Notes on Generalized Linear Models

compiled by D.Gueorguiev 12/1/2024

# Introductory Notes

## The Problem of Looking at Data

Suppose we have a number of measurements or counts, together with some associated structural or contextual information, such as the order in which the data were collected, which measuring instruments were used, and other differences in the conditions under which the individual measurements were made. To interpret such data, we search for a pattern, for example that one measuring instrument has produced consistently higher readings than another.

Such systematic effects are likely to be blurred or overwhelmed by other variation of a more haphazard nature. The latter variation is usually described in statistical terms, no attempt being made to model or to predict the actual haphazard contribution to each observation.

Statistical models contain both of those elements which are *systematic effects* and *random effects* accordingly. The problem of looking and understanding data demands the formulation of patterns that are thought capable of describing succinctly not only the systematic variation in the data under study, but also for describing patterns in similar data.

## Theory as Pattern

We shall consider theories as generating patterns of numbers, which in some sense can replace the data, and can themselves be described in terms of a small number of quantities. These quantities are called *parameters*. By giving the parameters different values, specific patterns can be generated.

# References

[1] [Generalized Linear Models, P. McCullagh, J.A. Nelder, 2nd Edition, 1983](https://github.com/dimitarpg13/probabilistic_machine_learning/blob/main/applied_statistics/books/GeneralizedLinearModels/GeneralizedLinearModels_McCullagh_Nelder_1983.pdf)

[2] [Generalized Linear Model, Wikipedia](https://en.wikipedia.org/wiki/Generalized_linear_model)