



Analisis Regresi Logistik

**Kampus
Merdeka**
INDONESIA JAYA

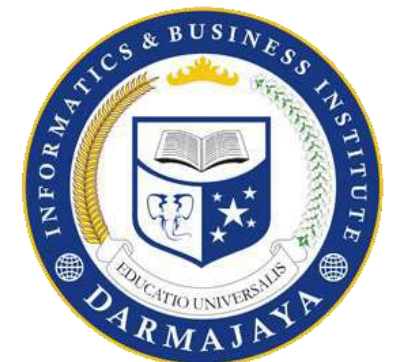
DISUSUN OLEH:

TRUFI MURDIANI

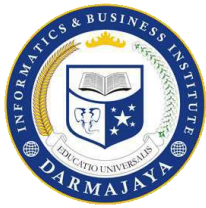
Dosen Program Bisnis Digital
Fakultas Ekonomi dan Bisnis (FEB) Darmajaya

trufimurdiani@darmajaya.ac.id

IG: @mudiani06



REGRESI LOGISTIK



Regresi logistik adalah sebuah pendekatan untuk membuat model prediksi seperti halnya regresi linear atau yang biasa disebut dengan istilah Ordinary Least Squares (OLS) regression. Perbedaannya adalah pada regresi logistik, peneliti memprediksi variabel terikat yang berskala dikotomi. Skala dikotomi yang dimaksud adalah skala data nominal dengan dua kategori, misalnya: Ya dan Tidak, Baik dan Buruk atau Tinggi dan Rendah.

Model persamaan aljabar layaknya OLS yang biasa kita gunakan adalah berikut:

$$Y = B_0 + B_1X + e.$$

Dimana e adalah error varians atau residual. Dengan model regresi ini, tidak menggunakan interpretasi yang sama seperti halnya persamaan regresi OLS. Model Persamaan yang terbentuk berbeda dengan persamaan OLS

ASUMSI REGRESI LOGISTIK



Regresi logistic tidak membutuhkan hubungan linear antara variabel independent dengan variabel dependen

