Assignment_11

John Bute

2024-10-14

Suppose we have a dataset with five predictors: X1 = GPA, X2 = IQ, X3 = Gender (1 for Female and 0 for Male, other genders are not represented in the data), X4 = Interaction between GPA and IQ, X5 = Interaction between GPA and Gender. The response is the starting salary after graduation (in thousands of dollars). We use the least squares to fit the model, and obtain the coefficients: $\beta 0 = 50$, $\beta 1 = 20$, $\beta 2 = 0.07$, $\beta 3 = 35$, $\beta 4 = 0.01$, $\beta 5 = -10$.

1. Which of the following statement is valid? Why?

i. For a fixed value of IQ and GPA, males earn more on average than females.

Let us split the model's coefficients as if we were calculating the starting salary for male, and a starting salary for females.

Females: As a female, X3 = 1. Thus the B3 coefficient is factored into the result. Furthermore, X5 is the interaction between GPA and Gender. Thus X5 = GPA * Gender = X1 * 1 = X1 Y(Female) = 50 + 20X1 + 0.07X2 + 35 + 0.01X4 - 10X1

Male: As a male, X3 = 0. Thus the B3 coefficient does not affect the final prediction on starting salary. Furthermore, X5 = GPA * Gender = X1 * 0 = 0. Therefore, X5 does not have an impact on the final prediction for starting salary. As a result Y(Male) = 50 + 20X1 + 0.07X2 + 0.01X4

The difference Y(female) - Y(male) = (50 - 50) + (20x1 - 20x1) + 35 + (0.01X4 - 0.01X4) - 10X1 =35 - 10X1

The difference between starting salaries of males and females is 35 - 10x1. However, if X1 > 3.5, males earn more than females. However, if the GPA is X1< 3.5, females earn more than males.

The statement "For a fixed value of IQ and GPA, males earn more on average than females" is only conditionally valid if the fixed GPA of these students is set at greater than 3.5. Otherwise it is false.

ii. For a fixed value of IQ and GPA, females earn more on average than males.

If we see the explanation above, we found that the difference between Males and females in terms of starting salary, provided that IQ and GPA stay fixed, is 35 - 10X1 where x1 is GPA. As we saw above, if the fixed GPA is X1> 3.5, then females earn more on avergae than males.

Thus, the statement is also conditionally valid if the fixed GPA of these students is set at less than 3.5. Otherwise it is false.

iii. For a fixed value of IQ and GPA, males earn more on average than females provided that the GPA is high enough.

As shown above, if the fixed value of the GPA is X1> 3.5, then females will earn less than males in terms of starting salary, then the statement For a fixed value of IQ and GPA, males earn more on average than females provided that the GPA is high enough is valid.

iv. For a fixed value of IQ and GPA, females earn more on average than males provided that the GPA is high enough.

We proved in iii that for a fixed value of IQ and GPA, males earn more on average than females provided that the GPA is high enough. When we analyzed the problem deeper, we saw that if the fixed value of GPA of students is X1<3.5, then females earn more than males in terms of starting salary. Therefore, the statement is invalid, it is actually the opposite: if the GPA is low enough, then women will earn more than males.

2. Predict the salary of a female with an IQ of 110 and a GPA of 4.0.

Y(Female) = 50 + 20X1 + 0.07X2 + 35 + 0.01X4 - 10X5

X1 = GPA = 4.0 X2 = IQ = 110 X3 = Gender = 1 X4 = GPA*IQ = 440 X5 = GPA* Gender = 4
Y(female) = 50 + 20 * 4 + 0.07 * 110 + 35 + 0.01 * 440 - 10 * 4 = 137.1

Thus, the predicted salary of a female with an IQ of 110 and a GPA of 4.0 is 137,100$

3. True or false: Since the coefficient for the GPA/IQ interaction term is very small, there is very little evidence of an interaction effect. Justify your answer

False, as the coefficient for the term B4 = 0.01 does not tell us the strength or significance of the interaction effect but rather tells us how much this influences the predicted salary. A small coefficient does not mean that the interaction effect is insignificant.

You would need further testing such as hypothesis tests and confidences intervals to determine whether there is evidence of an interaction effect.