

1. **Make a scatter plot** using the following data about Body Mass Index (BMI) and Average Hours of Exercise Per Week (AHEPW) observed from male subjects, and **explain the relation between the two variables**.

Subject ID	1	2	3	4	5	6	7
BMI	20	18	23	19	28	30	21
AHEPW	4	7	3	6	2	1	5

- a) Test for linear correlation at 5% level of significance, to see if the correlation is significantly different from zero, and state the conclusion. Use p-value approach.
H₀:
H_a:
p-value:
- b) Find the **least square linear regression equation** for estimating average **BMI** using **AHEPW** variable.
- c) What is the predicted average BMI for the sample population when AHEPW is 4? Use result from part b).
- d) Which of the two variables in a) is the **explanatory variable**?
- e) Find the 95% confidence interval for estimating the average BMI for those whose AHEPW is 4.
- f) Find the 95% confidence interval for estimating the BMI for an individual whose AHEPW is 4.

2. A company has two factories manufacturing their products. To study whether two factories have the similar performance in terms of their defective rates, two random samples of products were examined and the results are in the following contingency table. Out of products sampled from factory A, 6 were defective and 210 were non-defective. Out of products sampled from factory B, 25 were defective and 475 were non-defective. At $\alpha = 0.05$, perform a chi-square test for independence to see if there is correlation between quality of the product and where the product is from (or is there a significant difference in defective rates between the two factories). Use p-value approach.

	Defective	Non-defective
Factory A	6	210
Factory B	25	475

H_0 :

H_a :

p-value:

Conclusion: