

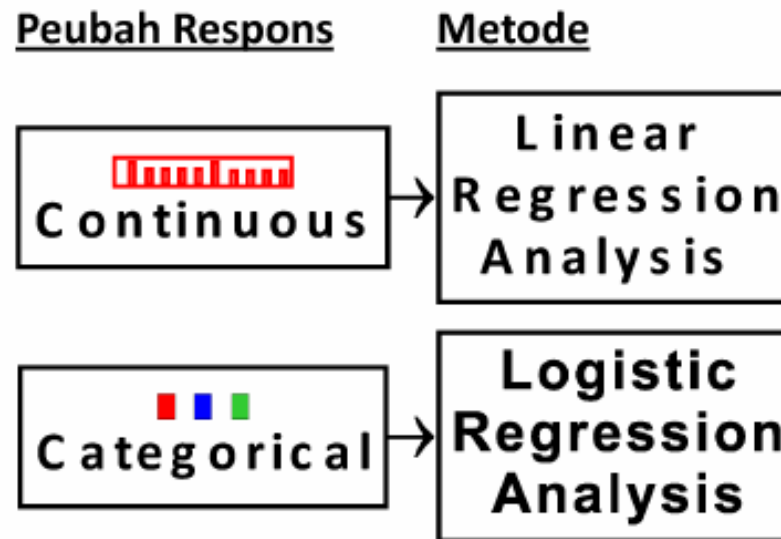
Model Regresi Logistik

Materi 11 - STK511 Analisis Statistika

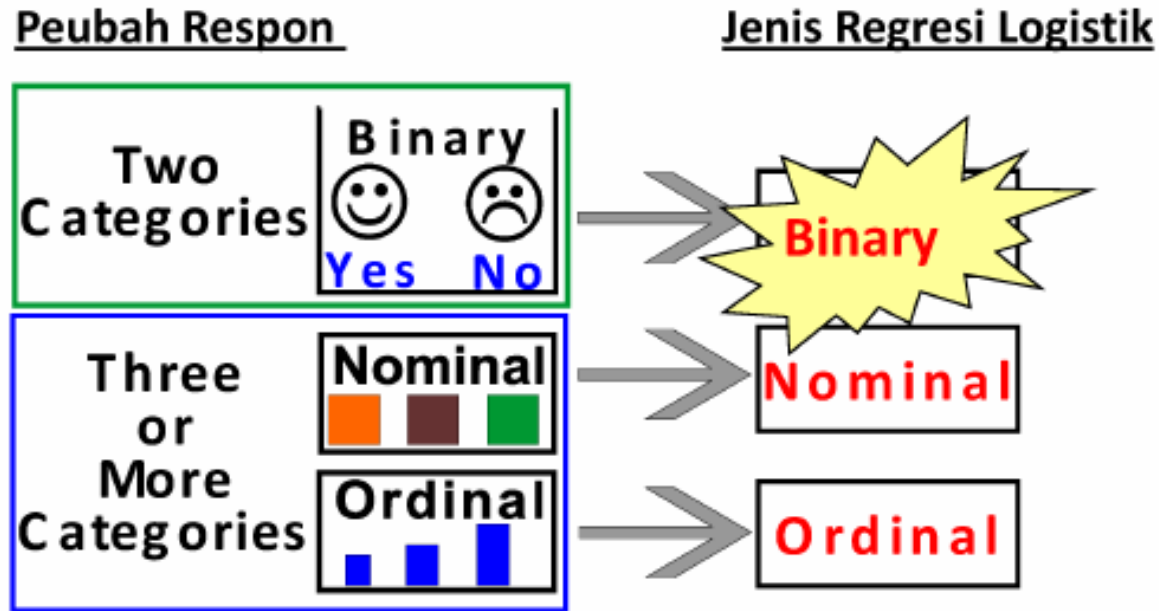
Des, 2017

Pengantar

- Model regresi logistik adalah salah satu model yang digunakan untuk mencari hubungan antara peubah respons kategori dengan satu atau lebih peubah penjelas yang kontinu ataupun kategori.



