Hypothesis Testing Assignment

1) A F&B manager wants to determine whether there is any significant difference in the diameter of the cutlet between two units. A randomly selected sample of cutlets was collected from both units and measured? Analyze the data and draw inferences at 5% significance level. Please state the assumptions and tests that you carried out to check validity of the assumptions.

Dataset: - Cutlet.csv

Null Hypothesis: - There is no significant difference between diameter of cutlet between 2 units.

Alternative Hypothesis:- There is significant difference between diameter of cutlet and action need to be taken.

Solution: -

There are two population, Unit A and Unit B whose value are continuous.

- First Step is Check whether Y1 and Y2 are Normal?

Null Hypothesis: - If Y1 and Y2 Normal proceed to comparing external conditions are same?

Alternative Hypothesis: - If both not normal or any one the population are not Normal, procced to Mann Whitney Test.

Normality Test

Unit A, P-Value = 0.32 > 0.05 . P high Ho Fly

Unit B, P-Value = 0.52 > 0.05 P high Ho Fly

Since both the population are accepted as Null hypothesis

- Check for external Condition are same?
  Yes, in this scenario External Condition are same.

- Perform, Paired T Test.

  Null Hypothesis: - If p-value > 0.05, go with null Hypothesis.

  Alternative Hypothesis: - If P-value < 0.05, go with Alternative Hypothesis.

  Since p-value is greater that > 0.05, we conclude with Null Hypothesis.

  Null hypothesis [p high Ho fly]: - There is no significant difference between diameter of cutlets from two different units.

---

2) A hospital wants to determine whether there is any difference in the average Turn Around Time (TAT) of reports of the laboratories on their preferred list. They collected a random sample and recorded TAT for reports of 4 laboratories. TAT is defined as sample collected to report dispatch. Analyze the data and determine whether there is any difference in average TAT among the different laboratories at 5% significance level.

Dataset: - LabTaT.csv

Solution: -

There contains 4 population with continuous values.

Goal:-
Null Hypothesis:- If there is no significance difference between Turn-around Time of 4 laboratory. Take no action.

Alternative Hypothesis: - If there is significant difference between TAT of 4 Laboratory take action.

- First step is to check whether Lab1, Lab2, Lab3 and Lab4 are normal?

  Null Hypothesis: - If all the population Y1 to Yn are normal proceed to equal variance test.

  Alternative Hypothesis: - If the population are not said the be normal, perform Mood's Median Test.

  Since P-value > 0.05, go with Null Hypothesis.

- Check whether Variance are equal ?

  Null Hypothesis :- If variance are equal perform One way Anova.
  Alternative Hypothesis:- If variance are not equal, perform only Anova.

  Since P-value > 0.05, we go with Null Hypothesis.

- Perform One way Anova test?

  Hull Hypothesis:- If P-value is > 0.05, go with Null Hypothesis.
  Alternative Hypothesis: - If P-value is < 0.05, go with alternative Hypothesis.

  Since P-Value < 0.05, we go with Alternative Hypothesis.

Alternative Hypothesis [p high Ho go]: - There is significant difference between turn-around time of 4 Laboratory. Take action.

---

3) Sales of products in four different regions is tabulated for males and females. Find if male-female buyer rations are similar across regions.

Dataset: - BuyerRation.csv

Goal: -

Null Hypothesis: - Male and Female buying ratio are similar across all regions.

Alternative Hypothesis: - Male and Female Buying Ration are different across religions.

Solution: -
There are only 4 Population with Discrete data. Hence, we go with two proportion test.

- Use Chi-Squared Test.

  Null hypothesis: - If p-value > 0.05, go with Null Hypothesis.

Alternative Hypothesis: - if p-value < 0.05, go with alternative hypothesis.

Since p-value < 0.05, we go with Alternative Hypothesis.

Alternative Hypothesis [p high Ho go]: - Male and Female Buying Ration are different across religions.

---

4) TeleCall uses 4 centers around the globe to process customer order forms. They audit a certain % of the customer order forms. Any error in order form renders it defective and has to be reworked before processing. The manager wants to check whether the defective % varies by center. Please analyze the data at *5%* significance level and help the manager draw appropriate inferences

Dataset :- CustomerOrderForm.csv

Goal:-

Null Hypothesis: - Defective doesn't varies be centre. Take no action.

Alternative Hypothesis: - Defective varies by centre. Take required action.

Solution: -

Since there are 4 population and discrete data. We go with Chi-Square Test.

- Perform Chi-Square Test.

  Null Hypothesis: - If P-value > 0.05, go with Null Hypothesis.

  Alternative Hypothesis: - If P-value < 0.05, go with Alternative Hypothesis.

  Since, P value > 0.05, we could go with Null Hypothesis.

Null Hypothesis [p high Ho fly]: - Defective doesn't varies be centre. Take no action.

5) Fantaloons Sales managers commented that *%* of males versus females walking in to the store differs based on day of the week. Analyze the data and determine whether there is evidence at *5 %* significance level to support this hypothesis.

Dataset: - Fantaloons.csv

Goal: -

Null Hypothesis: - Percentage of men and female working doesn't differ based on day of the week.

Alternative Hypothesis: - Percentage of men and females working differ by day.

Solution:-

Since Data Contains 2 populations and Discrete Data, we go with 2 Proportion test.

- 2 Proportion Test.
  Null Hypothesis:- If P-value > 0.05, we go with null hypothesis.
  Alternative Hypothesis:- If P-value < 0.05, we go with alternative Hypothesis.

  Alternative Hypothesis [p high Ho go]: - Since P-Value = 8.543e-05 < 0.05 we go with alternative Hypothesis.