Variabel independent tidak memerlukan asumsi multivariate normality

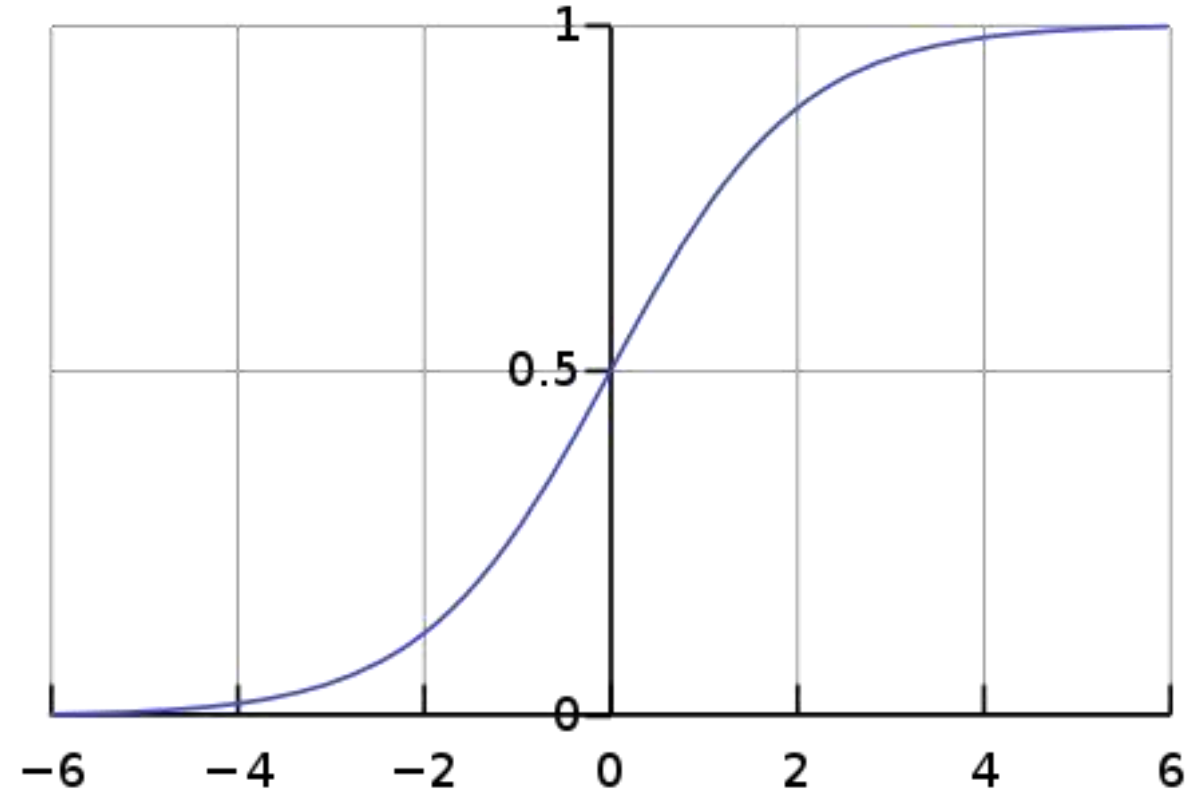
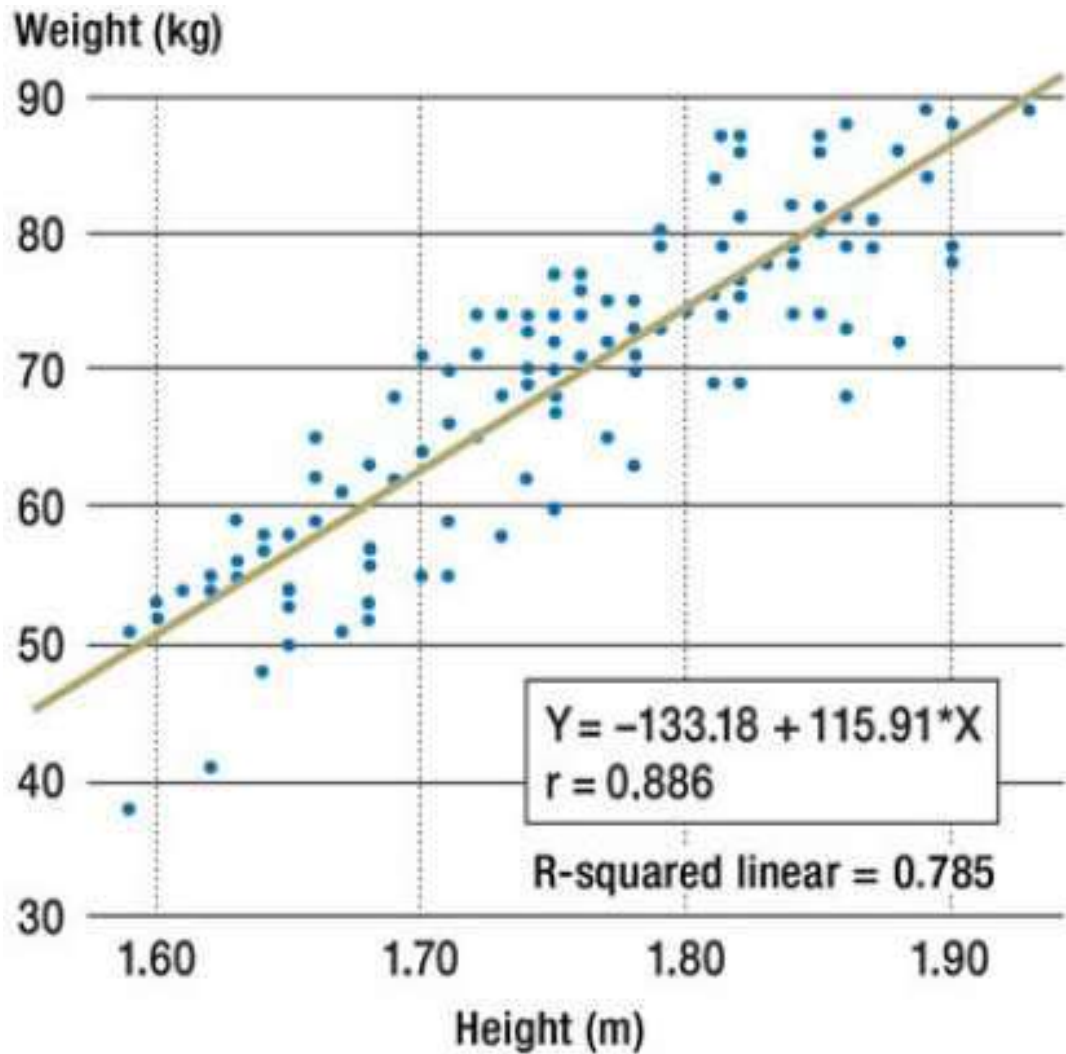
Asumsi Homoskedisitas tidak diperlukan

