

EKONOMETRIKA I

EKONOMI PEMBANGUNAN KELAS C

DOSEN PENGASUH:

IMAM AWALUDDIN, S.E., M.E.

M. MUFTI HUDANI S.E., M.EC.DEV.

FAKULTAS EKONOMI DAN BISNIS

UNIVERSITAS LAMPUNG

SEMESTER GENAP 2023/2024

Review Analisis Regresi

Analisis Regresi

- Ø Mempelajari hubungan ketergantungan dari satu peubah tak bebas (*dependent*) kepada satu atau lebih peubah penjelas (*explanatory*)
- Ø Menduga rata-rata populasi dari peubah tak bebas berdasarkan pengetahuan mengenai nilai peubah penjelas
- Ø Analisis regresi ada dua komponen:
 - Adanya kemungkinan peubah lain yang tidak terukur yang juga menjelaskan peubah tak bebas (komponen stokastik → error term)
 - Selain hubungan deterministik yang diasumsikan (komponen deterministik)

Bahan Kuliah Ekonometrika I

Analisis Regresi Dua Variabel: Masalah dalam Estimasi

IMAM AWALUDDIN

Review Analisis Regresi Sederhana

- Ø Menduga rata-rata peubah tak bebas berdasarkan nilai peubah (satu) bebas yang diketahui
- Ø Fungsi Regresi Populasi, PRF: garis yang menghubungkan nilai harapan bersyarat untuk seluruh kemungkinan nilai X
- Ø Di dalam analisis regresi sederhana, LINIER berarti linier dalam PARAMETER
- Ø Fungsi Regresi Populasi, PRF secara stokastik

$$E(Y|X_i) = b_1 + b_2 X_i$$

$$Y_i = E(Y|X_i) + u_i = b_1 + b_2 X_i + u_i$$

Review Analisis Regresi Sederhana

- Ø Asumsi utama fungsi regresi populasi adalah nilai harapan error term adalah nol

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

- Ø Faktor gangguan u_i merupakan pengganti dari semua variabel yang dihilangkan dari model, namun secara kolektif mempengaruhi Y .
- Ø Fungsi Regresi Sampel, SRF = regresi yang dibentuk dari sampel, untuk menduga regresi populasi

$$\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$$

- Ø Fungsi Regresi Sampel, SRF secara stokastik

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + \hat{u}_i$$

↓
Faktor residual/sisa, merupakan estimator dari u_i

Ordinary Least Square (OLS)

- Ø Metode OLS (Kuadrat Terkecil Biasa, KTB) dikemukakan oleh Carl Friedrich Gauss, seorang ahli matematika Jerman. → Gaussian → OLS → CLRM.
- Ø Secara umum dan jumlah, OLS merupakan analisis regresi yang paling sering digunakan, terutama menarik secara intuitif dan lebih sederhana secara matematis dibandingkan dengan metode ML dan MM.
- Ø Untuk asumsi tertentu (yang akan kita bahas kemudian), metode OLS memiliki beberapa sifat-sifat statistic yang sangat menarik dan telah membuat metode ini sebagai salah satu metode paling kuat dan dikenal dalam analisis regresi.

Metode-metode Estimasi

Pendugaan parameter pada regresi dengan dua peubah → Menduga PRF dengan SRF.

Ada 3 dengan menggunakan metode estimasi:

1. Ordinary Least Square (OLS) = kuadrat terkecil biasa (KTB)
2. maximum likelihood (ML) =
3. metode momen-momen (method of moments)

Dalam bagian ini kita hanya membahas metode OLS.

Pendugaan Parameter Pada Regresi dengan Dua Peubah

- Ø Menduga PRF dengan SRF
- Ø Menggunakan metode Ordinary Least Square (OLS)

PRF $Y_i = \beta_1 + \beta_2 X_i + u_i$

SRF $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i \rightarrow Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + \hat{u}_i = \hat{Y}_i + \hat{u}_i$

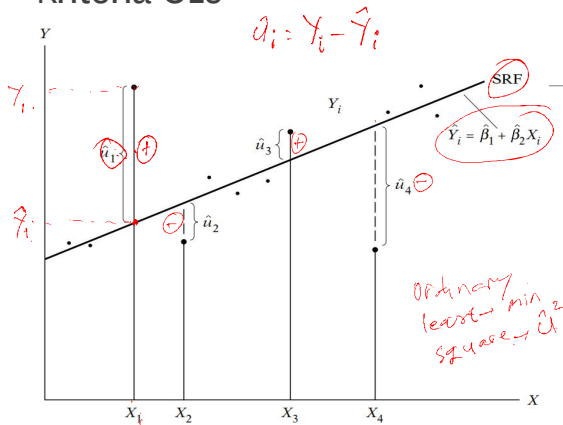
↳ Dari dua definisi tersebut:

