Homework 7 A20545853

Q (1) Given linear regression!

$$\text{Salary} = \beta_0 + \beta_1(\text{Years of Experience}) + \beta_2(\text{Management}) + \beta_3(\text{Years of Experience} \times \text{Management})$$

where;

Salary : Continuous variable (in thousands of dollars)

Years of Experience : Continuous variable (in years)

Management : categorical variable (1 = Management , 0 = Non-Management)

(a) : To explain, what the coefficient β_3 represents and what does it mean for the relationship between years of experience and salary.

⇒ The coefficient β_3 represents the interaction between the years of experience and management.

In general, it captures how the relationship between salary and years of experience changes, depending on whether an employee is in management or non-management.

we can say it as;

$\beta_3 =$

- positive: The management employees gets greater increase in salary than non-management for one year of experience.
- negative: The management employees gets smaller increase in salary than non-management for one year of experience.
- zero: The relationship between salary and years of experience is the same for both management and non-management.

(b) As we run the regression, we have obtain these estimated coefficients:

$$\hat{\beta}_0 = 35$$
$$\hat{\beta}_1 = 1.5$$
$$\hat{\beta}_2 = 10$$
$$\hat{\beta}_3 = 2.5$$

(i) For an employee with 5 years of Experience in a management position, what is their predicted salary?

For a management employee,

management $= 1$, Years of Experience $= 5$

substituting the above in the linear regression, we get

$$\text{Salary} = 35 + 1.5 \times 5 + 10 \cdot 1 + 2.5\,(5 \times 1)$$

$$= 35 + 7.5 + 10 + 12.5$$

$$\text{Salary} = 65$$

Therefore, the predicted salary is \$ 65,000.

(ii) For an employee with 5 years of Experience in a non-management position, what is their predicted salary?

for non-management employee,

Management $= 0$, Years of Experience $= 5$

substituting the above in the linear regression, we get

$$\text{Salary} = 35 + 1.5 \times 5 + 10 \cdot 0 + 2.5\,(5 \times 0)$$

$$= 35 + 7.5 + 0 + 0$$

$$\text{Salary} = 42.5$$

Therefore, the predicted salary is \$ 42,500.

(c) Study the effect of experience on salary for different groups,

(i) For non-management employees, the regression is written as:

$$\text{Salary} = \hat{\beta}_0 + \hat{\beta}_1 \cdot \text{Years of Experience}$$

As for non-management, management $= 0$

→ The effect of an additional year of experience is represented by the coefficient $\beta_1 = 1.5$.

Therefore, for non-management employees, for each additional year of experience, salary increases by \$1,500.

(ii) For management employees, the interaction coefficient β_3 adds to the prior effect of years of experience.

∴ The effect for each additional year of experience is equal to $\beta_1 + \beta_3 = 1.5 + 2.5 = 4$.

This does mean that for management employees, for each additional year of experience, salary increases by \$4,000.

→ The difference between management and non-management employees is: $4 - 1.5 = 2.5$

Hence, the management employees experience an additional \$2,500 increase in salary per year of experience compared to non-management employees

Q(2) Given regression model:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 (X_1 * X_2) + \varepsilon.$$

After fitting the model, the p-values for the estimated coefficients are as follows,

Terms	Coefficient $(\hat{\beta})$	p-value
β_0	50	0.001
β_1	0.2	0.35
β_2	1.5	0.40
β_3	5	0.02

(a) The hierarchy principle in regression analysis, states that if an interaction term is included in the model, the main effects that make up the interaction term (ie., X_1 and X_2) should also be included. As the interaction term depends on the main effects and removing them would make the interception of the interception invalid. The main effects provided necessary context to understand the interaction.

→ The p-value of the interaction term ($\beta_3 = 5$) is 0.02, which is statistically significant ($p < 0.05$)

→ The p-values for the main effects, Training Hours (X_1) and Experience (X_2) are 0.35 and 0.40, implies neither is statistically significant.

Based on the hierarchy principle, the main effects should still be included in the model along with the interaction term, even though they are not significant. Excluding them would make the interaction term diff-cult to interpret.

(b) The p-values for Training Hours (X_1) and Experience (X_2) are large as that mean that individually, these variables do not have a significant effect on productivity.

However, the p-value for the interaction term is small (0.02) shows that together, training and experience have a significant combined effect on productivity. This suggests that the impact of training hours depends on the level of experience and viceversa. for instance, training might be more beneficial for employees with more experience, or experiene might enchance the effectivenen of training.

Therfore, while neither training nor experience alone significantly affects productivity, their interaction shows that the combination of these factors is important.