Variabel Dependen harus bersifat dikotomi (2 kategori, missal: tinggi dan rendah, Ya dan Tidak)

Variabel independent harus terpisah satu sama lain atau bersifat eksklusif

Sampel yang diperlukan dalam jumlah relative besar, minimum dibutuhkan hingga 50 sampel data untuk sebuah variabel predictor (independent)

PERBEDAAN LINEAR DAN LOGISTIK



Regresi Logistik

Uji Signifikansi dan model



□ Pengujian secara statistik untuk melihat pengaruh secara bersama-sama dari variabel bebas terhadap variabel terikat digunakan Chi - Square. Jika nilai Chi - Square lebih besar dari nilai signifikansi (dari analisis), maka dapat diartikan ada pengaruh secara simultan dari variabel bebas terhadap variabel terikatnya.

□ Sedangkan untuk menguji secara parsial dilakukan dengan membandingkan antara nilai *Wald Statistic* dengan nilai signifikansinya (dari analisis). Jika nilai *Wald Statistic*nya lebih besar, maka dapat diartikan ada pengaruh secara parsial. Atau dapat juga dilihat berdasarkan nilai alpha yang ditentukan oleh peneliti dengan nilai signifikansinya. Jika nilai alpha lebih besar dari pada nilai signifikansinya, maka ada pengaruh.

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i}$$

Untuk mengetahui goodness of fit model yang telah dibuat

$$W = \frac{\hat{\beta}_j}{Se(\hat{\beta}_j)}$$

Untuk mengetahui variabel-variabel X yang mempengaruhi Y

Odds Ratio

- ❑ Nilai Odds Ratio Merupakan ukuran asosiasi yang memperkirakan **seberapa besar kecenderungan pengaruh** variabel bebas X terhadap variabel terikat Y.
- ❑ Jika suatu variabel bebas memiliki koefisien positif, maka nilai odds rasionya akan lebih besar dari satu. Sebaliknya jika suatu variabel bebas memiliki koefisien negative, maka nilai odds rasionya akan lebih kecil dari 1

Odd ratio (Ψ_j) = $\left[\frac{p(X_i)}{1 - p(X_i)} \right] = \exp(\beta_i)$

Keterangan:

- Ψ_j = Nilai odds ratio untuk kejadian responden memilih kembali pesawat yang pernah ditumpangnya (j)
- X_i = Variabel bebas
- P = Peluang variabel terikat (Y) terhadap variabel bebas (X)
- $\exp(\beta_i)$ = Exponential atau nilai pangkat koefisien variabel X ke-i

Interpretasi dari nilai odds ratio adalah sebagai berikut:

- Nilai odds ratio yang lebih dari 1 menunjukkan bahwa peluang responden untuk memilih YA (Y=1) lebih besar daripada responden yang memilih TIDAK (Y=0). Dengan kata lain, responden lebih cenderung memutuskan YA.
- Nilai odds ratio yang sama dengan 1 menunjukkan bahwa peluang responden untuk YA (Y=1) dengan responden memilih TIDAK (Y=0) adalah sama.
- Nilai odds ratio yang kurang dari 1 menunjukkan bahwa peluang responden untuk memilih YA (Y=1) lebih kecil daripada responden yang memilih TIDAK (Y=0). Dengan kata lain, responden lebih cenderung memutuskan memilih TIDAK.

Contoh Kasus



Perusahaan Penerbangan X melakukan sebuah survei mengenai loyalitas pelanggan. Perusahaan tersebut menyebarkan pertanyaan kepada 100 penumpang isi pertanyaan tersebut adalah:

Nama:.....