$\hat{u}_i = Y_i - \hat{Y}_i$
 $\hat{u}_i = Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i$

Residual adalah perbedaan antara nilai aktual (nyata) dengan nilai estimasi dari Y

$Y_i = \beta_0 + \beta_1 X_i + u_i$
 $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \hat{u}_i$
 $\hat{u}_i = Y_i - \hat{Y}_i$
 $\hat{u}_i = Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i$

Kriteria OLS



- Terdapat n pasang observasi dari Y dan X
- Kita ingin menentukan sedemikian rupa sehingga menghasilkan nilai sedekat mungkin dengan nilai Y aktual
- Terdapat residual untuk setiap X_i
- Prinsip metode OLS adalah memilih SRF sedemikian sehingga jumlah kuadrat dari residual sekecil mungkin

Pendugaan Parameter Pada Regresi dengan Dua Peubah

Diperoleh:

$$\hat{b}_2 = \frac{\hat{a}(X_i - \bar{X})(Y_i - \bar{Y})}{\hat{a}(X_i - \bar{X})^2}$$

Bagaimana rumus tersebut diperoleh?

Prinsip metode *Ordinary Least Square* (OLS):

- Memilih SRF sedemikian sehingga jumlah kuadrat dari residual sekecil mungkin *prinsip*

$$\mathring{a} \hat{u}_i^2 = \mathring{a} (y_i - \hat{b}_1 - \hat{b}_2 x_i)^2$$

- } Penduga parameter model dipilih berdasarkan metode optimasi:
- } Solusi dari turunan pertama dari masing-masing parameter yang disamadengankan nol

$$\begin{aligned}\sum \hat{u}_i^2 &= \sum (Y_i - \hat{Y}_i)^2 \\ &= \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2\end{aligned}$$

Diturunkan terhadap masing-masing parameter dan menyamakan dengan nol diperoleh

$$\frac{\partial (\sum \hat{u}_i^2)}{\partial \hat{\beta}_1} = -2 \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i) = -2 \sum \hat{u}_i = 0 \quad (1)$$

$$\frac{\partial(\sum \hat{u}_i^2)}{\partial \hat{\beta}_2} = -2 \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i) X_i = -2 \sum \hat{u}_i X_i = 0 \quad (2)$$

$$\sum Y_i = n\hat{\beta}_1 + \hat{\beta}_2 \sum X_i$$

$$\sum Y_i X_i = \hat{\beta}_1 \sum X_i + \hat{\beta}_2 \sum X_i^2$$

Ingat rumus derivative:
 $Y = U^n \dots\dots\dots dY/dX = nU^{n-1}(U')$

$$\sum \hat{u}_i^2 = \sum (Y_i - \hat{Y}_i)^2$$

$$= \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2$$

Ingat rumus derivative:
 $Y = U^n \dots \dots \dots dY/dX = nU^{n-1}(U')$

$$\sum n=30$$

$$\beta = 2$$

$$\frac{\partial \sum u^2}{\partial \beta_1} = -2 \sum (Y_i - \beta_1 - \beta_2 X_i) = 0$$

$$\frac{\partial \sum u^2}{\partial \beta_2} = -2 \sum (Y_i - \beta_1 - \beta_2 X_i) X_i = 0$$

$$-2 \sum Y + 2 \beta_1 n + 2 \beta_2 \sum X = 0$$

$$\sum Y = \beta_1 n + \beta_2 \sum X$$

$$\sum Y = \beta_1 n + \beta_2 \sum X$$

$$-2 \sum Y + 2 \beta_1 n + 2 \beta_2 \sum X = 0 \rightarrow \sum Y = \beta_1 n + \beta_2 \sum X$$

$$\hat{\beta}_1 = \frac{\sum X_i^2 \sum Y_i - \sum X_i \sum X_i Y_i}{n \sum X_i^2 - (\sum X_i)^2}$$

$$= \bar{Y} - \hat{\beta}_2 \bar{X}$$

$$\hat{\beta}_2 = \frac{n \sum X_i Y_i - \sum X_i \sum Y_i}{n \sum X_i^2 - (\sum X_i)^2}$$

