Inferensi dalam Analisis Regresi dan Korelasi

Persyaratan pada uji regresi linier

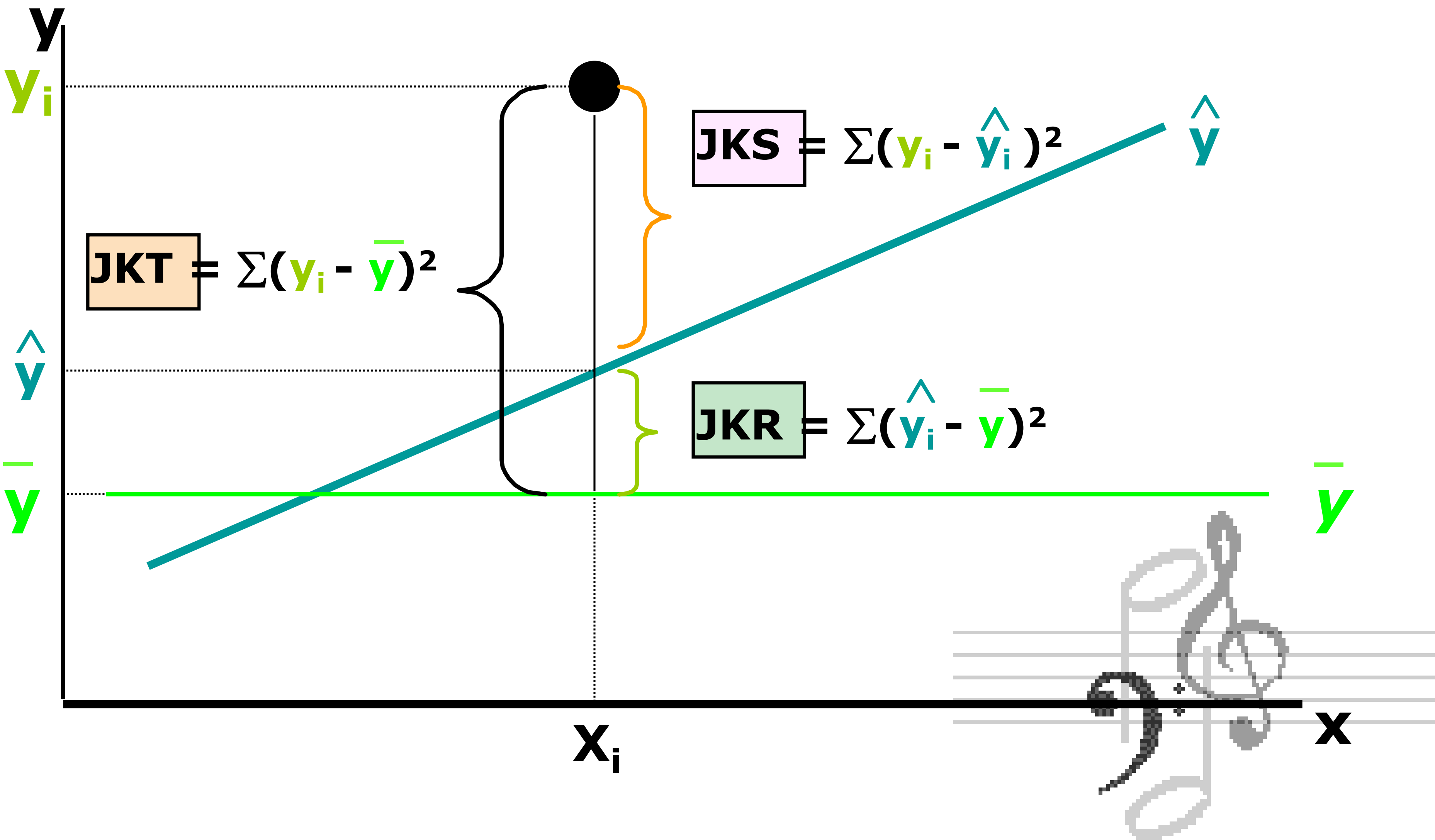
Perhatikan model regresi linier sederhana

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \qquad \varepsilon_i \sim \mathrm{NID}\left(0, \sigma^2\right)$$

