Homework 2

Rento Saijo

Department of Mathematics, Connecticut College

STA336: Statistical Machine Learning

Yan Zhuang, Ph.D.

February 6, 2026

Disclosure

ChatGPT-5.2 was used to create the `YAML` portion and some `LaTeX` code to format the text/equations nicely; I looked into the documentation for each of the package it used and added/removed unnecessary formattings. See the original `RMD` file here.

Problem 1 (a)

Given estimates

$$\hat{\beta}_0 = 50, \quad \hat{\beta}_1 = 20, \quad \hat{\beta}_2 = 0.07, \quad \hat{\beta}_3 = 35, \quad \hat{\beta}_4 = 0.01, \quad \hat{\beta}_5 = -10,$$

the fitted regression function is

$$\hat{Y} = 50 + 20\,\mathrm{GPA} + 0.07\,\mathrm{IQ} + 35\,\mathrm{Level} + 0.01(\mathrm{GPA} \cdot \mathrm{IQ}) - 10(\mathrm{GPA} \cdot \mathrm{Level}).$$

For fixed GPA and IQ, let us compare predicted salary for college versus high school:

$$\hat{Y}_C = 50 + 20\,\mathrm{GPA} + 0.07\,\mathrm{IQ} + 35 + 0.01(\mathrm{GPA} \cdot \mathrm{IQ}) - 10\,\mathrm{GPA},$$

$$\hat{Y}_{HS} = 50 + 20\,\mathrm{GPA} + 0.07\,\mathrm{IQ} + 0.01(\mathrm{GPA} \cdot \mathrm{IQ}).$$

The difference between them is

$$\hat{Y}_C - \hat{Y}_{HS} = 35 - 10\,\mathrm{GPA}.$$

College earns more when

$$\hat{Y}_C > \hat{Y}_{HS} \quad \iff \quad \hat{Y}_C - \hat{Y}_{HS} > 0 \quad \iff \quad 35 - 10\,\mathrm{GPA} > 0.$$

Solve:

$$35 > 10\,\mathrm{GPA} \quad \iff \quad \frac{35}{10} > \mathrm{GPA} \quad \iff \quad 3.5 > \mathrm{GPA}.$$

Thus:

$$\text{If GPA} < 3.5,\ \hat{Y}_C > \hat{Y}_{HS}; \qquad \text{if GPA} > 3.5,\ \hat{Y}_{HS} > \hat{Y}_C.$$

Performing the basic algebra shown above, we see that $\hat{Y}_C > \hat{Y}_{HS}$ when GPA < 3.5, and $\hat{Y}_{HS} > \hat{Y}_C$ when

GPA > 3.5. Therefore, (iii) is the correct statement: high school graduates earn more than college graduates provided that GPA is high enough (specifically, GPA > 3.5).

Problem 1 (b)

For a college graduate with IQ $= 110$ and GPA $= 4.0$:

$$\begin{aligned} \hat{Y} &= 50 + 20(4) + 0.07(110) + 35(1) + 0.01(4 \cdot 110) - 10(4 \cdot 1) && (1) \\ &= 50 + 80 + 7.7 + 35 + 4.4 - 40 && (2) \\ &= 137.1. && (3) \end{aligned}$$

Therefore, the predicted starting salary is

$$\boxed{\hat{Y} = 137.1 \text{ (thousand dollars)} = \$137{,}100.}$$

Problem 1 (c)

False. The numerical size of an interaction coefficient cannot, by itself, be used to judge whether an interaction is present. First, the practical impact of the interaction depends on the scale of the predictors: since the interaction term is GPA $\cdot$ IQ and IQ values are often around 100, the term (contribution)

$$\hat{\beta}_4(\text{GPA} \cdot \text{IQ}) = 0.01(\text{GPA} \cdot \text{IQ})$$

can be nontrivial. For example, at GPA $= 4$ and IQ $= 110$,

$$0.01(4 \cdot 110) = 4.4,$$

which corresponds to \$4,400 in predicted salary (because Y is measured in thousands of dollars). (I could certainly use an extra few thousand dollars lol.) Second, the "statistical evidence" for an interaction effect (or any effect) is assessed using inference for β_4, such as a hypothesis test

$$H_0 : \beta_4 = 0 \quad \text{vs.} \quad H_A : \beta_4 \neq 0,$$

which yields a t-statistic and corresponding p-value, or equivalently by checking whether a confidence interval for β_4 includes 0. All in all, even a coefficient that appears "small" could be statistically significant and practically meaningful, while a larger coefficient could fail to be significant if its standard error is large.

Problem 2 (a)