



# POLITEKNIK STATISTIKA STIS

*For Better Official Statistics*

Pertemuan 3

## ANALISIS REGRESI

**(Kode MK: K203313)**

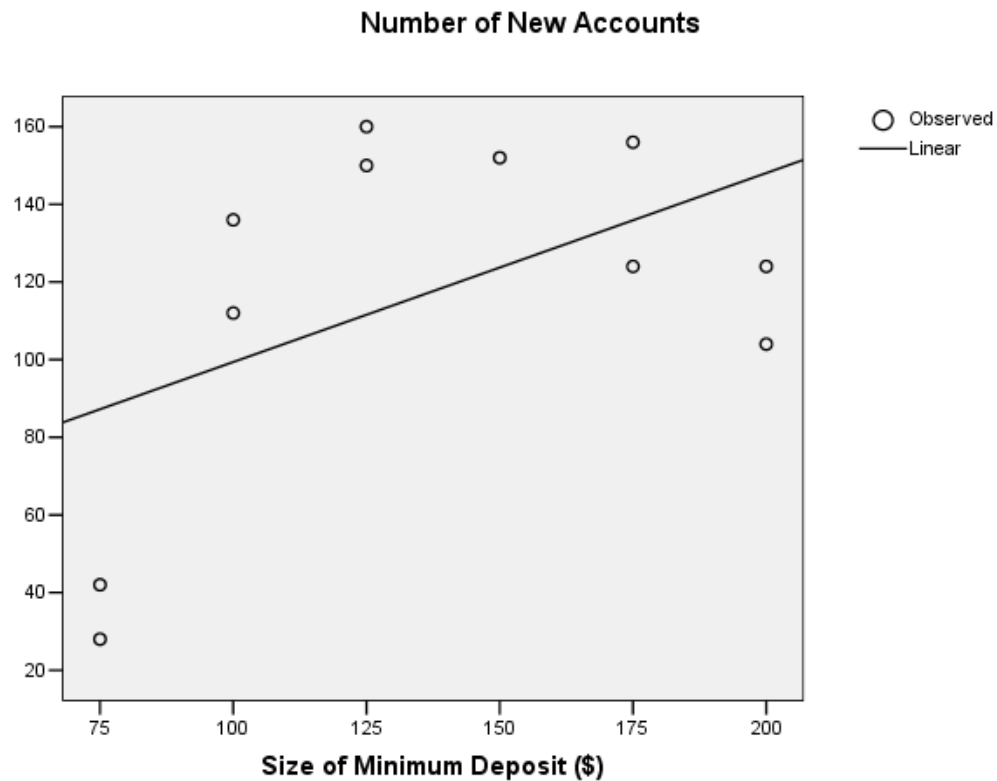
Lack of Fit Test pada RLS

Prodi DIV Komputasi Statistik - Tahun Akademik 2024-2025



# LACK OF FIT TEST

- Digunakan untuk pengujian kecocokan model untuk model dengan amatan berulang pada variabel bebas
- Contoh data yang bisa diperiksa dengan Lack of Fit Test (Kutner et al, p. 122)



# LACK OF FIT TEST

## ASUMSI

- $Y_i$  saling bebas pada setiap nilai  $X$
- $Y_i$  berdistribusi Normal
- $\text{Var}(Y_i) = \sigma^2$

## SYARAT:

Ada amatan berulang pada satu atau lebih nilai  $X$ .

# LACK OF FIT TEST

## HIPOTESIS

$$H_0: E(Y_i) = \beta_0 + \beta_1 X_i$$

Model linier cocok untuk menjelaskan hubungan antara X dan Y

$$H_1: E(Y_i) \neq \beta_0 + \beta_1 X_i$$

Model linier tidak cocok untuk menjelaskan hubungan antara X dan Y

Kesimpulan:

Bila  $H_0$  gagal ditolak, artinya model linier cocok untuk menjelaskan hubungan antara X dan Y

# ANOVA UNTUK LACK OF FIT TEST

Mendekomposisi  $SSE = SSPE + SSLF$

$$\sum \sum (Y_{ij} - \hat{Y}_{ij})^2 = \sum \sum (Y_{ij} - \bar{Y}_j)^2 + \sum \sum (\bar{Y}_j - \hat{Y}_{ij})^2$$

