Hypothesis Testing Lab 3

Question 1 – Sleep deprivation, CA vs. OR

A CDC report on sleep deprivation rates shows that the proportion of California residents who reported insufficient rest or sleep during each of the preceding 30 days is 8.0%, while this proportion is 8.8% for Oregon residents. These data are based on simple random samples of 11,545 California and 4,691 Oregon residents.

Goal: Conduct a hypothesis test to determine if these data provide strong evidence that the rate of sleep deprivation is different for the two states.

Part A: What type of hypothesis test is this? (proportion / diff. in proportions/ mean / diff. in means)

Part B: Compute $\widehat{p}_{pool}$.

Part C: Compute the test-statistic Z.

Part D: What is the p-value?

Part E: Write a conclusion to this question using 'strength of evidence'.

Question 2 – Find the P-value with t-distribution

A random sample is selected from an approximately normal population with an unknown standard deviation. Get the p-value using the `pt()` function in R. You may need to use `?pt()` in your console to see required inputs. It otherwise functions similarly to `pnorm()` and `qt()` used previously.

Part A: n = 26, T = 2.485, two-tail test

Part B: n = 18, T = 0.5, right-tail test

Question 3 – Piano

Georgianna claims that in a small city renowned for its music school, the average child takes less than 5 years of piano lessons. We have a random sample of 35 children from the city, with a sample mean of 4.6 years of piano lessons and a sample standard deviation of 2.2 years.

Evaluate Georgianna's claim using a hypothesis test.

- state hypotheses
- check conditions
- calculate test-statistic
- p-value
- conclusion

Question 4 – Diamonds

We have data on two random samples of diamonds: one with diamonds that weigh 0.99 carats and one with diamonds that weigh 1 carat. Each sample has 23 diamonds. Sample statistics for the price per carat of diamonds in each sample are provided below.

0.99 Carats

Sample mean = \$44.51, s.d. = \$13.32, n=23

1 Carat

Sample mean = \$57.20, s.d. = \$18.19, n=23

Assuming that the conditions for conducting inference using the t-distribution are satisfied, perform a hypothesis test to see if there is a difference in population prices per carat of diamonds that weigh 0.99 carats and 1 carat. (Wickham 2016)

Question 5 (Null Distributions)

Part A: Explain what a null distribution is and why we are using it.

For the following, explain what the following symbols mean. Not just the names, but the deeper meaning. Start your explanation with "For repeated samples...".

Part B: $\widehat{p} \sim N(p_0, \sqrt{p_0(1-p_0)/n})$

Part C: $\frac{\bar{x}-\mu_0}{s/\sqrt{n}} \sim t_{n-1}$

Part D: $\frac{(\bar{x}_1-\bar{x}_2)}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} \sim t_{min(n_1,n_2)-1}$

Question 6 (Conceptual)

Part A: Why do we calculate a test-statistic?

Part B: Why do we sometimes call tests left, right, or two-tailed tests?

Part C: Now that you have seen a few examples of p-values, in non-statistical jargon: What do they tell us about our data/sample?