$$= \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}$$

$$= \frac{\sum x_i y_i}{\sum x_i^2}$$

$$\sum Y = n \beta_1 + \beta_2 \sum X \rightarrow \beta_1 = \frac{\sum Y - \beta_2 \sum X}{n}$$

$$\sum Y X = \beta_1 \sum X + \beta_2 \sum X^2$$

$$\sum Y X = \left(\frac{\sum Y - \beta_2 \sum X}{n} \right) \sum X + \beta_2 \sum X^2$$

$$\sum Y = \bar{Y}$$

$$\sum X = \bar{X}$$

$$\sum Y X = \bar{Y} \bar{X}$$

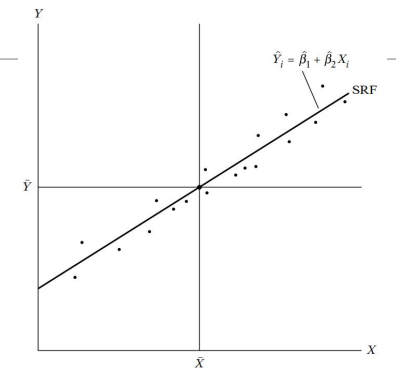
$$\sum Y X = \frac{\sum X \sum Y}{n} - \beta_2 \frac{(\sum X)^2}{n} + \beta_2 \sum X^2$$

$$n \sum Y X - \sum X \sum Y = \beta_2 (n \sum X^2 - (\sum X)^2)$$

$$n \sum Y X - \sum X \sum Y = \beta_2 (n \sum X^2 - (\sum X)^2)$$

$$n \sum Y X - \sum X \sum Y = \beta_2 (n \sum X^2 - (\sum X)^2)$$

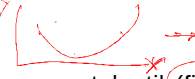
$$\beta_2 = \frac{n \sum Y X - \sum X \sum Y}{n \sum X^2 - (\sum X)^2}$$



Asumsi-asumsi yang mendasari Metode OLS (Model Gaussian atau CLRM)

Ada 10 asumsi yang mendasari model Gaussian, model standar, atau CLRM (Classical Linear Regression Model).

1. Regresi linier pada parameter
2. Nilai peubah bebas (eksogen) dianggap non stokastik (fixed)
3. Galat mempunyai nilai harapan nol
4. Homoskedastisitas: ragam yang sama pada galat
5. Galat tidak saling berkorelasi (tidak ada otokorelasi)
6. Peubah bebas (eksogen) dan galat saling bebas



Stokastik = random
= tidak tetap

Regresi Linier Pada Parameter

- Ø Hanya parameter yang bersifat linier
- Ø Peubah eksogen/bebas atau endogen/terikat boleh tidak linier

1. $Y_i = b_1 + b_2 X_i + u_i \rightarrow$ lin. parameter & lin. variabel
2. $Y_i = b_1 + b_2 X_i^2 + u_i \rightarrow$ lin. parameter, variabel non linier
3. $\ln Y_i = b_1 + b_2 \ln X_i + u_i \rightarrow$ lin. parameter, variabel non linier

$Y = b_1 + b_2 \frac{1}{X} \rightarrow$ lin. parameter, variabel non linier

Asumsi-asumsi yang mendasari Metode OLS

7. Jumlah pengamatan harus lebih besar daripada jumlah parameter yang akan diduga
8. Nilai peubah bebas harus bervariasi
9. Model regresi harus dispesifikasikan dengan tepat: *no specification bias*
10. Tidak ada multikolinieritas sempurna

Catatan: tambahan asumsi no. 10 adalah perluasan asumsi model regresi majemuk

$n > k$ lebih parameter

Nilai peubah bebas (eksogen) dianggap non stokastik (fixed)

- Ø Variabel X adalah variabel selalu tetap, yang non stokastik atau tidak random
- Ø Nilai X selalu dianggap tetap atau nilai X yang independen terhadap faktor kesalahan $\rightarrow \text{cov}(X_i, u_i) = 0$
- Ø Jika variabel-variabel independen (X) bersifat stokastik, maka model yang dihasilkan disebut neo-classical linear regression model (NLRM)

Galat mempunyai nilai harapan nol

Dengan syarat nilai X tertentu, galat mempunyai rata-rata atau nilai harapan sebesar nol

