Homework 4.1

The first 5 questions and the Part II question require the use of the HR dataset on AzureML. This is the same dataset used in the week 1 homework. You can continue by simply modifying your week 1 homework experiment, or you can create a new experiment where you can import the data by copying-and-pasting the following link:

https://docs.google.com/spreadsheets/d/1ub_0Y5CEj2HtKO6bgb8OIvMIG81Wca0r_q8G70OKUTc/pub?gid=1709232748&single=true&output=csv.

First, use an appropriate SQL query to get a dataset containing the satisfaction_level and last_evaluation of all employees in the sales department who stayed in the company (left = 0). We'll call this the "Stayed" group. Do a similar SQL query for employees in the sales department who left the company (left = 1). Call this the "Departed" group.

Verify that you have executed this correctly. If you did, you should have 3126 rows in your "Stayed" group, and 1014 in the "Departed" group. Do not proceed until you have completed this correctly.

Each question is worth 1 point.

Question 1

1/1 point (graded)

Please check the box below to confirm that you have correctly prepared the data sets, and verified that the two groups have 3126 and 1014 observations respectively.

☑ I have correctly prepared the data sets.

✔

Submit You have used 2 of 2 attempts

Question 2

1/1 point (graded)

Suppose you want to set up a hypothesis test to establish statistical support for the assertion that the mean satisfaction level for the Departed group is less than 0.47. What is the appropriate alternative hypothesis?

- (●) Mean satisfaction level for Departed group is less than 0.47. ✔
- ( ) Mean satisfaction level for Departed group is less than or equal to 0.47.
- ( ) Mean satisfaction level for Departed group is greater than 0.47.
- ( ) Mean satisfaction level for Departed group is greater than or equal to 0.47.

Submit You have used 2 of 2 attempts

Question 3

1/1 point (graded)

What is the p-value for the test in Question 2?

- 0.000
- 0.003 ✔
- 0.997
- 1.000

Submit You have used 2 of 2 attempts

Question 4

1/1 point (graded)

Based on the p-value, what do you conclude?

- We support the claim that mean satisfaction level for Departed group is less than 0.47. ✔
- We do not support the claim that mean satisfaction level for Departed group is less than 0.47.
- We support the claim that mean satisfaction level for Departed group is greater than 0.47.
- We do not support the claim that mean satisfaction level for Departed group is greater than 0.47.

Submit You have used 2 of 2 attempts

Question 5

1/1 point (graded)

Suppose you were looking for evidence that the average last evaluation for the Stayed group was greater than 0.70. Which of the following is the smallest level of significance at which you would conclude that the average last evaluation for the Stayed group is greater than 0.70?

- 0.50
- 0.10
- 0.05
- 0.01 ✔

Answer

Correct:

Answer explanation: We do a SingleSampleSet test with Null hypothesized μ set to 0.70 and the Hypothesis type set to OneTailGT. The resulting p-value is 0.00072. We reject the null, and therefore obtain the stated conclusion, if the p-value is less than the level of significance. So all of the listed levels of significance suffice, and the smallest among them is 0.01.

Submit

You have used 2 of 2 attempts

Question 6

1/1 point (graded)

Which of the following is the closest to the meaning of p-value?

- ○ It is the probability that the null hypothesis is false.
- ○ It is the probability that the null hypothesis is true.
- ◉ It is the probability of obtaining a sample at least as extreme as the observed data if the null hypothesis is true. ✔
- ○ It is the probability of obtaining a sample no more extreme than the observed data if the null hypothesis is true.

Submit You have used 2 of 2 attempts

Question 7

0/1 point (graded)

A medical research company has created a new drug. It tested it with a clinical trial, and found that according to the clinical trial, the new drug is no better than a placebo. Which of the following may have happened? The correct answer to this question requires selection of more than one of the following 4 options, and for full credit you must select all correct options and no incorrect options.

- ☑ Type I error: the drug is no better than the placebo, but clinical trial indicates it is better.
- ☑ Type II error: the drug is better than the placebo, but clinical trial indicates it is not. ✔
- ☑ No error: the drug is no better than the placebo, and clinical trial indicates it is no better. ✔
- ☐ No error: the drug is better than the placebo, and clinical trial confirms that.

✖

Submit You have used 2 of 2 attempts

Answers are displayed within the problem

Question 8

1/1 point (graded)

Which of the following is an advantage of doing a paired test when doing A/B testing?

- We can use a lower level of significance for paired tests, resulting in greater statistical accuracy.
- Paired tests can be conducted with larger sample sizes, so they can be more accurate.
- In a paired test, a Type-I error is impossible. Therefore, our result is more likely to be correct.
- In a paired test we control for all or almost all other sources of variation, so we can focus on the factor being tested. ✔

Submit You have used 2 of 2 attempts

Question 9

1/1 point (graded)

Suppose we were creating a new website, and wanted to do a full factorial design to figure out what combination of features is best. The features being tested are two: background color (for which there are 4 options: white, grey, blue, pink), and main graphic (for which there are 2 options: nature, and city). Assume that viewers' reactions to background color has no relationship with their reaction to main graphic. In a full factorial design, how many combinations do we need to test?

- ○ 2
- ○ 4
- ○ 6
- ◉ 8 ✔

Submit You have used 2 of 2 attempts

Question 10

0/1 point (graded)

Continuing with the context of Q9, suppose we tested only 3 combinations: (white, nature), (grey, nature), and (blue, city). Suppose we find that viewers rated the (grey, nature) combination as worse than the (blue, city) combination. Which of the following conclusions can we validly infer based on this?

- ◉ Viewers prefer the color blue to the color grey, regardless of main graphic. ✖
- ○ Viewers prefer the city graphic over the nature graphic, regardless of the background color.
- ○ Viewers are indifferent about both background color and main graphic.
- ○ None of the above. ✔

Submit You have used 2 of 2 attempts

Answers are displayed within the problem