

6. Suppose we collect data for a group of students in a statistics class with variables  $X_1$  = hours studied,  $X_2$  = undergrad GPA, and  $Y$  = receive an A. We fit a logistic regression and produce estimated coefficient,  $\hat{\beta}_0 = -6, \hat{\beta}_1 = 0.05, \hat{\beta}_2 = 1$ .

(a) Estimate the probability that a student who studies for 40 h and has an undergrad GPA of 3.5 gets an A in the class.

Model would be like :

$$P(Y = 1|X_1, X_2) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2)}}$$

- $\beta_0 = -6$
- $\beta_1 = 0.05$  (for  $X_1$ , hours studied)
- $\beta_2 = 1$  (for  $X_2$ , undergrad GPA)

$X_1 = 40$  and  $X_2 = 3.5$

Logit =  $-6 + (0.05 \times 40) + (1 \times 3.5) = -0.5$

So,

$$P(Y = 1|X_1 = 40, X_2 = 3.5) = \frac{1}{1 + e^{0.5}}$$

Calculating  $e^{0.5} \approx 1.6487$ :

$$P(Y = 1|X_1 = 40, X_2 = 3.5) = \frac{1}{1 + 1.6487} \approx \frac{1}{2.6487} \approx 0.3775$$

So, estimate that the probability is around **37.5%**

- (b) How many hours would the student in part (a) need to study to have a 50 % chance of getting an A in the class?

$$P(Y = 1|X_1, X_2 = 3.5) = 0.5$$

Correspond to logit = 0, so

$$\beta_0 + \beta_1 X_1 + \beta_2 X_2 = 0$$

Substitute the known values :

$$-6 + 0.05X_1 + 1 \times 3.5 = 0$$

Get :

$$1. -6 + 3.5 = -2.5$$

$$2. 0.05X_1 = 2.5$$

$$3. X_1 = \frac{2.5}{0.05} = 50$$

Therefore, the student would need to study **50 hours** to have a 50% chance of getting an A in the class.