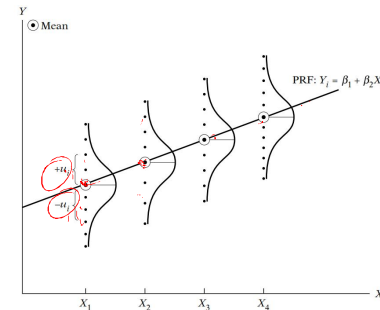
$$E(u_i | X_i) = 0$$

Atau jika X bersifat non stokastik

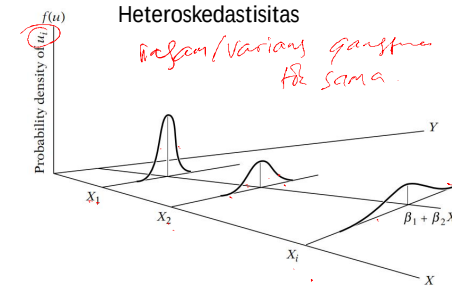
$$E(u_i)$$

$\rightarrow$ residual (a).
rata-rata = 0

Distribusi bersyarat dari faktor gangguan u_i



Ilustrasi grafis asumsi Heterokesdastisitas



Heteroskedastisitas

ragam/Varian gangguan
tidak sama.

$$\text{var}(u_i | X_i) = E(u_i^2 | X_i) = \sigma_{u_i}^2$$

$$X_1 < X_2 < X_3$$

$$\sigma_1^2 < \sigma_2^2 < \sigma_3^2$$

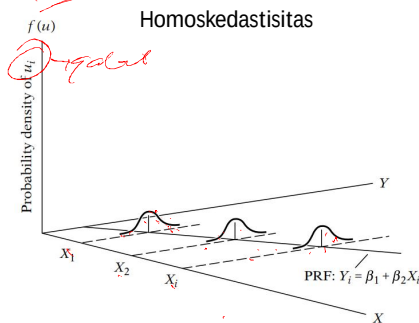
Homokedastisitas: ragam yang sama pada galat

Pada setiap nilai X , populasi Y mempunyai ragam yang sama

$$\begin{aligned} \text{var}(u_i) &= E[u_i - E(u_i | X_i)]^2 \\ &= E(u_i^2 | X_i) \text{ karena asumsi 3} \\ &= E(u_i^2), \text{ jika } X \text{ bersifat nonstokastik} \end{aligned}$$

$$\text{var}(u_i) = \sigma^2$$

$\rightarrow$ varians sama.
di mana var adalah varians



Pada kasus heterokesdastisitas

- ✗ Ragam galat meningkat seiring dengan meningkatnya nilai X
- ✗ Nilai-nilai Y pada X_1 lebih terpusat di garis regresi populasi (PRF) daripada nilai-nilai Y di X yang lainnya
- ✗ Pengamatan Y berasal dari $X = X_1$ akan lebih mungkin terletak di dekat PRF daripada Y yang berasal dari X yang lainnya.
- ✗ Pengamatan pada $X = X_1$ lebih akurat daripada pengamatan pada X selainnya.

Implikasi dari asumsi Homokedastisitas

Dari asumsi homokedastisitas, berlaku bahwa:

- Ragam dari Y dengan syarat nilai X juga sama untuk setiap kemungkinan nilai X

$$\text{var}(Y_i|X_i) = \text{Var}(b_1 + b_2 X_i + u_i|X_i)$$

Konstanta

Ragam dari konstanta adalah nol, dan kedua suku saling bebas, sehingga

$$\text{var}(Y_i|X_i) = \text{var}(u_i|X_i) = s^2$$

Galat Tidak Berkorelasi

Asumsi ini disebut dengan 'tidak ada autokorelasi' antar galat

Pada nilai X tertentu, penyimpangan nilai Y dari rata-rata tidak mempunyai pola tertentu (acak).

Jika terdapat autokorelasi, maka Y tidak hanya dipengaruhi oleh X , tapi juga dipengaruhi oleh galat dari X yang lainnya

$$Y_i = E(Y|X_i) + u_i = b_1 + b_2 X_i + u_i$$

$$u_i = f(u_{i-1})$$

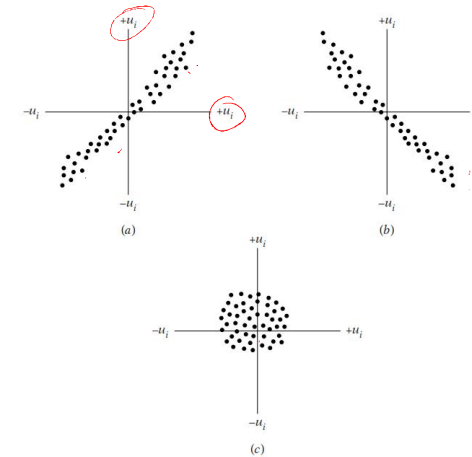
Galat Tidak Saling Berkorelasi (tidak ada otokorelasi di antara galat)