Umur:..... Thn (X1)

Harga Tiket:..... US\$ (X2)

Jika Cuaca sedang mendung/gelap di area penerbangan apakah anda masih ingin melanjutkan penerbangan?

YA/TIDAK

	Y	X1	X2	var
1	0	32	43	
2	1	73	88	
3	0	40	37	
4	0	32	35	
5	0	31	50	
6	1	78	89	
7	1	85	102	
8	0	35	39	
9	1	88	97	
10	0	33	28	
11	1	76	95	
12	1	72	88	

Keterangan Variabel Y:
1 = Tetap naik pesawat
0 = Membatalkan penerbangan

1. Apakah umur dan harga tiket mempengaruhi penumpang untuk tetap terbang/membatalkan penerbangan?
2. Berapa besar peluang penumpang yang sudah tua untuk tetap naik pesawat?
3. Berapa besar peluang penumpang yang telah membeli tiket lebih mahal untuk tetap naik pesawat?

Perintah Regresi Logistik



Regresi Logistik.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

	Y	X1
1	Batal	32
2	Naik	73
3	Batal	40
4	Batal	32
5	Batal	31
6	Naik	78
7	Naik	85
8	Batal	35
9	Naik	88
10	Batal	33
11	Naik	76
12	Naik	72
13	Batal	34
14	Batal	43
15	Batal	37
16	Batal	33
17	Naik	68
18	Batal	27
19	Naik	69
20	Batal	37
21	Naik	65
22	Batal	34
23	Naik	66
24	Naik	72
25	Batal	34
26	Naik	62
27	Naik	64
28	Batal	33

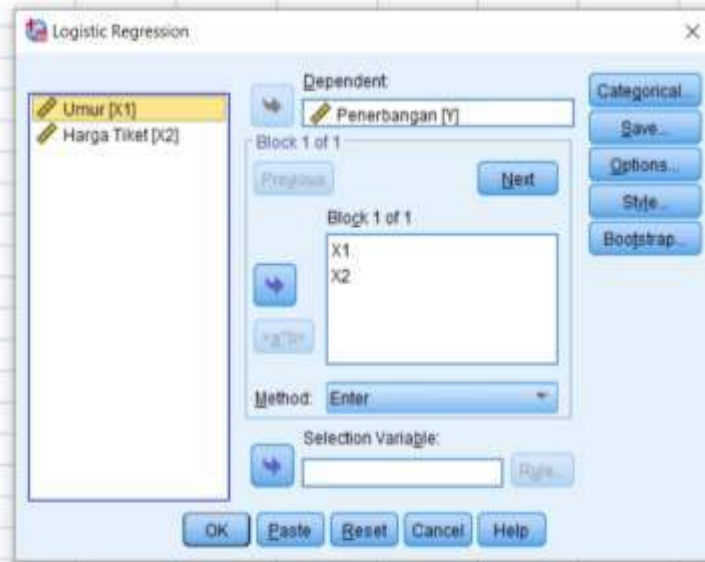
Reports
Descriptive Statistics
Bayesian Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Missing Value Analysis...
Multiple Imputation
Complex Samples
Simulation...
Quality Control
ROC Curve...
Spatial and Temporal Modeling...
Direct Marketing
IBM SPSS Amos...

Automatic Linear Modeling...
Linear...
Curve Estimation...
Partial Least Squares...
Binary Logistic...
Multinomial Logistic...
Ordinal...
Probit...
Nonlinear...
Weight Estimation...
2-Stage Least Squares...
Optimal Scaling (CATREG)...