Jenis Regresi Logistik



Model Linier

Model Linear:

$$y_i \sim N(\mu_i, \sigma^2) \text{ dengan } \mu_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \dots + \beta_p x_{pi}$$

- → Model Regresi Logistik:

Transformasi fungsi peluang

$$\text{logit}(p_i) = \log\left(\frac{p_i}{1-p_i}\right)$$

Model:

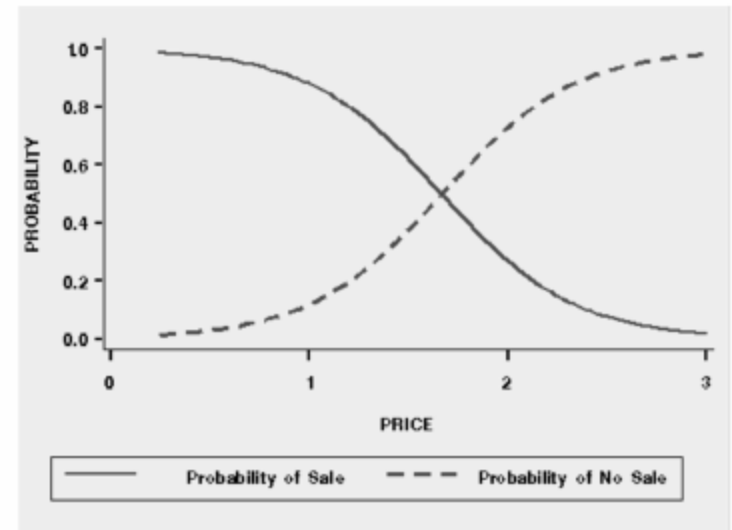
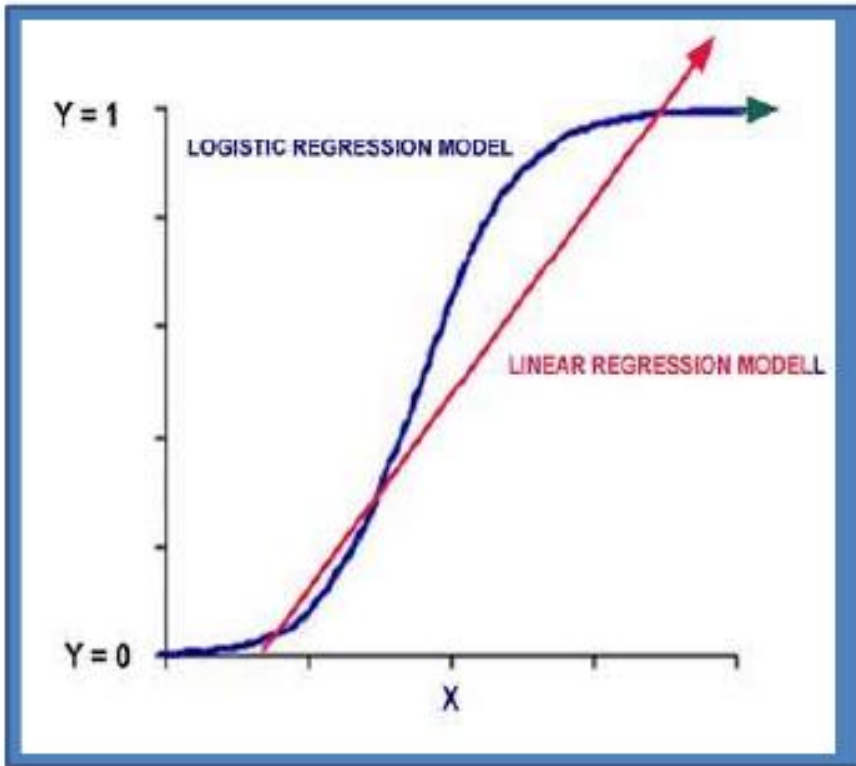
$$\text{logit}(p_i) = \beta_0 + \beta_1 x_1$$

$$P(Y = 1) = \pi = \frac{e^{\beta_0 + \beta_1 x_1}}{1 + e^{\beta_0 + \beta_1 x_1}}$$

Model regresi logistik menggunakan peubah penjelas, baik kategorik atau kontinu, untuk memprediksi peluang dari hasil yang spesifik.

Dengan kata lain, regresi logistik dirancang untuk menggambarkan peluang yang terkait dengan nilai-nilai peubah respon.

Kurva Model Regresi Logistik



Interpretasi β

- Odds akan meningkat secara multiplikatif sebesar e^β untuk setiap kenaikan 1 unit x
- $e^\beta \rightarrow$ rasio odds

$$\text{RasioOdds} = \frac{\text{odds}(X = x + 1)}{\text{odds}(X = x)}$$



Odd dan Rasio Odd

Odd (ukuran asosiasi pada regresi logistik) → rasio peluang kejadian sukses dengan kejadian tidak sukses dari peubah respon.