Pada dua nilai X yang berbeda, X_i dan X_j ($i \neq j$), korelasi / kovarians antar galat u_i dan u_j manapun ($i \neq j$) adalah 0.

$$\text{cov}(u_i, u_j | X_i, X_j) = 0$$

korelasi $\text{cov}(u_i, u_j) = 0$, jika X_i bersifat nonstokastik

di mana i dan j adalah dua observasi yang berbeda, dan cov adalah covarians yang bermakna korelasi



Pola korelasi antar gangguan:

- korelasi serial positif
- korelasi serial negatif
- korelasi nol

= 0 korelasi

Peubah bebas (eksogen) dan galat saling bebas

- Ø Kovarians di antara galat dan peubah eksogen = 0 atau $\text{cov}(u_i, X_i) = 0$
- Ø PRF dibentuk berdasarkan asumsi bahwa X dan u mempunyai efek aditif (yang terpisah) bagi Y
- Ø Jika kedua efek tersebut berkorelasi
 - Kesulitan dalam menganalisis efek individu dari X dan u
- Ø Jika keduanya tidak saling bebas
 - u semakin besar seiring peningkatan nilai X (korelasi positif)
 - u semakin kecil seiring peningkatan nilai X (korelasi negatif)

Nilai peubah bebas harus bervariasi

- Ø Karena tujuan dari analisis adalah mempelajari perubahan Y seiring dengan perubahan X
- Ø Dari rumus penduga slope model regresi, penyebut akan bernilai nol jika tidak ada variasi dari nilai X
- Ø Tidak ada solusi bagi penduga slope

$$\hat{b}_2 = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2} \rightarrow \neq 0$$

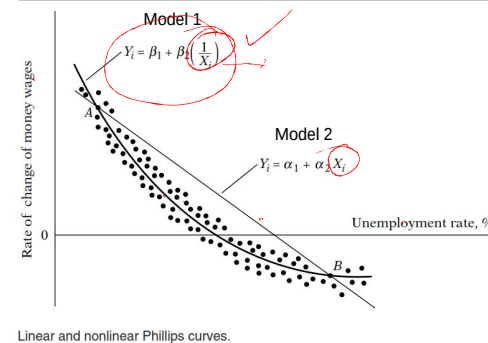
Y	X
10	10
15	10
18	10
24	10
30	10

*Agak aja
Resistensi*

Jumlah pengamatan harus lebih besar daripada jumlah parameter yang akan diduga

- Ø Degree of freedom, df (derajat kebebasan, $dk = db$) harus > 0
- Ø $df = n - k - 1$ (di mana k adalah jumlah variabel bebas)
- Ø Di beberapa buku, $df = n - k$ (di mana k adalah jumlah parameter) *Agus*
- Ø Dua parameter regresi bisa diduga jika dipunyai paling sedikit dua titik

Model regresi harus dispesifikasikan dengan tepat: *no specification bias*



Jika digunakan model 2, maka pada X tertentu, model akan *overestimate* rata-rata Y bagi titik-titik di antara A dan B

Tidak ada multikolinieritas sempurna

Ø Tidak ada hubungan linier di antara peubah-peubah eksogen yang digunakan

Ø Contoh multikolinieritas sempurna:

X_2	X_3
10	50
15	75
18	90
24	120
30	150

$$\hat{a} X_{3i} = 5X_{2i}$$

→ *tidak ada estimasi*
 à hubungan exact linear (multikolinieritas sempurna).

Sifat-sifat Estimator Kuadrat Terkecil: Teori Gauss-Markov

Sebuah estimator, misalkan estimator OLS $\hat{\beta}$, yang dikatakan merupakan estimator terbaik, linier, dan tidak bias β_2 , jika memiliki sifat-sifat berikut:

1. Bersifat linier, di mana merupakan fungsi linier dari sebuah variabel acak seperti variabel dependen Y dalam sebuah model regresi.
2. Bersifat tidak bias, di mana rata-rata atau nilai ekspektasinya $E(\hat{\beta})$ sama dengan nilai sebenarnya β_2 .
3. Memiliki varians minimum dari semua kelompok estimator-estimator yang linier dan tidak bias; sebuah estimator tidak bias dengan varians terkecil dikenal sebagai estimator yang efisien (*efficient estimator*).