Asumsi yang harus dipenuhi

1. Normalitas
2. Independensi
3. Homoskedastisitas

Variasi yang diterangkan dan Yang tidak dapat diterangkan



MENGUJI KOEFISIEN REGRESI DENGAN ANALISIS VARIANSI

i. Susun hipotesis

→ Perhatikan model $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$

$$H_0 : \beta_1 = 0$$

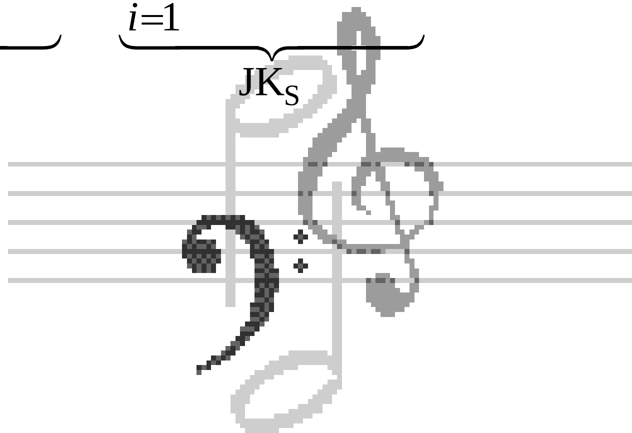
$$H_1 : \beta_1 \neq 0$$

ii. Pilih tingkat signifikansi α

iii. Menyusun Tabel ANAVA

$$\underset{\text{Total}}{\left(y_i - \overline{y}\right)} = \underset{\text{regresi}}{\left(\hat{y}_i - \overline{y}\right)} + \underset{\text{sisa}}{\left(y_i - \hat{y}_i\right)}$$

$$\underbrace{\sum_{i=1}^{n}(y_i - \bar{y})^2}_{\text{JK}_\text{T}} = \underbrace{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}_{\text{JK}_\text{R}} + \underbrace{2\sum_{i=1}^{n}(\hat{y}_i - \bar{y})(y_i - \hat{y}_i)}_{=0 \text{ (buktikan !!!)}} + \underbrace{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}_{\text{JK}_\text{S}}$$



Tabel ANAVA

Sumber Variasi	JK	dk	RK	FO
Regresi	JKR= $b^2\sum_{i=1}^{n}(x_i-\bar{x})^2$	1	RKR =JKR/1	F=RKR/RKS
Sesatan	JKS= JKT-JKR	n-2	RKS =JKS/n-2	Ftabel F(α, 1,n-2)
Total	JKT= $\sum_{i=1}^{n}y_i^2-\frac{\left(\sum y_i\right)^2}{n}$	n-1		

DK. Tolak H0 jika F0>Ftabel =Fα,1,n-2

Uji Keberartian Regresi (1)

1. Susun hipotesis

$H_0 : \beta_1 = 0$ (Hubungan linier X dan Y tidak berarti)

$H_1 : \beta_1 \neq 0$ (Hubungan linier X dan Y berarti)

2. Pilih tingkat signifikansi α**=0.05**

3. Susun Anava

Tabel Anava :

Sumber Variasi	JK	dk	RK	F Hitung
Regresi	541.193	1	541.193	29.04
Sesatan	186.557	12-2=10	18.6557	Ftabel F(alpha, 1,n-2)
Total	728.25	12-1=11		

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	541.693	1	541.693	29.036	.000[a]
	Residual	186.557	10	18.656		
	Total	728.250	11			

a. Predictors: (Constant), matematik

b. Dependent Variable: fisika

4. Kesimpulan :

Tolak H0 karena
Fobs=29.04>Ftabel=4.96

d.k.l regresi linier X dan Y berarti

Uji Keberartian Koef. Regresi (2)

1. Susun hipotesis

$$H_0 : \beta = 0$$
$$H_1 : \beta \neq 0$$

2. Pilih tingkat signifikansi
3. Kesimpulan : tolak Ho jika t> t tabel

$$t = \frac{b}{s_b}$$

Contoh Yll

1. Susun hipotesis

$$H_0 : \beta = 0$$
$$H_1 : \beta \neq 0$$

2. Pilih tingkat signifikansi α

3. Kesimpulan : tolak Ho jika t> t tabel=t (α/2, n-2)

$$b = 0.8972$$
$$s_b = 0.166504$$
$$t = \frac{0.8972}{0.166504} = 5.388$$

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	29.529	9.311		3.171	.010		
	matematik	.897	.167	.862	5.389	.000	1.000	1.000

a. Dependent Variable: fisika

Karena t=5.388>2.228 maka H0 ditolak jadi koefisien b berarti.
2.228 diperoleh dari tabel t dengan t(0.025,10)

