09/28/23

papergrid

Date: / /

HW-2

Name: Harin Rishab
Witname: harinris

3a) Let,

$x1$ = GPA
$x2$ = IQ
$x3$ = Level
x_4 = Interaction b/w GPA and IQ
x_5 = Interaction b/w GPA and Level

$$Salary = 50 + 20x1 + 0.07x2 + 35x3 + 0.01x4 - 10x5$$

Now for fixed IQ and GPA (x_1, x_2),

$$50 + 20x_1' + 0.07x_2' + 0.01(x_1' \cdot x_2')$$

$$Salary\ (college) = 50 + 20x_1' + 0.07x_2' + 35 \times 1 + 0.01(x_1' \cdot x_2') - 10(x_1' \cdot 1)$$

$$= 50 + 20x_1' + 0.07x_2' + 35 + 0.01(x_1' \cdot x_2') - 10(x_1')$$

$$= Salary\ (high\ school) + 35 - 10(x_1')$$

$\therefore$ Salary (college) - Salary (high school) $= 35 - 10x_1'$

If salary difference >= 0,

$$35 - 10x_1' >= 0 \quad x_1' <= 3.5$$

If salary difference <= 0,

$$35 - 10x_1' <= 0 \quad x_1' >= 3.5$$

Hence, for a fixed value of IQ and GPA, high school students earn more on avg if their GPA is greater or equal than 3.5

Option (iii) is correct.

b)

Salary = 50 + 20(4) + 0.07(110) + 35 + 0.01(110 × 4) − 10(4)

= 137.1 or $137,100

c) False. Coefficient does not indicate statistical significance

4a) It is difficult to say. Since cubic regression is more flexible it might have a smaller RSS compared to the linear regression. Cubic regression would do a better fit. However, the true relationship b/w X and Y is linear. This might cause linear regression to have a smaller RSS.

b) In this case, it is safe to say linear regression will have less RSS as cubic regression would overfit the data.

c) Here, cubic regression would better fit the data due to higher flexibility.

d) It is difficult to tell in this case because we do not know the underlying true relationship. It depends on how close the true relationship is to linear or cubic form.