Assignment 2
INFO H-515

An Introduction to Statistical Learning (with python application) Chapter 3

Exercise-3:

a)

True- iii. For a fixed value of IQ and GPA, high school graduates earn more, on average, than college graduates provided that the GPA is high enough.

Let,

- x1 = GPA
- x2 = IQ
- x3 = Level (College - 1, High School - 0)
- x4 = Interaction b/w GPA and IQ (x1.x2)
- x5 = Interaction b/w GPA and Level (x1.x3)
- Salary = b0 + b1x1 + b2x2 + b3x3 + b4x4 + b5x5 = 50 + 20x1 + 0.07x2 + 35x3 + 0.01x4 - 10x5

for fixed IQ and GPA at x1' and x2': - Salary (high school) = 50 + 20x1' + 0.07x2' + 35*(0) + 0.01(x1'.x2') - 10(x1'.0) = 50 + 20x1' + 0.07x2' + 0.01(x1'.x2')

- Salary (college) = 50 + 20x1' + 0.07x2' + 35*(1) + 0.01(x1'.x2') -10(x1'.1) = 50 + 20x1' + 0.07x2' + 35 + 0.01(x1'.x2') - 10(x1') = Salary (high school) + 35 - 10(x1')

From here:

Salary (college) - Salary (high school) = 35 - 10x1'

Assuming the salary difference to be more than equal to zero, we get:

35 - 10x1' >= 0 x1' <= 3.5

Assuming the salary difference to be less than equal to zero, we get:

35 - 10x1' <= 0 x1' >= 3.5

Hence, for a fixed value of IQ and GPA, high school graduates earn more, on average, than college graduates provided that the GPA is more than equal to 3.5

b)

Salary = 50+20(4)+0.07(110)+35+0.01(110x4)-10(4) = 137.1 (in thousands of dollars)

c)

False because the magnitude of coefficient is not an indicator of statistical significance.

Open Intro Statistics (4th Edition) Chapter 5:

Exercise-5.8:

52% of adult twitter users get at least some news on twitter.

standard error for this estimate = 2.4%

So, a normal distribution can be used here.

The critical value for 99% conf int=2.575829

Lower limit: 0.52 - (2.575829)(0.024) = 0.4582

Upper limit: 0.52 + (2.575829)(0.024) = 0.5818

0.4582< sample proportion < 0.5818

The interval extends 0.0618 in both directions, indicating a margin of error of 0.0618.

In terms of percentages, between 45.82% and 58.18% of U.S. adult Twitter users who receive news on the platform, with a margin of error of 6.18%.

It's important to note that the 6.18% margin of error pertains to the percentage of individuals who receive news on Twitter, and it does not represent the margin of error for the difference between the percentage of those receiving news on Twitter and those who do not.

Exercise-5.16:

a)

H_o : There is no change in the average calorie intake for diners.

H_o : $\mu=1100$

H_a : There is a difference in calorie intake for diners.

H_a : $\mu\neq1100$

b)

H_o : the rate is equal from the national rate of 70% =0.70

H_o : $p = 0.70$

H_a : the rate is different from the national rate of 0.70

H_a : $p \neq 0.70$