*SSE*                      =                      *SSPE*                      +                      *SSLF*

$$\underbrace{Y_{ij} - \hat{Y}_{ij}}_{\text{Error deviation}} = \underbrace{Y_{ij} - \bar{Y}_j}_{\text{Pure error deviation}} + \underbrace{\bar{Y}_j - \hat{Y}_{ij}}_{\text{Lack of fit deviation}}$$

Source of Variation	<i>SS</i>	<i>df</i>	<i>MS</i>
Regression	$SSR = \sum \sum (\hat{Y}_{ij} - \bar{Y})^2$	1	$MSR = \frac{SSR}{1}$
Error	$SSE = \sum \sum (Y_{ij} - \hat{Y}_{ij})^2$	$n - 2$	$MSE = \frac{SSE}{n - 2}$
Lack of fit	$SSLF = \sum \sum (\bar{Y}_j - \hat{Y}_{ij})^2$	$c - 2$	$MSLF = \frac{SSLF}{c - 2}$
Pure error	$SSPE = \sum \sum (Y_{ij} - \bar{Y}_j)^2$	$n - c$	$MSPE = \frac{SSPE}{n - c}$
Total	$SSTO = \sum \sum (Y_{ij} - \bar{Y})^2$	$n - 1$	

$c$  = banyaknya amatan variable bebas yang berulang

$$F^* = \frac{SSLF}{c - 2} \div \frac{SSPE}{n - c}$$

$$= \frac{MSLF}{MSPE}$$

# ANOVA UNTUK LACK OF FIT TEST

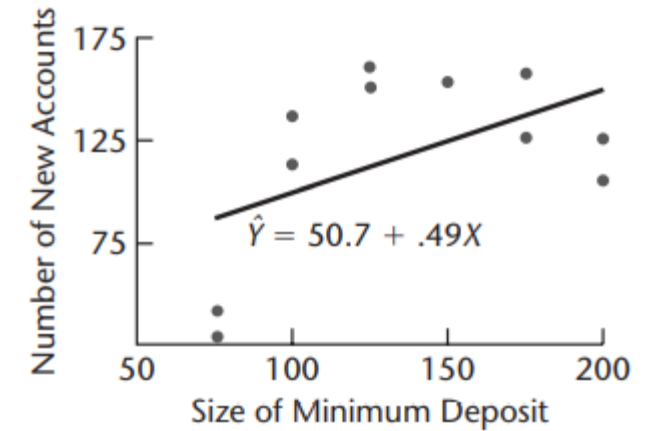
$$\begin{aligned} F^* &= \frac{SSLF}{c - 2} \div \frac{SSPE}{n - c} \\ &= \frac{MSLF}{MSPE} \end{aligned}$$

If  $F^* \leq F(1 - \alpha; c - 2, n - c)$ , conclude  $H_0$

If  $F^* > F(1 - \alpha; c - 2, n - c)$ , conclude  $H_a$

# CONTOH (Bank example, Kutner et al, p. 120)

(a) Data					
Branch $i$	Size of Minimum Deposit (dollars) $X_i$	Number of New Accounts $Y_i$	Branch $i$	Size of Minimum Deposit (dollars) $X_i$	Number of New Accounts $Y_i$
1	125	160	7	75	42
2	100	112	8	175	124
3	200	124	9	125	150
4	75	28	10	200	104
5	150	152	11	100	136
6	175	156			



Replicate	Size of Minimum Deposit (dollars)					
	$j = 1$ $X_1 = 75$	$j = 2$ $X_2 = 100$	$j = 3$ $X_3 = 125$	$j = 4$ $X_4 = 150$	$j = 5$ $X_5 = 175$	$j = 6$ $X_6 = 200$
$i = 1$	28	112	160	152	156	124
$i = 2$	42	136	150		124	104
Mean $\bar{Y}_j$	35	124	155	152	140	114

# CONTOH

$$\begin{aligned}
 SSPE &= (28 - 35)^2 + (42 - 35)^2 + (112 - 124)^2 + (136 - 124)^2 + (160 - 155)^2 \\
 &\quad + (150 - 155)^2 + (152 - 152)^2 + (156 - 140)^2 + (124 - 140)^2 \\
 &\quad + (124 - 114)^2 + (104 - 114)^2 \\
 &= 1,148
 \end{aligned}$$

(b) Bank Example

Source of Variation	SS	df	MS
Regression	5,141.3	1	5,141.3
Error	14,741.6	9	1,638.0
Lack of fit	13,593.6	4	3,398.4
Pure error	1,148.0	5	229.6
Total	19,882.9	10	

