Regression Methods: Problems MATH-408
Anthony Davison

Problem 1 Independent data $y_1, \dots, y_n$ are assumed to follow a binary logistic model in which y_j takes value 1 with probability $\pi_j = \exp(x_j^{\mathrm{T}}\beta)/\{1 + \exp(x_j^{\mathrm{T}}\beta)\}$ and value 0 otherwise.

(a) Show that the deviance for a model with fitted probabilities $\widehat{\pi}_j$ can be written as

$$D = -2\left\{ y^{\mathrm{T}} X \widehat{\beta} + \sum_{j=1}^{n} \log(1 - \widehat{\pi}_j) \right\}$$

and that the likelihood equation is $X^{\mathrm{T}} y = X^{\mathrm{T}} \widehat{\pi}$. Deduce that the deviance is a function of the $\widehat{\pi}_j$ alone. Comment on the implications for using D to measure goodness of fit.

(b) If $\pi_1 = \cdots = \pi_n$, show that Pearson's statistic equals n. Comment.

Problem 2 Data $y_1, \dots, y_n$ are a realisation of independent Bernoulli variables with

$$\mathrm{P}(Y_j = 1) = 1 - \mathrm{P}(Y_j = 0) = \frac{1}{1 + \exp(-x_j^{\mathrm{T}}\beta)}, \quad j = 1, \dots, n,$$

where $x_1, \dots, x_n$ are known $p \times 1$ vectors of constants and $\beta \in \mathbb{R}^p$ is to be estimated.

(a) Find the log likelihood $\ell(\beta)$ and show that it is never positive.

(b) If there exists a vector γ such that $x_j^{\mathrm{T}}\gamma > 0$ when $y_j = 1$ and $x_j^{\mathrm{T}}\gamma < 0$ when $y_j = 0$, then show by considering $\ell(t\gamma)$, where t is scalar, that the maximum likelihood estimate has at least one component that equals $\pm\infty$. What is then the value of the log likelihood?

(c) The panels below show binary responses $y = 1/0$ (solid black/circle) for data with $n = 50$ and $p = 2$. For each panel say whether you expect to have difficulties with likelihood estimation, and explain what you would expect when fitting a model.