Adapun rasio odd mengindikasikan seberapa lebih mungkin, dalam kaitannya dengan nilai odd, munculnya kejadian sukses pada suatu kelompok dibandingkan dengan kelompok lainnya. Sebagai contoh, seberapa lebih besar peluang wanita untuk membeli produk dengan harga tertentu dibandingkan dengan pria.

Odd dan Rasio Odd

Jenis kelamin	Membeli produk		Total
	Ya	Tidak	
Pria	10	90	100
Wanita	20	60	80
Total	30	150	180

$$\text{Odd}_{\text{pria}} = \frac{P(\text{membeli})}{P(\text{tidak membeli})} = \frac{0.1}{0.9} = 0.11$$

$$\text{Odd}_{\text{wanita}} = \frac{P(\text{membeli})}{P(\text{tidak membeli})} = \frac{0.25}{0.75} = 0.33$$

Rasio odd antara pria dengan wanita adalah:

$$\text{Rasio Odd} = \frac{\text{Odd}_{\text{pria}}}{\text{Odd}_{\text{wanita}}} = \frac{0.11}{0.33} = 0.33$$

Uji Hipotesis: Simultan

Statistik uji-G adalah uji rasio kemungkinan (likelihood ratio test) yang digunakan untuk menguji peranan peubah penjelas di dalam model secara bersama-sama (Hosmer & Lemeshow, 1989).

Rumus umum uji-G untuk menguji hipotesis :

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$$

H₁ : minimal ada satu β yang tidak sama dengan 0

adalah

$$G = 2 \ln \left[\frac{\text{likelihood tanpa peubah bebas}}{\text{likelihood dengan peubah bebas}} \right]$$

Statistik G ini, secara teoritis mengikuti sebaran χ^2 dengan derajat bebas k.

Uji Hipotesis: Parsial

Sementara itu, uji Wald digunakan untuk menguji parameter β_i secara parsial. Hipotesis yang diuji adalah:

$$H_0 : \beta_i = 0$$

$$H_1 : \beta_i \neq 0$$

Formula statistik Wald adalah:

$$W = \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)}$$

Secara teori, statistik W ini mengikuti sebaran normal baku jika H_0 benar.

ilustrasi

- Suatu penelitian mengenai faktor-faktor yang mempengaruhi banyaknya satellite yang dipunyai kepiting betina (Y)
- $Y = 1$ jika kepiting betina memiliki paling tidak 1 satellite
 $Y = 0$ jika tidak memiliki satellite.
- X = lebar cangkang kepiting betina (dalam cm)

Output

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Exp(Est)
Intercept	1	-12.3508	2.6287	22.0749	<.0001	0.000
width	1	0.4972	0.1017	23.8872	<.0001	1.644

$$\hat{\pi}(x) = \frac{\exp(-12.351 + 0.497x)}{1 + \exp(-12.351 + 0.497x)}$$

At the minimum width in this sample of 21.0 cm, the estimated probability is $\exp(-12.351 + 0.497(21.0)) / [1 + \exp(-12.351 + 0.497(21.0))] = 0.129$

At the maximum width of 33.5 cm, the estimated probability equals $\exp(-12.351 + 0.497(33.5)) / [1 + \exp(-12.351 + 0.497(33.5))] = 0.987$

- lebar minimum $x = 21$ cm,

$$\begin{aligned}\hat{\pi}(x) &= \frac{\exp(-12.351 + 0.497(21))}{1 + \exp(-12.351 + 0.497(21))} \\ &= 0.129\end{aligned}$$

- lebar maksimum $x = 33.5$ cm

$$\begin{aligned}\hat{\pi}(x) &= \frac{\exp(-12.351 + 0.497(33.5))}{1 + \exp(-12.351 + 0.497(33.5))} \\ &= 0.987\end{aligned}$$

Interpretasi Output

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Exp(Est)
Intercept	1	-12.3508	2.6287	22.0749	<.0001	0.000
width	1	0.4972	0.1017	23.8872	<.0001	1.644

- Dugaan $\pi(x) = 0.5$ saat $x = -\hat{\alpha} / \hat{\beta} = 12.351 / 0.497 = 24.8$
- Dugaan odds = $\exp(\hat{\beta}) = \exp(0.497) = 1.64$
 → kepiting betina yang memiliki lebar 1 cm lebih besar, memiliki kecenderungan 1.64 kali mempunyai satelit

Selesai...