Chapter 3

Generalized linear models (GLMs)

3.1 Introduction

I begin this chapter with the re-analysis of a small example from Finney (1978). Table 3.1 gives the data for the number of plates (out of five) on which growth of *Bacillus mesentericus* was successful in different dilutions of a potato flour suspension. Figure 3.1 shows a plot of the data. Several features of the experiment and the plot are worth noting. First, the way the experiment was conducted (positive/negative response for each of five independent trials) guarantees that the distribution of the response is binomial. Second, the form of the response is nonlinear, exhibiting somewhat of an S-shape. Third, extension of the range of the predictor in the positive direction would likely lead to more responses of 1.0 and extension in the negative direction would lead to additional responses of 0.0.

To account for these features, we hypothesize the following model:

$$
\begin{aligned}
Y_i &\sim \text{indep. binomial}[5, p(x_i)], \\
\log[p(x_i)/\{1-p(x_i)\}] &= \alpha + \beta x_i, \text{ or equivalently,} \\
p(x_i) &= 1/(1+\exp\{-[\alpha+\beta x_i]\}).
\end{aligned}
\tag{3.1}
$$

Thus, p_i is modeled as an S-shaped function of the *linear predictor*, $\alpha + \beta x_i$. The log likelihood for this model is easily specified and is proportional to

$$
\log L = \sum_i y_i(\alpha + \beta x_i) - \log[1 + \exp(\alpha + \beta x_i)]. \tag{3.2}
$$

This can be easily maximized numerically as a function of α and β to find the maximum likelihood estimates of $\hat{\alpha} = 4.17$ and $\hat{\beta} = 1.62$. Figure 3.1 shows the fit of this model to the data.