Lecture 16

Logistic regression

(Ch.16 in the supplementary chapters)

The model

LOGISTIC REGRESSION MODEL

The **statistical model for logistic regression** is

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x$$

where p is a binomial proportion and x is the explanatory variable. The parameters of the logistic model are β_0 and β_1.

Definition, pg 16-5
Introduction to the Practice of Statistics, Fifth Edition

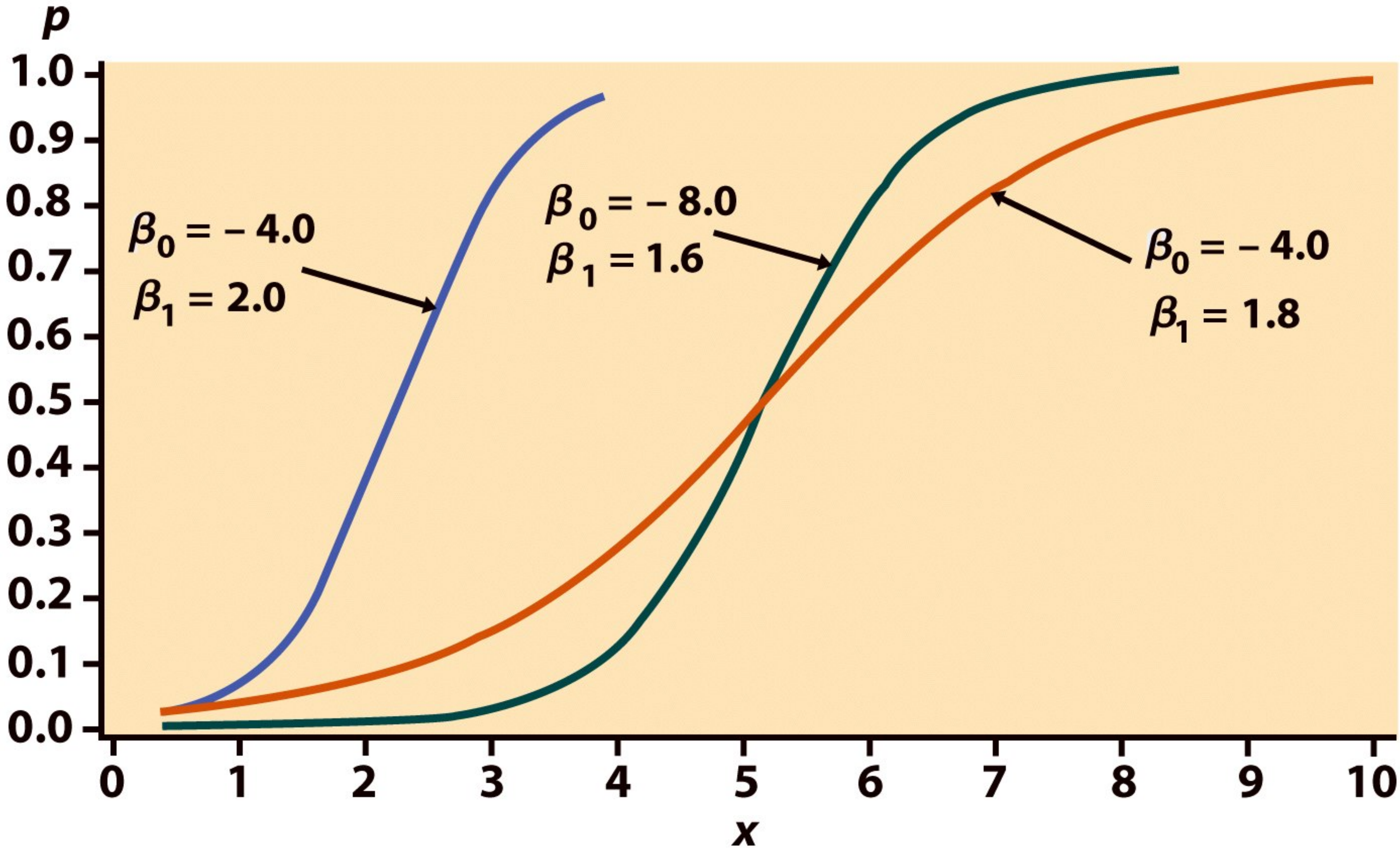

Figure 16-1
Introduction to the Practice of Statistics, Fifth Edition

CONFIDENCE INTERVALS AND SIGNIFICANCE TESTS FOR LOGISTIC REGRESSION PARAMETERS

A **level C confidence interval for the slope β_1** is

$$b_1 \pm z^* \mathrm{SE}_{b_1}$$

The ratio of the odds for a value of the explanatory variable equal to $x + 1$ to the odds for a value of the explanatory variable equal to x is the **odds ratio.**

A **level C confidence interval for the odds ratio e^{β_1}** is obtained by transforming the confidence interval for the slope

$$(e^{b_1 - z^* \mathrm{SE}_{b_1}},\ e^{b_1 + z^* \mathrm{SE}_{b_1}})$$

In these expressions z^* is the value for the standard normal density curve with area C between $-z^*$ and z^*.

To test the hypothesis H_0: $\beta_1 = 0$, compute the **test statistic**

$$z = \frac{b_1}{\mathrm{SE}_{b_1}}$$

The P-value for the significance test of H_0 against H_a: $\beta_1 \neq 0$ is computed using the fact that when the null hypothesis is true, z has approximately a standard normal distribution.

Definition, pg 16-08

Introduction to the Practice of Statistics, Fifth Edition

Example

- We use the example CSDATA from your textbook on page D-2
- We run through the exercises in chapter 16.
- Please see the accompanying R code.