Generalized Linear Models – Part V

Johns Hopkins Engineering

625.461 Statistical Models and Regression

Module 12 – Lecture 12B

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

Testing hypotheses on subsets of parameters using deviance

$$\eta = \mathbf{X}\beta = \mathbf{X}_1\beta_1 + \mathbf{X}_2\beta_2$$

where the **full model** contains p parameters, $\boldsymbol{\beta}_1$ contains $p - r$ parameters, $\boldsymbol{\beta}_2$ contains r parameters.

Testing hypotheses on subsets of parameters using deviance

Wish to test $H_0 : \boldsymbol{\beta}_2 = \mathbf{0}, \quad H_1 : \boldsymbol{\beta}_2 \neq \mathbf{0}$

The **reduced model** is $\boldsymbol{\eta} = \mathbf{X}_1 \boldsymbol{\beta}_1$

The deviance of the reduced model is $D(\boldsymbol{\beta}_1)$. The difference in the deviance is

$$D(\boldsymbol{\beta}_2 \mid \boldsymbol{\beta}_1) = D(\boldsymbol{\beta}_1) - D(\boldsymbol{\beta})$$

which is chi-square distributed with df $= n - (p - r) - (n - p) = r$

Testing hypotheses on subsets of parameters using deviance

$$\text{if } D(\boldsymbol{\beta}_2 \mid \boldsymbol{\beta}_1) \geq \chi^2_{\alpha,r} \quad \text{reject the null hypothesis}$$

$$\text{if } D(\boldsymbol{\beta}_2 \mid \boldsymbol{\beta}_1) < \chi^2_{\alpha,r} \quad \text{do not reject the null hypothesis}$$

Pneumoconiosis Data (Ex 13.3, page 434)

Based on the data of Table 13.1, we get

$$\hat{y} = \hat{\pi} = \frac{1}{1 + e^{+4.7965 - 0.0935x}}$$

Pneumoconiosis Data (Ex 13.3, page 434)

Suppose that we wish to determine whether adding a quadratic term improves the model fit. So the full model considered is

$$y = \frac{1}{1+e^{-(\beta_0+\beta_1 x+\beta_{11}x^2)}}$$

$$\begin{aligned}\eta &= \mathbf{X}\boldsymbol{\beta} \\ &= \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 \\ &= \beta_0 + \beta_1 x + \beta_{11}x^2\end{aligned}$$

Pneumoconiosis Data (Ex 13.3, page 434)

$$D(\boldsymbol{\beta}) = 3.28164 \qquad D(\boldsymbol{\beta}_1) = 6.05077$$

$$\begin{aligned} D(\boldsymbol{\beta}_2 \mid \boldsymbol{\beta}_1) &= D(\boldsymbol{\beta}_1) - D(\boldsymbol{\beta}) \\ &= 6.05077 - 3.28164 \\ &= 2.76913 \end{aligned}$$

df = 1 and p-value = 0.0961

$\Rightarrow$ some marginal value in including quadratic term, i.e., (yrs of exposure)^2

JOHNS HOPKINS

WHITING SCHOOL
of ENGINEERING