Logistic Regression

11 March 2022 12:19

$$Z_i = \ln\left(\frac{p_i}{1-p_i}\right) \quad \text{--- (1)}$$

$p \Rightarrow$ probability of success

$(1-p) \Rightarrow$ " " failure

$$\left(\frac{p}{1-p}\right) \rightarrow \cancel{Y}[p=1 / X_i]$$

$\rightarrow P[Y=1 / X_i]$ → set of independent variables.

Train Arrival:

$Y \Rightarrow$ Train Arriving / Not Arriving.

$X_1 \Rightarrow$ Reaching station on time (Y/N)

$X_2 \Rightarrow$ No delay in earlier departures (dep_time)

$X_3 \Rightarrow$ Ticket Price

→ independent variables trying to explain the PROBABILITY OF OCCURRENCE (Train arrival).

$$Z_i = \ln\left(\frac{p_i}{1-p}\right) = [\alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n] + \epsilon$$

LINK FUNCTION

$Z \Rightarrow$ LINK FUNCTION

Probabilities — Independent vars.

$$Z \Rightarrow g(y) = \alpha + \beta_1 x_1 + \beta_2 x_2 \quad (2)$$

$\ln\left(\frac{p}{1-p}\right) \Rightarrow$ direct relationship of logarithmic transformations with exponential values.

LOG-ODDS $\rightarrow$ chance of occurrence of an event.

$$Z = \left(\frac{p}{1-p}\right) = \alpha + \beta_1 x_1 + \beta_2 x_2$$

$$\ln\left(\frac{p}{1-p}\right) = \ln(\alpha + \beta_1 x + \beta_2 x_2)$$

LINEAR REGRESSION:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \epsilon \quad (1)$$

$\ln\left(\frac{p}{1-p}\right) \Rightarrow$ dependent variable

... predict the

$\hookrightarrow \ln(1-p)$

I need to predict the probability.

$$\Rightarrow \ln\left(\frac{p}{1-p}\right) = y$$

Take exponents on both the sides,

$$\Rightarrow \left(\frac{p}{1-p}\right) = e^{y}$$

$$\Rightarrow \left(\frac{p}{1-p}\right) = e^{(\alpha + \beta_1 x_1 + \beta_2 x_2 + \epsilon)} \quad \text{--- (2)}$$

$$Z_i = \ln\left(\frac{p_i}{1-p_i}\right) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \epsilon.$$

Taking ~~log~~ exponential on both sides. —

$$\exp\left[\ln\left(\frac{p_i}{1-p_i}\right)\right] = \exp(\alpha + \beta_1 x_1 + \beta_2 x_2 + \epsilon)$$

$$\Rightarrow \left(\frac{p_i}{1-p_i}\right) = e^{(\alpha + \beta_1 x_1 + \epsilon)}$$

$$\Rightarrow p_i = (1-p_i)\left[e^{(\alpha + \beta_1 x_1 + \epsilon)}\right]$$

$$\Rightarrow p_i = \frac{\exp[\alpha + \beta_1 x_1 + \epsilon]}{1 + \exp[\alpha + \beta_1 x_1 + \epsilon]}$$

$\Rightarrow \quad p_i = \overline{\qquad\qquad}$

Dividing by $e^{(\alpha+\beta_1 x_1+\epsilon)}$ — $1 + \exp\left[(\alpha+\beta_1 x_1+\epsilon)\right]$

$$\Rightarrow \quad \boxed{p_i = \frac{1}{1+\exp^{-(\alpha+\beta_1 x_1+\epsilon)}}}$$

SIGMOID CURVE (S-curve)