Koefisien Determinasi, r^2

- Koefisien Determinasi adalah bagian dari variasi total dalam variabel dependen yang dijelaskan oleh variasi dalam variabel independen
- Disebut juga dengan r-squared dan dinotasikan dengan r^2

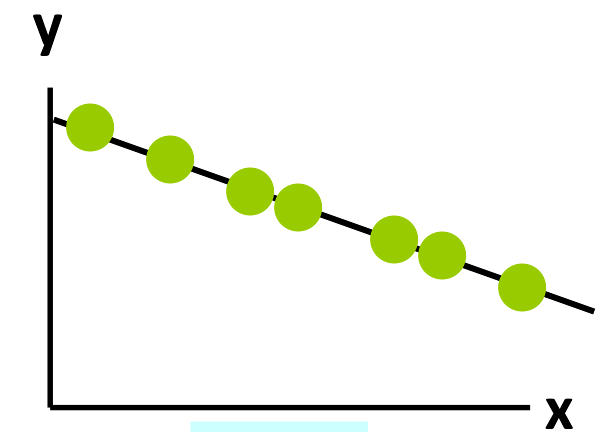
$$\boxed{r^2 = \frac{JK_R}{JK_T} = \frac{\text{Jumlah kuadrat yang dijelaskan oleh regresi}}{\text{Jumlah kuadrat total}}}$$

dengan $\boxed{0 \le r^2 \le 1}$

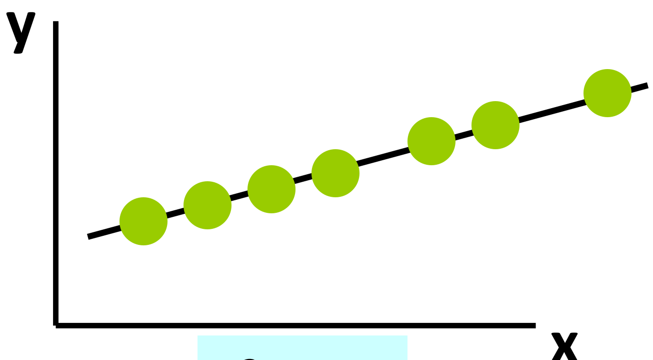
$$\boxed{r = \pm\sqrt{r^2}, \quad -1 \le r \le 1}$$

r = Koefisien Korelasi

Gambaran nilai r^2



$r^2 = 1$

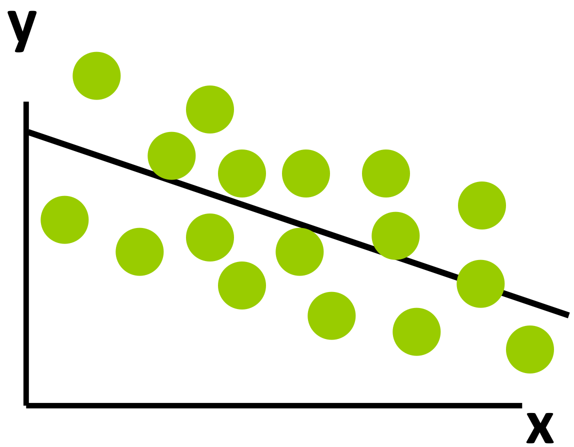


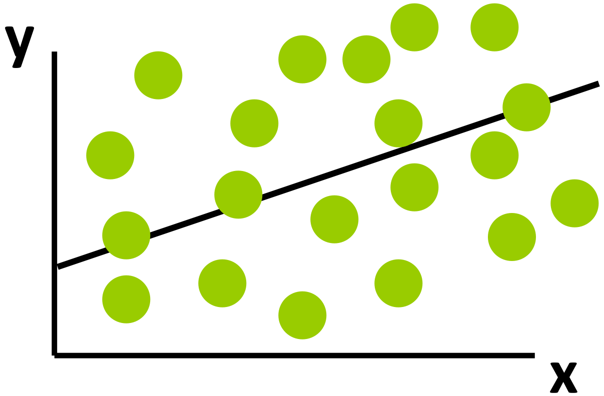
$r^2 = +1$

$r^2 = 1$

Hubungan linier sempurna antara x dan y :
100% variasi dalam y dijelaskan oleh variasi dalam x

