus 3: Bila  $\alpha_2 = 0$  dan  $\beta_2 \neq 0$   $\Rightarrow$  Intercept sama, slope beda
- Kasus 4: Bila  $\alpha_2 \neq 0$  dan  $\beta_2 \neq 0$   $\Rightarrow$  Intercept dan slope berbeda