Classical Linier Regression Model

- Ø Asumsi-asumsi tersebut disebut dengan asumsi pada *Classical Linier Regression Model* (CLRM)
- Ø Asumsi tersebut mendasari sifat-sifat penduga kuadrat terkecil (OLS) secara statistika.
- Ø Berdasarkan asumsi CLRM, dalam estimator kuadrat terkecil terkandung sifat ideal atau optimum.
- Ø Dinyatakan dalam Teorema Gauss-Markov

1972 1960

Teorema Gauss-Markov

Teorema Gauss-Markov: mengacu pada asumsi-asumsi CLRM, estimator/penduga kuadrat terkecil (OLS), pada estimator-estimator yang linier dan tidak bias, memiliki varians minimum, sehingga estimasinya disebut BLUE (*best linear unbiased estimator*)

Var. min.

Keakuratan dan galat baku dari penduga OLS

- Ø Mempelajari sebaran penarikan contoh dari penduga regresi
- Ø SRF tidak pernah sama dari sampel satu ke sampel yang lain
- Ø Nilai penduga juga tidak pernah sama dari satu sampel ke sampel yang lain
- Ø Penduga dinyatakan akurat jika mempunyai ragam/simpangan baku yang kecil pada sebaran penarikan contohnya.

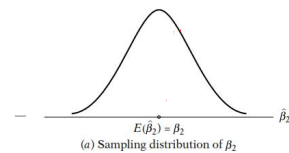
Keakuratan dan galat baku dari penduga OLS

Penduga ragam dari Penduga OLS

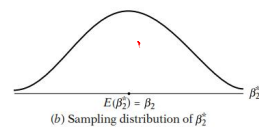
$$\text{var}(\hat{\beta}_2) = \frac{\hat{s}^2}{\hat{a} (X_i - \bar{X})^2} \quad \text{se}(\hat{\beta}_2) = \frac{\hat{s}}{\sqrt{\hat{a} (X_i - \bar{X})^2}}$$

$$\text{var}(\hat{\beta}_1) = \frac{\hat{s}^2 \hat{a} X^2}{n \hat{a} (X_i - \bar{X})^2} \quad \text{se}(\hat{\beta}_1) = \hat{s} \sqrt{\frac{\hat{a} X_i^2}{n \hat{a} (X_i - \bar{X})^2}}$$

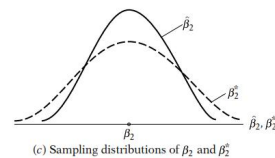
$$\hat{\sigma} = \sqrt{\frac{\sum \hat{u}_i^2}{n-2}} \quad \text{Standard error estimasi atau standard error regresi}$$



Sebaran penarikan sampel penduga 1
- tepat, tidak bias
- Cukup akurat, ragam kecil



Sebaran penarikan sampel penduga 2
- tepat, tidak bias
- Kurang akurat, ragam besar



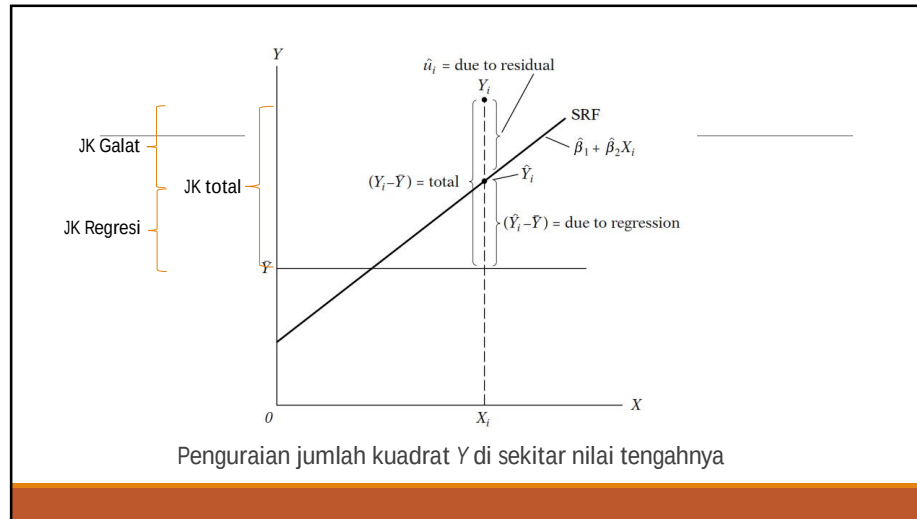
Goodness of Fit dari garis regresi

- Ø Sebagai alat untuk:
 - Menentukan apakah tidak ada alternatif garis lain yang dapat menjelaskan hubungan X dan Y
 - Mengukur seberapa baik model yang diperoleh menjelaskan Y
- Ø Diperlukan penguraian nilai JK (jumlah kuadrat) Y di sekitar nilai tengahnya.

