MCP261 IE Lab I

Exercise 8: Analytics III: Hypothesis Testing and Linear Regression

Due date: 6 April 2024, 11:59 PM

Solve problems 1-3 using Excel. Submit one Excel file for problems 1-3. The Excel file should contain three worksheets, one for each problem. For example, your Excel file should contain worksheets named `Q1 solution', `Q2 solution' and so on. All your answers need to be in the Excel file itself. The solution to each question, including calculations, must be clearly annotated, with appropriate explanatory text. Worksheets that are not readable or comprehensible will not be graded.

The file name format for your submission must be: EntryNumber_Name_Ex8.xlsx.

1. (5 marks) Dataset for this problem is available here: https://tinyurl.com/ex8q1data

Assume that the random sample (`Sample 1', in column A) is sampled from a Gaussian distribution. For every hypothesis test below, report the p-value, the critical value of the test statistic and the confidence interval for the population parameter that is the subject of the hypothesis test.

(a) Conduct hypothesis tests to check whether the population mean μ is 10 or not, assuming that the population variance is 3 at significance levels of 0.05 and 0.10. Report the p-value, the critical value of the test statistic and the confidence intervals for both cases.

(b) Conduct the above tests (i.e., at $\alpha_0 = 0.10$ and 0.05) to determine whether the population mean is greater than equal to 11 or not assuming the population variance is not known.

(c) Conduct the test in question 1(a) varying μ_0 from 11 to 13 in increments of 0.1 and plot the p-value for each test against the value of μ_0. The plot must be neatly labelled, and the p-values used for the plot must be displayed in an appropriately labelled column. From the plot, what can you infer regarding the `true' value of the population parameter μ?

2. (5 marks) Dataset for this problem is available here: https://bit.ly/2FDvKG6

Copier maintenance. The Tri-City Office Equipment Corporation sells an imported copier on a franchise basis and performs preventive maintenance and repair service on this copier. The data below have been collected from 45 recent calls on users to perform routine preventive maintenance service; for each call, X is the number of copiers serviced and Y is the total number of minutes spent by the service person. Assume that first-order regression model (1.1) is appropriate.

i:	1	2	3	...	43	44	45
X_i:	2	4	3	...	2	4	5
Y_i:	20	60	46	...	27	61	77

a. Obtain the estimated regression function.

b. Plot the estimated regression function and the data. How well does the estimated regression function fit the data?

c. Interpret b_0 in your estimated regression function. Does b_0 provide any relevant information here? Explain.

d. Obtain a point estimate of the mean service time when $X = 5$ copiers are serviced.

3. (5 marks) Solve the following problem, with respect to the data from the previous problem.

a. Estimate the change in the mean service time when the number of copiers serviced increases by one. Use a 90 percent confidence interval. Interpret your confidence interval.

b. Conduct a t test to determine whether or not there is a linear association between X and Y here; control the α risk at .10. State the alternatives, decision rule, and conclusion. What is the P-value of your test?

c. Are your results in parts (a) and (b) consistent? Explain.