Gambaran nilai r^2

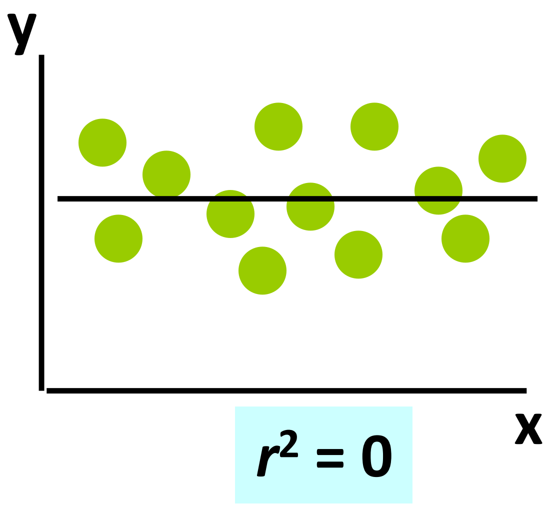




$0 < r^2 < 1$

Hubungan linier antara x dan y lemah :
Beberapa tapi tidak semua variasi y dijelaskan oleh variasi dalam x

Gambaran nilai r^2



$r^2 = 0$

Tidak ada hubungan linier antara x dan y

Nilai Y tidak tergantung x

Contoh yang lalu

Mat (X)	Fis (Y)	XY	X2	Y2
60	80	4800	3600	6400
45	69	3105	2025	4761
50	71	3550	2500	5041
60	85	5100	3600	7225
50	80	4000	2500	6400
65	82	5330	4225	6724
60	89	5340	3600	7921
65	93	6045	4225	8649
50	76	3800	2500	5776
65	86	5590	4225	7396
45	71	3195	2025	5041
50	69	3450	2500	4761
665	951	53305	37525	76095

$$a = 29.529, \quad b = 0.897$$

$$JK_R = b^2 \sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2 = 541.693$$

$$JK_T = \sum_{i=1}^{n} y_i^2 - \frac{\left(\sum y_i\right)^2}{n} = 728.25$$

$$JK_S = JK_T - JK_R = 728.25 - 541.693 = 186.557$$

Jadi bisa kita hitung

$$r^2 = \frac{JK_R}{JK_T} = \frac{541.693}{728.25} = 0.743828$$

jadi persamaan regresi $\hat{Y}_i = 29.5294 + 0.8972 X_i$, nilai Y dapat dijelaskan 74.4% oleh X, sedangkan sisanya 25.6% diterangkan oleh variabel lain yang tidak dimasukkan dalam model

Total variasi dalam Y diturunkan sebanyak 74,4% oleh X

$$\boxed{r = \sqrt{0.744} = 0.863}$$

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.862[a]	.744	.718	4.319

a. Predictors: (Constant), matematik

Kesalahan Baku Taksiran
(*Standard Error of Estimate*)

- Merupakan ukuran variabilitas antara Y dengan nilai Y prediksi

$$s_{y.x} = \sqrt{\frac{\mathrm{JK_S}}{\mathrm{n-2}}}$$

- Contoh yll:

$$\mathrm{JK_S} = 186.557$$

$$s_{y.x} = \sqrt{\frac{186.557}{12-2}} = 4.319$$

Kesalahan Baku Koef. Regresi

definisi

$$s_b = \sqrt{\frac{s_{y.x}^2}{c}}, \quad c = \sum x^2 - \frac{\left(\sum x\right)^2}{n}$$

Contoh yll

$$\sum x^2 = 37525, \quad \left(\sum x\right)^2 = (665)^2 = 442225$$

$$s_b = \sqrt{\frac{s_{y.x}^2}{c}}, \quad c = \sum x^2 - \frac{\left(\sum x\right)^2}{n} = 37525 - \frac{442225}{12} = 672.9167$$

$$s_b = \sqrt{\frac{s_{y.x}^2}{c}} = \sqrt{\frac{18.6557}{672.9167}} = 0.166504$$