Problem Set 3 - Simple Linear Regression

ECO 204 - Statistics for Business and Economics - II, Summer 2025
Faculty: Shaikh Tanvir Hossain, **TA:** Habiba Afroz
Due Date: 7th August, 10:00 PM, 2025

§

Instructions : Please form a group of three (max) and submit on Google Classroom by the due date. Submit the Excel / R file with the calculations and your answers in a single PDF file (handwritten solutions are fine). Please write all group members names and ID numbers on the first page of the PDF file.

1. You have a data set titled `data_msales_aprice.xlsx`, it has an independent variable average price in BDT, and a dependent variable monthly sales (measured in 1000 taka), now answer following questions,
   (a) How do you think the average price and monthly sales are related? Just give a qualitative answer without any data work?
   (b) Calculate the covariance between average price and monthly sales. Write the formula you used to calculate the covariance. What can you say about the relationship based on the covariance?
   (c) Calculate the correlation between average price and monthly sales. Write the formula you used to calculate the correlation. What can you say about the relationship based on the correlation?
   (d) Why do we need correlation when we have covariance?
   (e) Plot the scatter plot between average price and monthly sales. Do you see any linear relationship?
   (f) Calculate the slope and intercept of the regression line between average price and monthly sales. What is the slope and intercept?
   (g) Write the regression equation or equation of the best fitted linear line.
   (h) How would you interpret the intercept and slope in this case?
   (i) Calculate the in-sample predicted values (or fitted values) of monthly sales using the regression equation. What is the sum of squared error in this case?
   (j) Calculate the predicted values of monthly sales when average price is 300 BDT and also 1000 BDT.
   (k) What is the R-squared value of the regression model? How do you interpret this value?

   Additional Questions - don't have to submit, but please practice

   (l) What is the formula for the Standard Error of $\widehat{\beta_1}$? Using the formula explain if the variance of X variable decreases (so we have less variance of X variable) what happens to the Standard Error, is this good if we consider interval estimation or testing.
   (m) What is the formula for the Estimate of the Standard Error of $\widehat{\beta}_1$, calculate this using MSE and $\sum_{i=1}^{n}(x_i - \overline{x})^2$
   (n) Construct the 90% and 95% confidence of the slope coeffcient assuming Normality of conditional error?
   (o) Is Aprice a statistically significant variable, write down the hypothesis clearly first and do the significance testing at 5% level.
   (p) Is the slope coefficient of Apice in population is at least -0.5? Write down the hypothesis clearly first and do the significance testing at 5% level.