Hypothesis Testing Assignment 2

1.Hypothesis Formulation:

A company claims that their new energy drink increases focus and alertness.
Formulate the null and alternative hypotheses for testing this claim.

Null Hypothesis: difference in alertness level after and before having new energy drink <= 0

Alternate Hypothesis: difference in alertness level after and before having new energy drink >0

2. Significance Level Selection:

A researcher is conducting a study on the effects of exercise on weight loss. What
significance level should they choose for their hypothesis test and why?

Significance level or **α** value can be chosen based on 2 factors broadly :

a. sample size: in general statistical formulations, alpha should be a decreasing function of sample size
b. cost type I and type II error: **α** or significance level can also be defined as the probability of rejecting the true null hypothesis (Type I error) and **β** as the probability of accepting false null hypothesis (Type II error). So significance level should be chosen such as the expected losses from the type I and type II error are minimized, or else as this may have serious consequences when decision making of accepting or rejecting null hypothesis, and overall sensitivity of the test

In the above scenario of the study of effects of exercise on weigh loss, lets break down the cost of type I and type II error.

Null hypothesis: difference in mean weights before and after exercise=0

Alternate hypothesis: difference in mean weights before and after exercise ≠0

Type I error scenario: test result interprets that there is a difference in mean weight after exercise, when actually there is no difference in weights.

Type II error scenario: test result interprets that there is no difference in weights after exercise, when actually there is.

From a basic sense of judgement we can tell that type I error in this study is of greater serious consequence than type II error. Type I error will falsely provide evidence that the exercise leads to weight loss, and people undergoing the exercise may have adverse effects on health, lose their time, money etc, when there is actually no outcome. Whereas for type II error scenario, at the most the exercise will not be proved as significant, which however would not lead to a serious consequence.

Hence, in this experiment it is desirable to have lower Type I error probability, and hence a ***lower* α** value is to be chosen, even if it leads to a bit higher **β** value, which would minimize the error costs and also increase the experiment sensitivity.

So, in the light of above discussion we can conclude, that if the experiment is performed for a sample size greater than 30, we can assume an alpha value of **0.1** so that the experiment minimum error costs.

3.Interpreting p-values:

In a study investigating the effectiveness of a new teaching method, the calculated
p-value is 0.03. What does this p-value indicate about the null hypothesis?

A p-value in a hypothesis test indicates the probability value of the defined null hypothesis to be true, or in other words how much significant is the result of the test, when compared to the pre-defined significance level (alpha) of the test.

In the above experiment p-value=0.03, indicates the probability of the defined null hypothesis to come true is 3%.

If the experiment assumes as significance level of 0.05 (CI: 95%), we can reject the null hypothesis as p-value is less that desired significance level, and the result is non-significant

However, if the same experiment assumes a significance level of 0.01(CI: 99%), we cannot reject the null hypothesis, as now p-value is significant as it is above the defined significance level of the experiment.

4.Type I and Type II Errors:

Describe a scenario in which a Type I error could occur in hypothesis testing. How
does it differ from a Type II error?

Type I error is the event of rejecting a null hypothesis when the null hypothesis is actually true. This a false positive scenario.

Scenario: Considering a test is being done on a patient to check for pregnancy. The doctor assumes the patient as 'not pregnant' until certain medical tests provides some significant evidence. So in this case the doctor is testing for null hypothesis that the women is not pregnant against the alternate hypothesis that the patient is pregnant.

So, if the conducted medical test provides significant evidence for the patient being pregnant when she is actually not, then we can call this event as a Type I error, or false positive. In this case we are rejecting the null hypothesis where actually alternate hypothesis should have been rejected

In corollary, to the above, if the conducted medical tests fails to provide significant evidence for the patient being pregnant, when she actually is , then we can call this event as a Type II error, or false negative . In this case we are rejecting the alternate hypothesis where actually null hypothesis should have been rejected.

6.Two-Tailed Hypothesis Test:

A researcher wants to determine if there is a difference in mean exam scores between
two groups of students. Formulate the null and alternative hypotheses for this study as a
two-tailed test.

Null Hypothesis: difference in mean scores of two groups of students=0

Alternate Hypothesis: difference in mean scores of two groups of students≠0

10.Interpreting Results:

After conducting a hypothesis test, the calculated p-value is 0.02. What can you
conclude about the null hypothesis based on this result, assuming a significance level of
0.05?

A p-value of 0.02 indicates the probability of null hypothesis to be true is 2%

As this p-value (0.02) is lesser than the assumed significance level (0.05), hence we need to reject the null hypothesis.

It also shows that considering the assumed significance level of the test, there is evidence that the probability of null hypothesis to be true is not significant.