$$\hat{a} (Y_i - \bar{Y})^2 = \hat{a} (\hat{Y}_i + u_i - \bar{Y})^2$$

$$\hat{a} (Y_i - \bar{Y})^2 = \hat{a} (\hat{Y}_i - \bar{Y})^2 + \hat{a} (Y_i - \hat{Y}_i)^2$$

JK total	JK Regresi	JK Residual/Galat
TSS	ESS	RSS



Rentang Nilai Koefisien Determinasi

Dari hubungan: $1 = R^2 + \frac{\text{JK Galat}}{\text{JK Total}}$

} Jika model regresi gagal menjelaskan keragaman nilai Y maka:

$$\text{JK Galat} = \text{JK Total} \quad R^2 = 0$$

} Jika model regresi menjelaskan keragaman nilai Y dengan sempurna maka:

$$\text{JK Galat} = 0 \quad R^2 = 1$$

$$0 \leq R^2 \leq 1$$

Koefisien Determinasi (R^2)

Dari penguraian JK tersebut dapat diturunkan koefisien determinasi berikut:

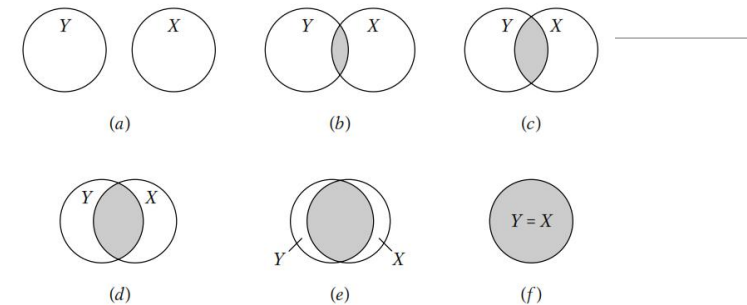
$$R^2 = \frac{\text{JK Regresi}}{\text{JK Total}}$$

} Koefisien determinasi: Sebagai ukuran seberapa besar (dalam proporsi/persen) keragaman total Y dapat dijelaskan oleh model regresi.

$$\text{JK Total} = \text{JK Regresi} + \text{JK Galat}$$

$$1 = \frac{\text{JK Regresi} + \text{JK Galat}}{\text{JK Total}} \quad \longrightarrow \quad 1 = R^2 + \frac{\text{JK Galat}}{\text{JK Total}}$$

Rentang Nilai Koefisien Determinasi



Pandangan Ballentine (diagram Venn) mengenai r^2 : (a) $r^2 = 0$; (f) $r^2 = 1$.

Koefisien Determinasi dan Koefisien Korelasi

$$1 = \frac{ESS}{TSS} + \frac{RSS}{TSS}$$

$$= \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} + \frac{\sum \hat{u}_i^2}{\sum(Y_i - \bar{Y})^2}$$

$$r^2 = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} = \frac{ESS}{TSS}$$

$$r^2 = 1 - \frac{\sum \hat{u}_i^2}{\sum(Y_i - \bar{Y})^2}$$

$$= 1 - \frac{RSS}{TSS}$$

Koefisien Korelasi

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

$$= \frac{n \sum X_i Y_i - (\sum X_i)(\sum Y_i)}{\sqrt{[n \sum X_i^2 - (\sum X_i)^2][n \sum Y_i^2 - (\sum Y_i)^2]}}$$

Koefisien Determinasi

$$r^2 = \frac{[\sum(Y_i - \bar{Y})(\hat{Y}_i - \bar{Y})]^2}{\sum(Y_i - \bar{Y})^2 \sum(\hat{Y}_i - \bar{Y})^2}$$

Uji Keberartian Penduga OLS

$$H_0 : b_i = 0$$

$$H_1 : b_i \neq 0$$

Uji satu arah jika
dipunyai wawasan
'a priori'

$$H_0 : b_i = 0$$

$$H_1 : b_i > 0$$

Statistik uji:

$$t = \frac{\hat{b}_i - b_i}{se(\hat{b}_i)} = \frac{\hat{b}_i}{se(\hat{b}_i)}$$

} Tolak atau terima H_0 berdasarkan nilai p untuk tingkat nyata tertentu dan sifat uji, satu arah atau dua arah

