

Generalized linear models (GLMs) were a great advance in statistical modeling. The original manuscript with the GLM framework was from [Nelder and Wedderburn in 1972](#). in the Journal of the Royal Statistical Society. The McCullagh and Nelder book¹ is the famous standard treatise on the subject.

Recall linear models. Linear models are the most useful applied statistical technique. However, they are not without their limitations. Additive response models don't make much sense if the response is discrete, or strictly positive. Additive error models often don't make sense, for example, if the outcome has to be positive. Transformations, such as taking a cube root of a count outcome, are often hard to interpret. In addition, there's value in modeling the data on the scale that it was collected. Particularly interpretable transformations, natural logarithms in specific, aren't applicable for negative or zero values.

The generalized linear model is family of models that includes linear models. By extending the family, it handles many of the issues with linear models, but at the expense of some complexity and loss of some of the mathematical tidiness. A GLM involves three components

- An *exponential family* model for the response.
- A systematic component via a linear predictor.
- A link function that connects the means of the response to the linear predictor.

The three most famous cases of GLMs are: linear models, binomial and binary regression and Poisson regression. We'll go through the GLM model specification and likelihood for all three. For linear models, we've developed them previously. The next two modules will be devoted to binomial and Poisson regression. We'll only focus on the most popular and useful link functions.