CHAPTER 1

Link Functions

In generalized linear models, the link function provides the relationship between the linear predictor and the mean of the distribution function. Different choices of link function can be used to model different types of relationships. Here are a few commonly used link functions:

1. **Identity link:** The identity link function is the simplest form of link function, where the response variable is expected to be the linear combination of the predictors. This is the default link function for Gaussian family distributions.

$$g(\mu) = \mu$$

2. **Log link:** The log link function is used when modeling positive data and count data. This link function is the default for Poisson and exponential family distributions.

$$g(\mu) = \log(\mu)$$

3. **Logit link:** The logit link function is often used when modeling binary response data, and is the default link function for binomial family distributions. It gives the log-odds, or the logarithm of the odds $p/(1-p)$.

$$g(\mu) = \log\left(\frac{\mu}{1-\mu}\right)$$

4. **Probit link:** The probit link function is another common choice for binary response data. It is based on the cumulative distribution function of the standard normal distribution.

$$g(\mu) = \Phi^{-1}(\mu)$$

where $\Phi^{-1}(\cdot)$ is the inverse cumulative distribution function of the standard normal distribution.

5. **Inverse link:** The inverse link function is often used in modeling rates or times. It is the canonical (or default) link function for the Gamma family distributions.

$$g(\mu) = \mu^{-1}$$

Different link functions can substantially impact the model's interpretation, so it's crucial to choose a link function that aligns with the nature of the data and the scientific question at hand.

Appendix A

Answers to Exercises

Index