Uji Hipotesis dan Selang Kepercayaan

Dengan asumsi *Classical Linier Regression Model* (CLRM)
penduga OLS menyebar secara normal:

$$\hat{b}_1 \sim N(b_1, \text{var}(\hat{b}_1))$$

$$\hat{b}_2 \sim N(b_2, \text{var}(\hat{b}_2))$$

$$\text{var}(\hat{b}_1) = \frac{s^2 \mathbf{a} X^2}{n \mathbf{a} (X_i - \bar{X})^2} \quad \text{var}(\hat{b}_2) = \frac{s^2}{\mathbf{a} (X_i - \bar{X})^2}$$

Selang Kepercayaan

Selang di mana nilai β yang sebenarnya terletak, pada tingkat kepercayaan tertentu

$$\hat{b}_i \pm t_{\frac{\alpha}{2}, n-2} se(\hat{b}_i)$$

Contoh Data 2.6 Buku Gujarati

Obs	Tahun pendidikan X	Rerata Upah, \$ Y	$x - \bar{x}$	y	x^2	y^2	xy	x^2	Y estimate	residual	residual ²
1	6	4.4567	-6	-4.21801	36	17.79159	25.30895	36	4.330127	0.126573	0.016021
2	7	5.77	-5	-2.90471	25	8.437327	14.52354	49	5.054224	0.715776	0.512335
3	8	5.9787	-4	-2.69601	16	7.268457	10.78403	64	5.778321	0.200379	0.040152
4	9	7.3317	-3	-1.34301	9	1.80367	4.029023	81	6.502418	0.829282	0.687709
5	10	7.3182	-2	-1.35651	4	1.840113	2.713015	100	7.226514	0.091686	0.008406
6	11	6.5844	-1	-2.09031	1	4.369386	2.090308	121	7.950611	-1.36621	1.866532
7	12	7.8182	0	-0.85651	0	0.733605	0	144	8.674708	-0.85651	0.733605
8	13	7.8351	1	-0.83961	1	0.704941	-0.83961	169	9.398804	-1.5637	2.445171
9	14	11.0223	2	2.347592	4	5.51119	4.695185	196	10.1229	0.899399	0.808918
10	15	10.6738	3	1.999092	9	3.99637	5.997277	225	10.847	-0.1732	0.029997
11	16	10.8361	4	2.161392	16	4.671617	8.645569	256	11.57109	-0.73499	0.540217
12	17	13.615	5	4.940292	25	24.40649	24.70146	289	12.29519	1.319809	1.741895
13	18	13.531	6	4.856292	36	23.58357	29.13775	324	13.01929	0.511712	0.261849
Jumlah	156	112.7712	0	0	182	105.1183	131.7856	2054	112.7712	0	9.69281
Rerata	12	8.674707692									
carilah	$\hat{\beta}_2$	0.724096703		$\text{var } \hat{\beta}_2$	0.004842		$\text{se}(\hat{\beta}_2)$	0.069581			
	$\hat{\beta}_1$	-0.014452747		$\text{var } \hat{\beta}_1$	0.764967		$\text{se}(\hat{\beta}_1)$	0.874624			
	$\hat{\sigma}^2$	0.881164513									
	$\hat{\sigma}$	0.938703634									
	r^2	0.90779144									
	r	0.952780898									

Tugas

Kerjakan salah satu contoh 3.1, 3.2, dan 3.3 (bagian a dan b) buku Gujarati dengan cara manual (seperti contoh sub bab 3.6) dengan menggunakan bantuan MS Excel.

Jawaban ditulis tangan per kelompok, discan dan disatukan dalam format .pdf serta diunggah di assignment pertemuan ke-3 Vclass.

Output Estimasi dengan Excel

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.952781							
R Square	0.907791							
Adjusted R Square	0.899409							
Standard Error	0.938704							
Observations	13							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	1	95.4255185	95.42552	108.2947816	4.96E-07			
Residual	11	9.692809647	0.881165					
Total	12	105.1183281						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-0.01445	0.874623916	-0.01652	0.987111847	-1.93949	1.910582	-1.93949	1.910582
X	0.724097	0.069581343	10.40648	4.95788E-07	0.570949	0.877244	0.570949	0.877244

Sumber Bacaan

Gujarati, Damodar N. and Dawn C. Porter (2009). *Basic Econometrics*, Fifth Edition. McGraw-Hill/Irwin. New York. Chapter 3.

Gujarati, Damodar N. (2004). *Basic Econometrics*, Fourth Edition. McGraw-Hill/Irwin. New York. Chapter 3.