ML

Runi Malladi

June 26, 2023

1 appendix

1.1 logistic regression

The purpose of logistic regression is to take data associating various values of the independent variables to binary outcomes and produce a model which takes values of the independent variables and returns a probability of a binary outcome occuring.

Background

Consider p to be the probability of an event occurring. We can further assume only two outcomes: either the event occurs, or it doesn't. We define the *odds ratio* to be

$$\begin{aligned} \text{odds ratio} : [0,1) &\to [0,\infty) \\ p &\mapsto \frac{p}{1-p}. \end{aligned}$$

We define the *log-odds ratio*, or *logit*, to be

$$\begin{aligned} \text{logit} : (0,1) &\to (-\infty,\infty) \\ p &\mapsto \log\left(\frac{p}{1-p}\right). \end{aligned}$$

The graphs of these functions are depicted below (Figure 1):

Assumptions

The fundamental assumption of logistic regression is a linear relationship between the independent variables and the log-odds.

For instance, consider a situation with two independent variables X_1, X_2 which determine a binary outcome (either 0 or 1). We assume

- it is reasonable to model the probability of an input (x_1, x_2) resulting in the binary outcome 1. That is, each outcome y_i is Bernoulli distributed.
- this relationship is linear: letting p denote the probability of (x_1, x_2) producing 1, we have
$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2.$$
for some $\beta_0, \beta_1, \beta_2 \in \mathbb{R}$. Note that the β_i do not depend on the x_i.

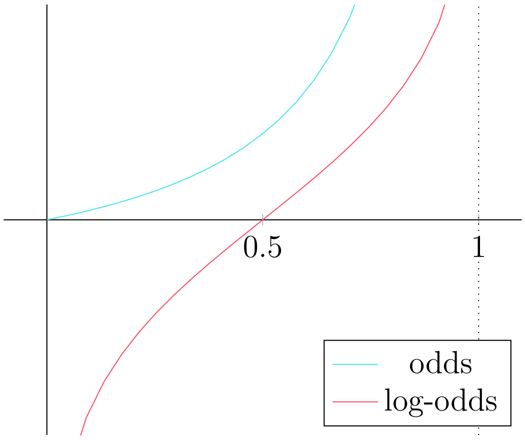

Figure 1: Graphs of odds and log-odds functions.

Objective

Assuming a linear relationship between log-odds and the independent variables

$$\operatorname{logit}(p(x)) = \beta \cdot \begin{pmatrix} 1 \\ x \end{pmatrix} = \beta_0 + \beta_1 x_1 + \cdots \beta_n x_n,$$

the objective of logistic regression is to determine (or approximate) the coefficients β in the above linear combination. As a matter of convention, by $\beta \cdot x$ or $\beta^T x$ we will mean the above dot product, where we have added an $x_0 = 1$ term to the original x.

As we will demonstrate, once the coefficients β have been determined, we can determine the probability $p(x)$ of input x succeeding using the following formula:

$$p(x) = \frac{1}{1 + e^{-\beta^T x}}.$$