Discrete Response Model
Lecture 2

datascience@berkeley

Parameter Estimation

Likelihood Function

Maximum likelihood estimation is used to estimate the parameters of the model. The likelihood function is

$$L(\beta_0, \ldots, \beta_p) = \prod_{i=1}^{n} \pi_i^{y_i}(1-\pi_i)^{(1-y_i)}$$

where

$$\begin{aligned}\pi_i &= P(y=1|\mathbf{x}) \\ &= \frac{exp(\beta_0 + \mathbf{x}\beta)}{1+exp(\beta_0+\mathbf{x}\beta)} \\ &= \frac{exp(\beta_0+\beta_1 x_1 + \cdots + \beta_k x_p)}{1+exp(\beta_0+\beta_1 x_1 + \cdots + \beta_p x_p)}\end{aligned}$$

and the corresponding log-likelihood function is

$$\begin{aligned}log(L(\beta_0, \ldots, \beta_p|y_1, \ldots, y_n)) &= log\left(\prod_{i=1}^{n} \pi_i^{y_i}(1-\pi_i)^{(1-y_i)}\right) \\ &= \sum_{i=1}^{n} y_i log(\pi_i) + (1-y_i) log(1-\pi_i)\end{aligned}$$

Deriving the MLE and the glm( ) Function in R

Taking derivatives with respect to $\beta_0, \ldots, \beta_p$, setting them equal to 0, and then solving for the parameters lead to the MLEs. These parameter estimates are denoted by $\hat{\beta}_0, \ldots, \hat{\beta}_p$. Corresponding estimates of are

$$\hat{\pi} = \frac{exp(\hat{\beta}_0 + \hat{\beta}_1 x_1 + \cdots + \hat{\beta}_k x_p)}{1 + exp(\hat{\beta}_0 + \hat{\beta}_1 x_1 + \cdots + \hat{\beta}_p x_p)}$$

Unfortunately, there are no closed form expressions that can be written out for $\hat{\beta}_0, \ldots, \hat{\beta}_p$. The MLEs are estimated using numerical procedures.

The Newton-Raphson procedure can be used for finding the MLE of π in a homogeneous population setting. (See the text for an example.)

In this course, we will use extensively the R function, glm(). Though we will explain the utility of this function, you are expected to read the documentation of this function, which can be shown in R by typing "?glm".

Iterated Reweighted Least Squares

glm() in R uses iteratively reweighted least squares as the default numerical method, which is shown in an excerpt of R documentation below.

`method`

the method to be used in fitting the model. The default method `"glm.fit"` uses iteratively reweighted least squares (IWLS): the alternative `"model.frame"` returns the model frame and does no fitting. User-supplied fitting functions can be supplied either as a function or a character string naming a function, with a function which takes the same arguments as `glm.fit`. If specified as a character string it is looked up from within the stats namespace.

This is an iterative process, which is typical for many numerical methods, and continues until convergence is found or a prior-specified maximum number of iterations is reached.

The function for Iterated Reweighted LS is documented, among many places, in the following website: http://www.inside-r.org/packages/cran/Rfit/docs/irls.

Berkeley SCHOOL OF INFORMATION