Logistic Regression

Reporter : Aiqing Jiang

What is logistic regression?

- In statistics, logistic regression, or logit regression, or logit model is a regression model where the dependent variable (DV) is categorical.

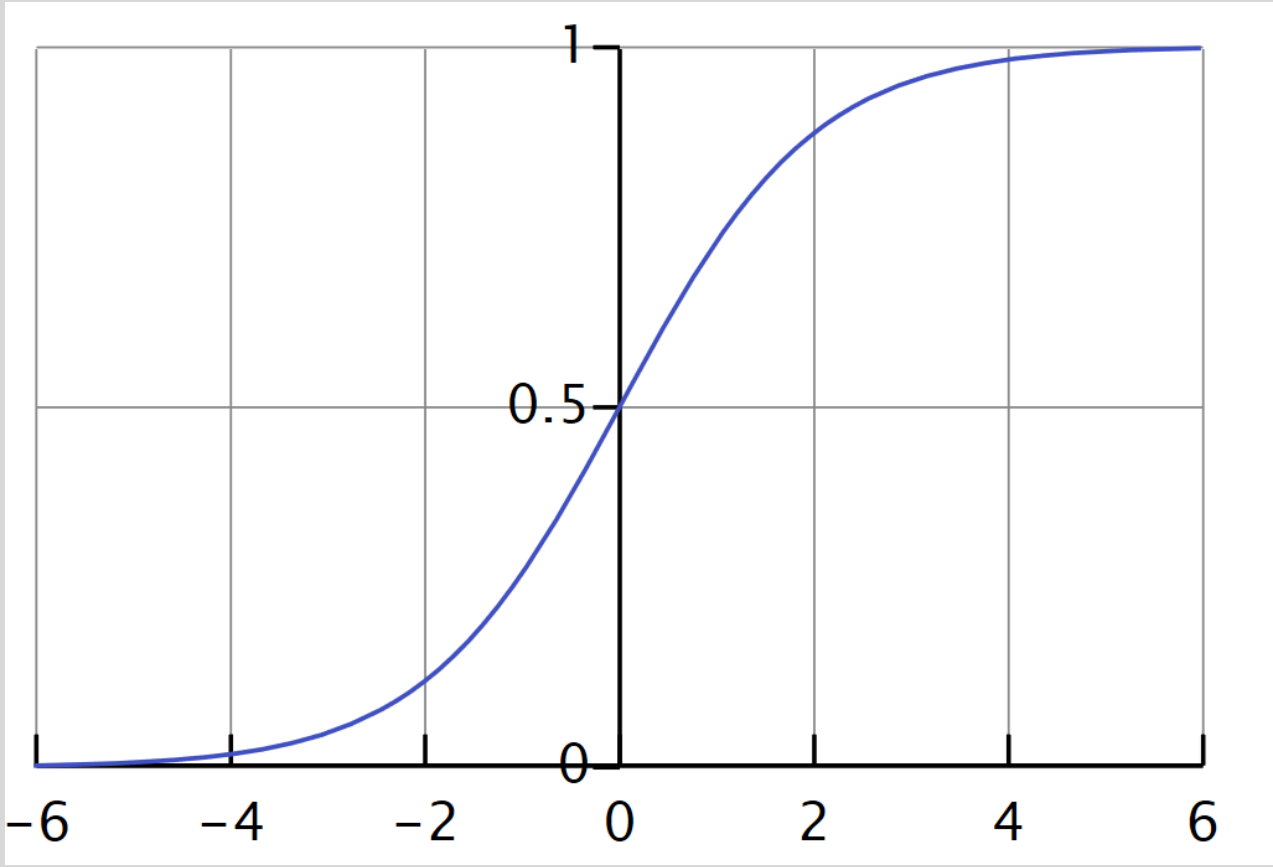
Which means the dependent variable can be binary—where the output can take only two values, "0" and "1", represents outcomes such as pass/fail, win/lose, alive/dead or healthy/sick. Cases where the dependent variable has more than two outcome categories may be analysis in multinomial logistic regression, or, if the multiple categories are ordered, in ordinal logistic regression.

- The logistic function is useful because it can take any real input t whereas the output always takes values between zero and one and hence is interpretable as a probability.

Definition of Logistic Regression

- The logistic function $\sigma(t)$ is defined as follows:

$$\sigma(t) = \frac{e^t}{e^t+1} = \frac{1}{1+e^{-t}}$$



Logistic Function on the *t*-interval (-6,6)

Definition of Logistic Regression

- Let us assume that t is a linear function of a single explanatory variable x, which means $t = \beta_0 + \beta_1 x$

- And the logistic function can now be written as:

$$F(x) = \frac{1}{1+e^{-(\beta_0+\beta_1 x)}}$$

Note that F (x)is interpreted as the probability of the dependent variable equaling a "success" or "case" rather than a failure or non-case.

Applications

- Logistic regression is used in various fields, including machine learning, most medical fields, and social sciences.
- Logistic regression may be used to predict whether a patient has a given disease (e.g. diabetes; coronary heart disease), based on observed characteristics of the patient (age, sex, body mass index, results of various blood tests, etc.).
- Another example might be to predict whether an American voter will vote Democratic or Republican, based on age, income, sex, race, state of residence, votes in previous elections, etc.
- It is also used in marketing applications such as prediction of a customer's propensity to purchase a product or halt a subscription, etc. In economics it can be used to predict the likelihood of a person's choosing to be in the labor force, and a business application would be to predict the likelihood of a homeowner defaulting on a mortgage.