Casper Kristiansson, 0012240453
2023-12-19

DA2210 Final (D-C Quiz)

Question 1

Evaluating a hypothesis critically can be extremely useful in specific situations. In the court of law, it is extremely important that an innocent person is not convicted of a crime rather than that a criminal is convicted. In terms of law, the goal is to prevent false positives from happening.

When we make a critical evaluation of a hypothesis it will involve two different types of errors, type I error (hypothesis is true when it is false) and type II error (hypothesis is false when it is true). In the situation of the court of law, it is extremely important to reach the goal that no type 1 errors happen, which is equivalent to the situation in that we do not want to convict an innocent person. If an innocent person is convicted without conclusive evidence the court of law risks destroying a person's life. In this process, the moral cost of convicting an innocent person is higher than the conviction of a guilty person.

Another good example where we want to avoid type 1 errors is in medical research. In medicine research, if a type 1 error happens is the result that an ineffective or dangerous drug might be accepted as a working. While type 2 errors in this situation might be bad where we might reject working medicine, it is far worse than we accept a non-working medicine.

There is a situation where we might favor type 2 errors happening than a type 1 error. These situations are especially during screening and testing. We favor the situation that we might label a person as having potential cancer when the person does not have it rather than labeling a person noncancer when they indeed have cancer. This has to do with that if we get a false positive, we can always prove that it is not true with further testing. What we state with this is that it is acceptable to increase the probability of type 2 errors happening by 10 folds even if that decreases the chance of type 1 errors happening.

This kind of concept also involves other theories and subjects. We know that if a theory or concept hypothesis is true but is indeed false it will over time by high probability of being disproven. We see this happening all the time in science where we experience paradigm shifts. Thomas Kuhn states that a paradigm is a set of practices that define a scientific discipline. This paradigm over time will be disproven. A good example of a paradigm shift is when it was believed that we had a geocentric model where the earth was the center of the universe which later got disproven.

Overall, this means that in certain situations we might favor type 1 errors over type 2 errors and vice versa. In science when type 1 errors happen it is likely that in the future with more extent research and experiment testing it gets disproven to show that the hypothesis is wrong. But in cases of example medical development or in a court of law it might be too late to discover that a type 1 error has happened. Therefore, depending on different situations, we might increase the probability of either type 1 or 2 of happening to reduce the other one.

Casper Kristiansson, 0012240453
2023-12-19

Question 2

A confidence interval is a range of values that is calculated based on sample data. The statement that the 95% confidence interval for c is [14.03, 16.27] means that the researchers are 95% confident that the true value of c lies within this interval. This statement means that there is not a 95% chance that the true value of c is in that interval but rather that across many data samples the interval has been calculated that the interval will contain the true value of c 95% of the time.

What can be concluded from the result of the algorithm is that the mean value of c which is 15.94 is inside the confidence interval which supports the statement. What we can see is that the confidence interval is relatively small which shows it has a high precision in predicting the algorithm's performance. Estimating and finding a confidence interval, especially in algorithm performance can be extremely important in specific use cases.

An important note is that the algorithm has a runtime of c*n^2. If this is wrong the analysis of the value of c might be wrong. But by the given runtime of c*n^2, we can state that the mean runtime of this algorithm will be 15.94*n^2.

Casper Kristiansson, 0012240453
2023-12-19

Question 3

The study described performs an analysis of the physiological behavior of nurses during training and how it changes over time. We see that the study measures 20 different physiological behaviors of each of the nurses. The study then collected a p-value for each of the measurements. The study found that the psychological test for self-efficacy had a p-value of 0.02. In terms of p-value, it is commonly found that there is a threshold of about 0.05 if a p-value is lower than that the threshold the result drawn from it can be considered significant, and it is unlikely that it occurred by chance. Based on the study the researchers concluded that "there was a significant change over time in self-efficacy between the two groups, where peer learning resulted in an improvement, but that other differences were not significant".

What could be concluded from the study is that no other measurement besides self-efficacy shows significant changes (their p-value being above 0.05). This shows that the peer-learning approach especially impacts the self-efficacy of the nurses. Based on this study there are a few things that the study could have improved to get better results and draw better conclusions.

Sample size and division: First, the study mentions that there are two groups of a total of 70 students the first group is 28 people, and the second group is 42 people. While this division of groups is a bit uneven, I would still think it is acceptable (the result can of course be improved by increasing the number of students in each group). However, the study doesn't mention anything about how the student was selected for the different groups. A student for example with good knowledge and experience in self-efficacy will most likely not experience an increase while someone with bad knowledge and experience with self-efficacy might. This can lead to the result being biased depending on the students in the different groups.

It could also be interesting to see how prior knowledge, demographics, and age would affect the result of the different learning approaches. Therefore, collecting data as sub-groups within the different learning approaches might also be effective in understanding the results of the study.

Control group: Another way of improving the study is by adding another group of students that has another learning approach rather than the two mentioned. This could help with finding more psychological measurements that were improved in the two learning approaches.

Comparisons: In the study, there are a lot of comparisons between the student's improvement. This means that if a recorded measurement of a student is irregular and not significant it might affect the result of the study. This means applying a correction on the data example using Bonferroni correction could help.

Length of testing: It is not mentioned how long the test was performed for the different groups. It is important to see how the learning approach affects the psychological measurements over a longer period.