IPB University
Bogor Indonesia

Statistics and Data Science Study Program
School of Data Science, Mathematics, and Informatics

DIKTISAINTEK BERDAMPAK

PRAKTIKUM 4

Analisis Regresi Linier Berganda (Part 1)

Cameliya Ulya Hidayah (G1~026)
Helmi Falah (G1~049)

Model Regresi Linier Berganda

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \varepsilon$$

- **Linier** dalam parameter
- Memiliki lebih dari satu peubah penjelas/peubah bebas
- Hubungan antara setiap **X dan Y** dinyatakan dalam **fungsi linier** atau **berordo/berderajat satu**
- Peubah penjelas memiliki **pangkat sama dengan 1**

Model Regresi Linier Berganda

Model Regresi Linier Berganda dalam notasi matriks dapat dituliskan sebagai berikut:

$$Y = X\beta + \varepsilon$$
$$Y_{n\times 1} = X_{n\times(k+1)}\beta_{(k+1)\times 1} + \varepsilon_{n\times 1}$$

$$\boldsymbol{y}_{n\times 1} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \boldsymbol{x}_{n\times(k+1)} = \begin{bmatrix} 1 & X_{11} & \dots & X_{k1} \\ 1 & X_{12} & \dots & X_{k2} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & X_{1n} & \dots & X_{kn} \end{bmatrix}, \boldsymbol{\beta}_{(k+1)\times n} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \boldsymbol{\varepsilon}_{n\times 1} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$$

Sifat-sifat Penduga MKT

1. $\hat{\beta}$ merupakan penduga tak bias bagi β
2. $Var(\hat{\beta}) = \sigma^2(X'X)^{-1}$, misalkan $C = (X'X)^{-1}$ maka $Var(\hat{\beta}) = \sigma^2 C$
3. Penduga $\hat{\beta}_j$ menyebar normal dengan nilai tengah β_j dan ragam $\sigma^2 c_{jj}$
4. Dugaan bagi ragam y atau ragam $\varepsilon(\sigma^2)$

$$\hat{\sigma}^2 = \frac{SSResidual}{n-p} = \frac{y'y - \hat{\beta}x'y}{n-p}$$

Asumsi Regresi Linier Berganda

1. Kondisi Gauss-Marcov

- $E[\varepsilon_i] = 0$, nilai harapan/rataan galat sama dengan nol
- $E[\varepsilon_i^2] = var[\varepsilon] = \sigma^2 I$, ragam sisaan homogen untuk setiap x (homoskedastisitas)
- $E[\varepsilon_i \varepsilon_j] = 0$, sisaan saling bebas/tidak ada autokorelasi

1. Galat menyebar normal
2. Galat bebas terhadap peubah bebas, $cov(x_i, \varepsilon_i) = 0$
3. Tidak ada multikolineritas pada peubah bebas

Interpretasi Koefisien Regresi

$$\hat{y} = \widehat{\beta_0} + \widehat{\beta_1}X_1 + \widehat{\beta_2}X_2 + \cdots + \widehat{\beta_k}X_k$$

$\widehat{\beta_0}$ disebut sebagai **intersep** yang merupakan **nilai dugaan rataan y ketika semua peubah X bernilai 0** atau dugaan nilai harapan Y yang tidak dipengaruhi oleh peubah penjelas X, *jika x sama dengan 0 terdapat dalam selang pengamatan*

$\widehat{\beta_j}$, $\mathrm{j} = 1,2,\ldots,\mathrm{k}$ adalah **nilai dugaan perubahan rataan y** atau nilai harapan y jika X_j berubah satu satuan dan *peubah penjelas lainnya dianggap tetap*

Selang Kepercayaan Parameter Regresi

Selang kepercayaan $(1-\alpha)100\%$ bagi parameter β_j:

$$\boldsymbol{b_j \pm t_{\frac{\alpha}{2},(n-p)} s \sqrt{c_{jj}}} \qquad , j = 0,1,2,\dots,k$$

$$\boldsymbol{s = \sqrt{\frac{SSRes}{n-p}} = \sqrt{\frac{e'e}{n-(k+1)}} = \sqrt{\frac{\sum e_i^2}{n-(k+1)}}}$$

$\boldsymbol{s}$ adalah dugaan simpangan baku σ. $\boldsymbol{n}$ adalah banyaknya amatan. $\boldsymbol{p}$ adalah banyaknya parameter. $\boldsymbol{k}$ adalah banyaknya peubah penjelas. $\boldsymbol{c_{jj}}$ adalah nilai dari **baris ke-j** dan **kolom ke-j** pada matriks

$$\boldsymbol{(X'X)^{-1}} = \begin{bmatrix} c_{00} & c_{01} & \cdots & c_{0k} \\ c_{10} & c_{11} & \cdots & c_{1k} \\ \vdots & \vdots & \vdots & \vdots \\ c_{ko} & c_{k1} & \dots & c_{kk} \end{bmatrix}.$$

Uji Hipotesis Parameter Regresi (Uji F Simultan)

Hipotesis

$H_0: \beta_1 = \beta_2 = \cdots = \beta_k = 0$ (semua peubah X tidak berpengaruh linier terhadap Y)

$H_0: \beta_1 \neq 0$ (minimal ada satu peubah X yang berpengaruh linier terhadap Y)

Sumber Keragaman	**Derajat Bebas**	**Jumlah Kuadrat**	**Kuadrat Tengah**	**F**
Regresi	k	$b'X'Y - \frac{Y'11Y'}{n}$	JKR/k	KTR/KTG
Galat	n-(k+1)	$Y'Y - b'X'Y$	JKG/n-(k+1)	
Total	n-1	$Y'Y - \frac{Y'11Y'}{n}$		

Tolak H_0 jika $F > F_{k,n-k-1,\alpha}$ atau $p - value < \alpha$

Uji Hipotesis Parameter Regresi (Uji Parsial)

Hipotesis

$H_0: \beta_j = 0$ (Peubah penjelas x_j tidak berpengaruh terhadap Y)

$H_0: \beta_j \neq 0$ (Peubah penjelas x_j berpengaruh terhadap Y setelah peubah penjelas lainnya ada dalam model)

Statistik Uji

$$t = \frac{b_j}{se(b_j)} = \frac{b_j}{\sqrt{s^2 c_{jj}}} \qquad , s^2 = KTG$$

Tolak H_0 jika $|t| > t_{\frac{\alpha}{2}, n-p}$ atau $p - value < \alpha$

Ukuran Kebaikan Model

Koefisien Determinasi

Nilai R^2 digunakan untuk mengukur kesesuaian model linier dengan data yang ada. Nilai R^2 yang tinggi atau mendekati satu tidak selalu berarti bahwa model tersebut cocok dengan data, karna semakin banyak peubah penjelas maka akan semakin tinggi nilai R^2 . Untuk mengatasi hal tersebut digunakan $adj.\,R^2$

$$R^2 = \frac{JKR}{JKT} = 1 - \frac{JKG}{JKT} = 1 - \frac{\sum(y_i - \hat{y})^2}{\sum(y_i - \bar{y})^2}$$

$$Adj.\,R^2 = 1 - \frac{\frac{JKG}{n-k-1}}{\frac{JKT}{n-1}} = 1 - \frac{n-1}{n-k-1}(1 - R^2)$$

IPB University
Bogor Indonesia

DIKTISAINTEK
BERDAMPAK

TERIMA KASIH

IPB University
Bogor Indonesia

DIKTISAINTEK
BERDAMPAK

MARI PRAKTIK