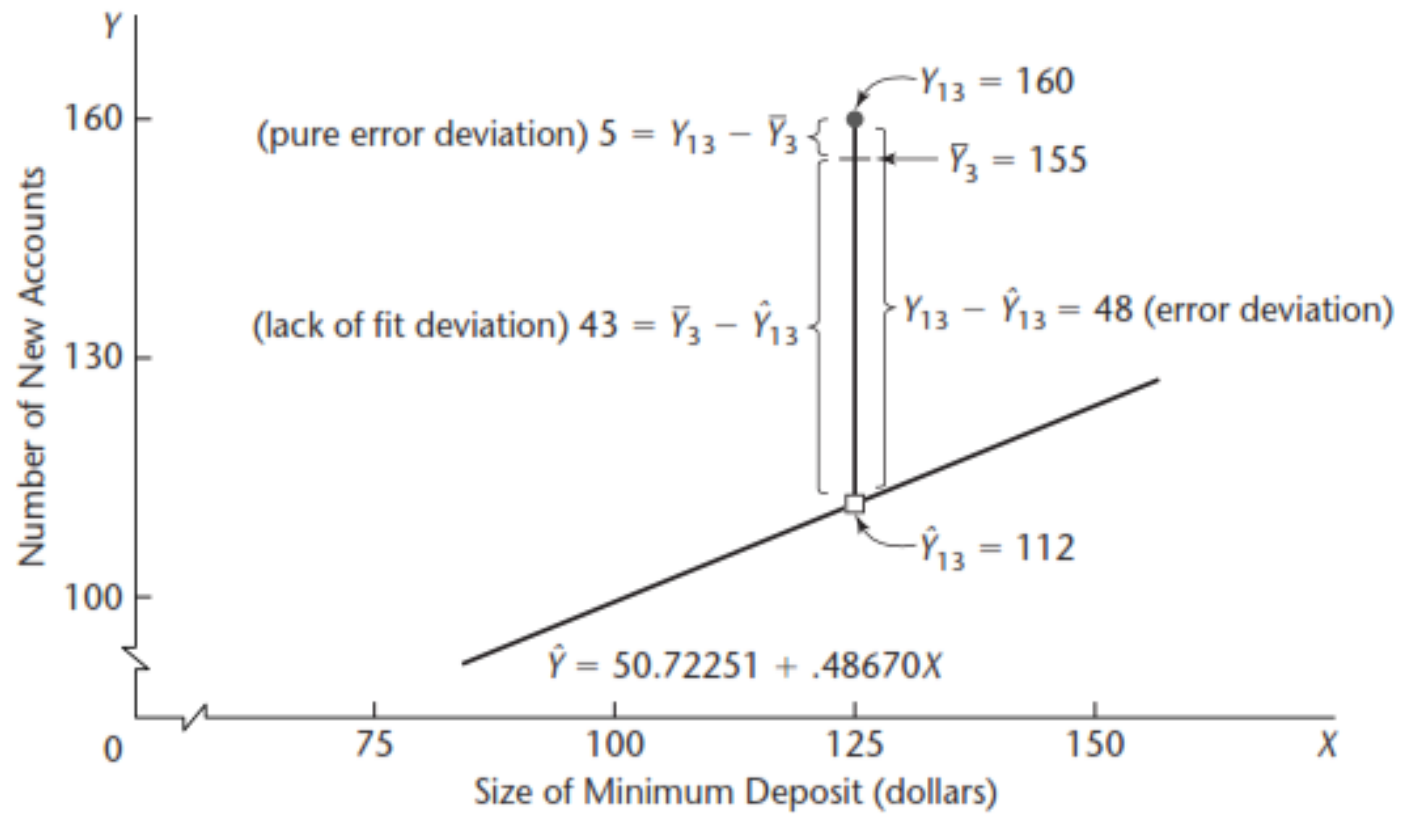
$$\begin{aligned}
 F^* &= \frac{13,593.6}{4} \div \frac{1,148.0}{5} \\
 &= \frac{3,398.4}{229.6} = 14.80
 \end{aligned}$$

Misal  $F_{(0.99;4,5)} = 11,4 \rightarrow \text{Tolak } H_0$

Maka dengan tingkat signifikansi 1 % cukup bukti menyatakan bahawa fungsi regresi tidak linier / model linier tidak cocok menjelaskan hubungan X dan Y.



# CONTOH



# Terimakasih