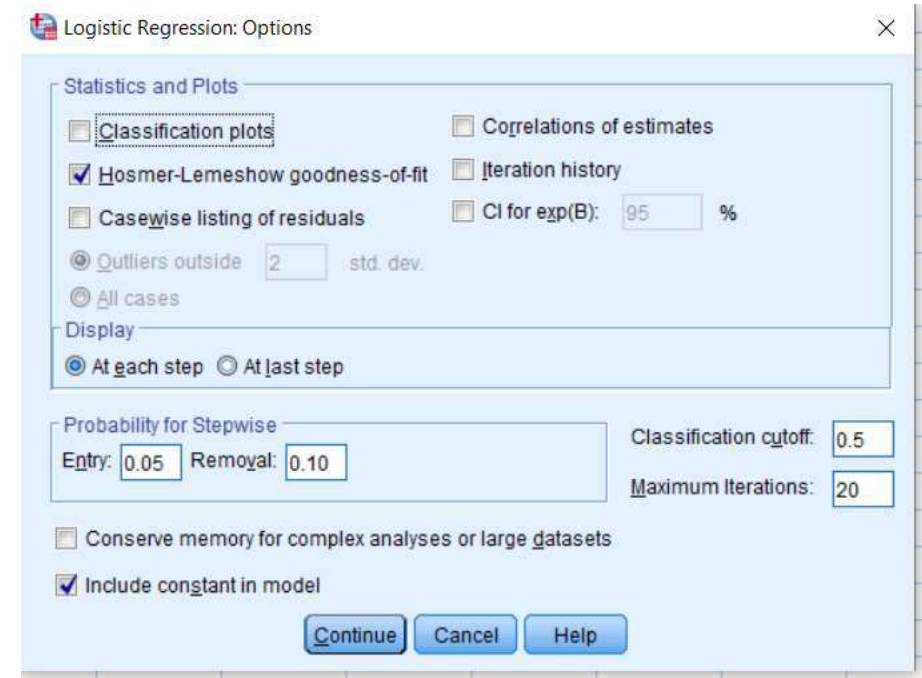
Data View Variable View

Perintah Regresi Logistik

	Y	X1	X2	var	var	var	var	var	var	var	var	var
1	Batal	32	43									
2	Naik	73	88									
3	Batal	40	37									
4	Batal	32	35									
5	Batal	31	50									
6	Naik	78	89									
7	Naik	85	102									
8	Batal	35	39									
9	Naik	88	97									
10	Batal	33	28									
11	Naik	76	95									
12	Naik	72	88									
13	Batal	34	40									
14	Batal	43	52									
15	Batal	37	40									
16	Batal	33	50									
17	Naik	68	87									
18	Batal	27	48									
19	Naik	69	76									
20	Batal	37	40									
21	Naik	65	88									
22	Batal	34	38									
23	Naik	66	95									
24	Naik	72	85									
25	Batal	34	31									
26	Naik	62	82									
27	Naik	64	80									
28	Batal	33	44									



Klik Options dan centang hosmer-lemeshow goodness of fit



Case Processing Summary

Unweighted Cases ^a		N	Percent
Selected Cases	Included in Analysis	100	100.0
	Missing Cases	0	.0
	Total	100	100.0
Unselected Cases		0	.0
Total		100	100.0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
Batal	0
Naik	1

Classification Table^{a,b}

Observed		Predicted		Percentage Correct	
		Batal	Naik		
Step 0	Penerbangan	Batal	62	0	100.0
		Naik	38	0	.0
Overall Percentage					62.0

a. Constant is included in the model.

b. The cut value is .500

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	.118	8	1.000

Model regresi logistic yang dibua telah memenuhi asumsi kelayakan model dimana nilai SIG 1.000 > 0.05

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	Umur	1.082	.479	5.095	1	.024	2.949
	Harga Tiket	.219	.090	5.926	1	.015	1.245
	Constant	-87.253	36.800	5.622	1	.018	.000

Berdasarkan uji wald variabel independent secara parsial memiliki pengaruh terhadap Y dimana Umur ($0.024 < 0.05$) dan Harga Tiket ($0.015 < 0.05$)

a. Variable(s) entered on step 1: Umur, Harga Tiket.

Interpretasi



- ❑ Pada table output variabel in equation dapat diketahui nilai odds ratio dari masing-masing variabel bebas
- ❑ Untuk variabel Umur (X1), nilai odds ratio-nya sebesar 2,949, hal ini menunjukkan bahwa semakin tua para penumpang, maka peluang mereka untuk tetap ikut naik pesawat lebih besar 2,949 kali dibandingkan membatalkan membatalkan pesawat
- ❑ Untuk variabel Harga Tiket (X2), nilai odds ratio sebesar 1,245. hal ini menunjukkan bahwa semakin mahal tiket, maka peluang mereka untuk tetap ikut naik pesawat lebih besar 1,245 kalo dibandingkan membatalkan pesawat

END OF SLIDE



STARTUP



STRATEGY



BUSINESS MANAGEMENT



PARTNERSHIP