Problem Set 2 - Testing, Proportion Problems

ECO 204 - Statistics for Business and Economics - II, Summer 2025
Faculty: Shaikh Tanvir Hossain, **TA:** Habiba Afroz
Due Date: 16th July, 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.

Testing for Populaion Mean

[Note: For the following questions, you can use the data file **Student_Lunch.xlsx** where the amounts are recorded in an Excel file. Although for the testing problems you can do the testing either following the **critical value approach** or **p-value approach**. For all the problem below please use both approaches. As you know in general, both method will give you the same conclusion, but it is good to practice both methods here.]

1. Suppose someone named **Rocky Balboa** visits East West University and had a lunch in the cafeteria. He is worried about the quality of the food here and wants to do some hypothesis testing regarding the amount spent on lunch by the students at the Economics Dept. at EWU. To do this he collected a sample of $n = 64$ students from the Economics Dept. and records the amount spent on lunch by each student. The data is provided in the file **Student_Lunch.xlsx**. Assuming the amount spent on lunch is Normally distributed with a population standard deviation of $\sigma = 6$, Rocky wants to do the following hypothesis tests:
   (a) Suppose Rocky wants to test the hypothesis that **the average amount spent for lunch by the students at the Economics Dept. is** 20 **TK**. Clearly write both the hypotheses and check whether Rocky is correct at 5% significance level. Do the test using both the critical value approach and the p-value approach.
   (b) Now suppose Rocky wants to test the hypothesis that the average amount spent for lunch by the students at the Economics Dept. is at least 20 TK. Now test whether Rocky is correct at 1% significance level. Do the test using both the critical value approach and the p-value approach.
   (c) Finally, suppose Rocky wants to test the hypothesis that the average amount spent for lunch by the students at the Economics Dept. is at most 20 TK. Again test whether Rocky is correct at 10% significance level. Do the test using both the critical value approach and the p-value approach.
2. Continuing from question 1, do the same three tests (in (a), (b), and (c)) as in the previous question, but now assume that the population standard deviation is unknown, however, you can continue to assume that the amount spent on lunch is Normally distributed. Do the test using both the critical value approach and the p-value approach. .
3. Finally, do the same three tests as in the previous two questions, but now assume that the amount spent on lunch follows some arbitrary probability distribution (i.e., the population distribution is not known). Again check whether Rocky is correct at, you can use the same significance levels as in the previous two questions. Do the test using both the critical value approach and the p-value approach.

Estimation and Testing for Population Proportion

4. Suppose now Rocky collected a dataset from the same students at the Economics Dept. at EWU, and asked them *whether they are satisfied with the quality of food served at the cafeteria*, the answers are recorded as $Yes = 1$ or $No = 0$. The data is contained in the file `Student_Satisfaction.xlsx`. Now answer the following questions:
   (a) What is the population of this study?
   (b) If our goal is to find out whether the students in general, not just from the Economics dept. are satisfied with the quality of food served at the cafeteria, is this a good sample? Why or why not?
   (c) Suppose our target parameter is the **propotion of all students at the Econmics Dept. who are satisfied with the quality of food served at the cafeteria**, let's call it p, in theory how can we calculate this number?
   (d) Of course it will be difficult to calculate the population propotion p, this is why Rocky collected a random sample of 64 students from the Economics Dept. Now calculate the sample propotion $\overline{p}$ of students who are satisfied with the quality of food served at the cafeteria.
   (e) Based on this sample what is the **point estimate of the population propotion of students who are satisfied** with the quality of food served at the cafeteria?

(f) If we think each data point (or each row) of the sample is a Bernoulli random variable, what is the mean and variance of the Bernoulli random variable? What is the estimated variance of the Bernoulli random variable in this case?

(g) If we assume i.i.d. assumption, this means all rows in the random sample follow the same distribution and they are independent of each other. Then what is the mean and variance of sample propotion $\overline{p}$ when it is a random variable (i.e., $\overline{p}$ is random when we think about repeated sampling)?

(h) What is the standard error of the sample propotion?

(i) Construct a 90% confidence interval for the population propotion of students who are satisfied with the quality of food served at the cafeteria. Interpret the confidence interval.

(j) Now suppose we want to test the hypothesis that the propotion of students who are satisfied with the quality of food served at the cafeteria is at least 0.7. Test this hypothesis at 5% significance level.

"The Only Relevant Test of the Validity of a Hypothesis is Comparison of Prediction with Experience."

- Milton Friedman