Discrete Response Model
Lecture 3

datascience@berkeley

Quadratic Term:
An Introduction

Odds ratios involving polynomial terms are dependent on the explanatory variable of interest. For example, to find OR for x_1 in

$$\text{logit}(\pi) = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2$$

The corresponding odds ratio is

$$OR = \frac{Odds_{x_1+c}}{Odds_{x_1}} = \frac{e^{\beta_0+\beta_1(x_1+c)+\beta_2(x_1+c)^2}}{e^{\beta_0+\beta_1 x_1+\beta_2 x_1^2}} = e^{c\beta_1+2cx_1\beta_2+c^2\beta_2} = e^{c\beta_1+c\beta_2(2x_1+c)}$$

The standard interpretation becomes

The odds of a success change by $e^{c\beta_1+c\beta_2(2x_1+c)}$ times for a c-unit increase in x_1 when x_1 is at a value of ___.

Because the odds ratio is dependent on the explanatory variable value, it is better to change the interpretation to

The odds of a success are $e^{c\beta_1+c\beta_2(2x_1+c)}$ times as large for x_1 = __ + c than for x_1 = __,

where you need to put in the appropriate value of x_1. Also, this means multiple odds ratios may be needed to fully understand the effect of x_1 on the response.

Wald Confidence Interval

Wald confidence intervals are found in a similar manner as for interaction terms.

For the model of $\text{logit}(\pi) = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2$, the interval is

$$e^{c\hat{\beta}_1 + c\hat{\beta}_2(2x_1+c) \pm cZ_{1-\alpha/2}\sqrt{\text{Var}(\hat{\beta}_1+\hat{\beta}_2(2x_1+c))}}$$

where

$$\text{Var}(\hat{\beta}_1 + \hat{\beta}_2(2x_1 + c)) = \text{Var}(\hat{\beta}_1) + (2x_1 + c)^2\,\text{Var}(\hat{\beta}_2) + 2(2x_1 + c)\text{Cov}(\hat{\beta}_1, \hat{\beta}_2)$$

Profile likelihood ratio intervals can be calculated as well, but they are subject to the same problems as before with the mcprofile package.

Berkeley SCHOOL OF INFORMATION