Chapter 45
Logodds, the Basis of Logistic Regression

1 General Purpose

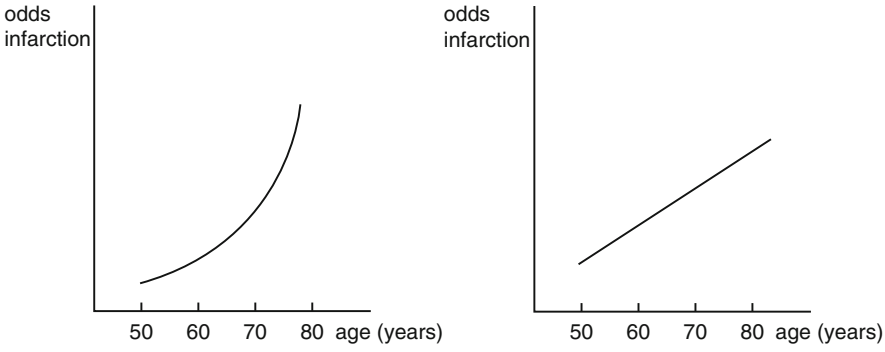
Logistic regression is much similar to linear regression (see Chap. 8). The difference is the type of outcome variable, which is continuous with linear regression and binary with logistic regression. In order for logistic regression to work, we need to transform the binary outcome into the odds of responding, or rather the logodds of responding. In a population

$$\text{the odds of an infarction} = \frac{\text{the number of patients with infarct}}{\text{number of patients without.}}$$

The easiest way to understand the term odds is to think of it as though it is the risk.

The odds or risk of an infarction is correlated with age: the older, the larger the odds.

Now how does it correlate with age? As shown underneath it is not at all linear.



However, if we transform the underneath linear model

T.J. Cleophas, A.H. Zwinderman, *Clinical Data Analysis on a Pocket Calculator*,
DOI 10.1007/978-3-319-27104-0_45

$$y = a + bx$$

into a loglinear model

$$\ln \text{odds} = a + b \times (x = \text{age}),$$

then, all of a sudden, we will observe a close to linear relationship (the above right graph). This present from heaven can be used for statistical testing. The current chapter assesses how ln odds, often called logodds, can be used for testing studies with binary outcome data, like numbers of responders to treatment yes or no.

2 Primary Scientific Question

How can we use logodds for modeling binary outcome data for the purpose of making predictions from them.

3 Data Example

In a random population of subjects between 50 and 75, we register 5 year age classes for each subject and follow them for 3 years. After that the numbers of infarctions are counted per age class.

ageclass	population size	infarctions	odds of infarct	logodds
50–55	977	8	8/969	−4.80
55–60	1010	42	42/968	−3.14
60–65	999	67	67/932	−2.63
65–70	990	112	112/878	−2.06
70–75	1061	190	190/871	−1.52

For the pocket calculation of logodds, first, compute the above odds, and, then, press the ln button. We will also use the pocket calculator to compute the intercept and regression coefficient of the linear regression between the age class (x-variable) and the logodds of infarction (y-variable). This linear regression is otherwise called logistic regression.

We will use the Electronic Calculator (see Chap. 1) for computations. First, we will calculate the b and r values.

Command:

click ON....click MODE....press 3....press 1....press SHIFT, MODE, and again 1.... press = ...start entering the data.... [55, −4.80]....[60, −3.14]....[65, −2.63]. . . [70. −2.06]. . .[75, −1.52]

In order to obtain the a value, press: shift, S-VAR, ▶, ▶, 1, = .
In order to obtain the b value, press: shift, S-VAR, ▶, ▶, 2, = .

The underneath values are obtained.

a = −12.8
b = 0.15

logodds = a + b (age class)

Now, we can use the above equation for making predictions about the risk of having an infarction in the upcoming 3 years in future subjects.

For someone 50–55 years	the logodds of infarction	= −12.8 + 0.15 (55)
		= −4.55
	the odds of infarction	= 0.0106 = 1.06 %
For someone 70–75 years	the logodds of infarction	= −12.8 + 0.15 (75)
		= −1.55
	the odds of infarction	= 0.212 = 21.2 %

The risk is pretty much similar to the odds, particularly, with small risks, and is equal to risk = 1/(1 + 1/odds)

e.g., with	odds	= 21.2 %
	risk	= 1/(1 + 1/0.212)
		= 0.175 = 17.5 %.

4 Conclusion

Logistic regression is not in pocket calculators, but can be used even so if you transform your outcome odds values into logodds values. We should add that logistic regression is a magnificent methodology with plenty applications, most of whom require the use of advanced statistical software. For example it can be used not only for predictive models of age and infarction, but also for predictive models with multiple predictors like risk factors and patient characteristics for risk management assessments. It also can be used for efficacy analysis of survival studies and exploratory purposes. More information is given in the Chaps. 17, 19, 21, 49, and 65, in Statistics applied to clinical studies 5th edition, Springer Heidelberg Germany, 2012, from the same authors.

5 Note

More background, theoretical and mathematical information of logistic regression is given in SPSS for starters and 2nd levelers 2nd edition, Chap. 36, 37, 38, and 39, Springer Heidelberg Germany, 2015, from the same authors.