Generalized linear models

Kristian Wichmann

October 20, 2017

A *generalized linear model* is an extension of the general linear model framework. It introduces a non-linear component to the mean function, and allows the distribution of the response variable to be non-Gaussian.

1 Components

A generalized linear model consists of three parts:

- A stochastic component.
- A systematic component.
- A link function.

1.1 Example: Logistic regression

In simple logistic regression, we try to model the probability of a Bernoulli variable as a function of the explanatory variable input x, such that:

$$p(x) = \sigma(\beta_1 x + \beta_0) \tag{1.1}$$

Here σ is the *logistic function*:

$$\sigma(z) = \frac{1}{1 + e^{-z}} \tag{1.2}$$

Figure 1 graphs the function.

There's a few important things to note here. First, the parameter p for a Bernoulli distribution is equal to the expectation value. It turns out that it's really the expectation value we wish to model, more generally. Second, the argument of σ is linear in the explanatory variable x.

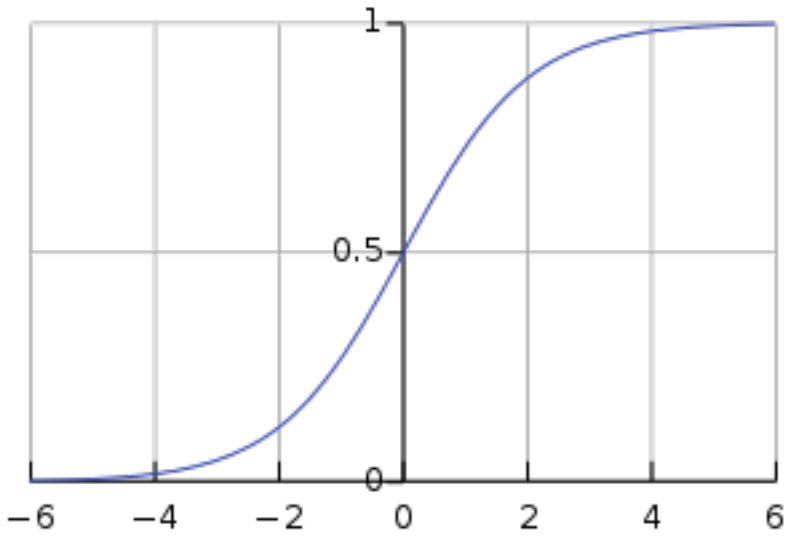

Figure 1: The logistic function σ.

1.2 Stochastic component

The stochastic component is a distribution from the exponential family. The values which the response variable take on for a given