Generalized linear models

Kristian Wichmann

October 21, 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 response variable y 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.

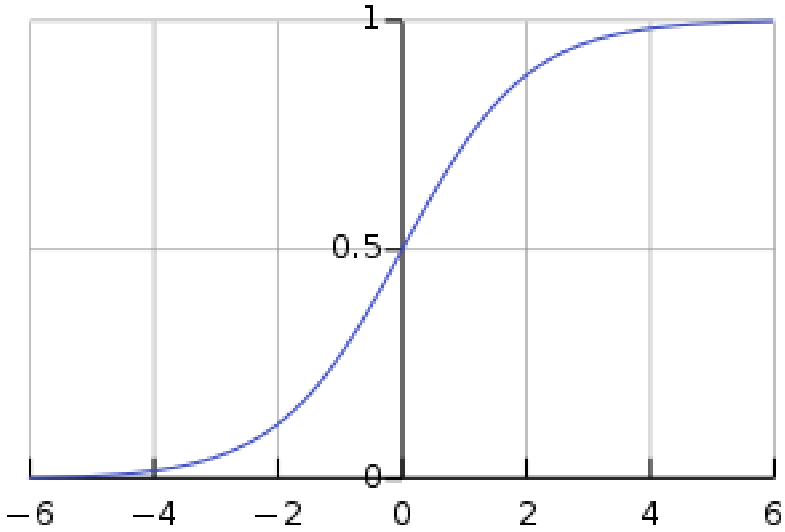

Figure 1: The logistic function σ.

1.2 Stochastic component

The stochastic component in a generalized linear model is a distribution from the exponential family. The response variable is distributed according to this, with a mean value that is given by the model.

In the logistic regression example above, the stochastic component is the Bernoulli distribution.

1.3 Systematic component

The systematic component is the 'linear' part of of the model. I.e. a linear function of the explanatory variables: $X\beta$. Here β is a vector of coefficients to be modelled.

In the logistic regression example above, the systematic component is the expression $\beta_1 x + \beta_0$. It is assumed that the constant 1 is an explanatory variable to fit the constant term β_0 as usual.

1.4 Link function

Finally, the link function g is the inverse of the (generally non-linear) function that is applied to the systematic component to get the predicted mean value:

$$\mathbb{E}[y] = g^{-1}(X\beta) \tag{1.3}$$

This of course means, that the function has to be invertible.

In the logistic regression example above, the inverse of the link function is the logistic function:

$$g^{-1}(z) = \sigma(z) \tag{1.4}$$

The link function itself is the inverse, which is known as the *logit* function:

$$g(p) = \ln\left(\frac{p}{1-p}\right) \tag{1.5}$$