

REGRESI LOGISTIK

EKONOMETRIKA



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



Asumsi Regresi Logistik

- ❖ Regresi logistik tidak membutuhkan hubungan linier antara variabel independen dengan variabel dependen.
- ❖ Variabel independen tidak memerlukan asumsi *multivariate normality* (asumsi bersifat normal)
- ❖ Asumsi homokedastisitas tidak diperlukan
- ❖ Variabel dependen harus bersifat dikotomi (2 kategori, misal: tinggi dan rendah atau baik dan buruk)
- ❖ Variabel independen tidak harus memiliki keragaman yang sama antar kelompok variabel
- ❖ Kategori dalam variabel independen harus terpisah satu sama lain atau bersifat eksklusif
- ❖ Sampel yang diperlukan dalam jumlah relatif besar, minimum dibutuhkan hingga 50 sampel data untuk sebuah variabel prediktor (independen).
- ❖ Dapat menyeleksi hubungan karena menggunakan pendekatan non linier log transformasi untuk memprediksi odds ratio. Odd dalam regresi logistik sering dinyatakan sebagai probabilitas.



Model Persamaan Regresi Logistik

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = B_0 + B_1X$$

Ln: Logaritma Natural.

Di mana:

$B_0 + B_1X$: Persamaan yang biasa dikenal dalam OLS



Model Persamaan Regresi Logistik

P Aksen adalah probabilitas logistik yang didapat rumus sebagai berikut

$$\hat{p} = \frac{\exp(B_0 + B_1 X)}{1 + \exp(B_0 + B_1 X)} = \frac{e^{B_0 + B_1 x}}{1 + e^{B_0 + B_1 x}}$$

Probabilitas Regresi Logistik

Di mana:

exp atau ditulis “e” adalah fungsi exponen.

(Perlu diingat bahwa exponen merupakan kebalikan dari logaritma natural. Sedangkan logaritma natural adalah bentuk logaritma namun dengan nilai konstanta 2,71828182845904 atau biasa dibulatkan menjadi 2,72).



Model Persamaan Regresi Logistik

Dengan model persamaan tersebut, tentunya akan sangat sulit untuk menginterpretasikan koefisien regresinya. Oleh karena itu maka diperkenalkanlah istilah Odds Ratio atau yang biasa disingkat Exp (B) atau OR. Exp (B) merupakan exponen dari koefisien regresi. Jadi misalkan nilai slope dari regresi adalah sebesar 0,80 maka Exp(B) dapat diperkirakan sebagai berikut:

$$2,72^{0,8} = 2,23$$



Nilai Odds Ratio

Besarnya nilai Exp (B) dapat diartikan sebagai berikut:

- ❖ Misalnya nilai Exp (B) pengaruh rokok terhadap kanker paru adalah sebesar 2,23 maka disimpulkan bahwa orang yang merokok lebih beresiko untuk mengalami kanker paru dibandingkan dengan orang yang tidak merokok.
- ❖ Interpretasi ini diartikan apabila pengkodean kategori pada tiap variabel sebagai berikut:
 - Variabel bebas adalah Rokok: Kode 0 untuk tidak merokok, kode 1 untuk merokok.
 - Variabel terikat adalah kanker Paru: Kode 0 untuk tidak mengalami kanker paru, kode 1 untuk mengalami kanker paru.



Pseudo R Square

- ❖ Perbedaan lainnya yaitu pada regresi ini tidak ada nilai “R Square” untuk mengukur besarnya pengaruh simultan beberapa variabel bebas terhadap variabel terikat.
- ❖ Dalam regresi logistik dikenal istilah *Pseudo R Square*, yaitu nilai **R Square Semu** yang maksudnya sama atau identik dengan R Square pada OLS.
- ❖ Jika pada OLS menggunakan uji F Anova untuk mengukur tingkat signifikansi dan seberapa baik model persamaan yang terbentuk, maka pada regresi ini menggunakan Nilai Chi-Square. Perhitungan nilai Chi-Square ini berdasarkan perhitungan Maximum Likelihood.

