8.1.1: Exercises

Confidence Intervals for Paired Data

Recall the five step process for constructing a confidence interval for a paired mean difference:

1. Verify that the sampling distribution of sample mean differences is approximately normal by showing the number of differences in the set is more than 30 OR the population of differences is normally distributed.
2. Find the critical value from the Student's T distribution that corresponds to the given confidence level. Let $\bar{x}$ represent the mean of sample differences, let s represent the standard deviation of sample differences, let n represent the number of differences, and let $df = n - 1$ represent the degrees of freedom.
3. Compute the margin of error: $E = T_c \cdot \frac{s}{\sqrt{n}}$
4. Compute the upper and lower limit of the interval and write the answer in interval notation: $(\bar{x} - E, \bar{x} + E)$
5. Interpret the interval and state a conclusion in context.

Hypothesis Tests for Paired Data

Recall the four step process testing a hypothesis about a paired mean difference:

1. Define a variable to represent the population mean difference and determine the hypotheses. The null hypothesis will always be $\mu = 0$.
2. Verify that the sampling distribution of sample mean differences is approximately normal by showing the number of differences in the set is more than 30 OR the population of differences is normally distributed.
3. Let $\bar{x}$ represent the mean of sample differences, let s represent the standard deviation of sample differences, let n represent the number of differences, and let $df = n - 1$ represent the degrees of freedom. Compute the T-score and the corresponding P-value. $T = \frac{\bar{x}}{\frac{s}{\sqrt{n}}}$
4. Make a decision and state a conclusion in context.

1. An educational website offers a practice program for the Law School Admissions Test (LSAT). Users of the program take a pretest and posttest. Here are the scores and gains for a random sample of 6 users:

User ID	1	2	3	4	5	6
Pretest	140	152	153	159	150	146
Posttest	150	159	170	164	148	166
Difference (Posttest - Pretest)						

   Assume the population of differences in test scores is approximately normal. Construct a 90% confidence interval for the mean difference in test scores. Use the interval to create an advertisement for the educational website.

2. Mariah wants to know if growth-minded language influences performance on mathematics exams. She randomly selects 35 students and gives each student two assessments. One assessment has growth-minded language throughout, and the other does not. She records the scores, finds the difference between the assessment scores (with growth-minded language minus without growth-minded language) for each student. She constructs a 99% confidence interval for the mean difference is scores (out of 10) to be (1.16, 2.725). Improve the following conclusions:
   a. Since 0 is not in the interval, we can conclude that there is a difference, on average.

   b. Since the two numbers in the interval are positive, we can conclude that there is a positive difference on average.

3. Theo wants to know if using melatonin helps with insomnia so he will construct a 95% confidence interval. He randomly selects 10 individuals with insomnia. He records how long each individual sleeps without any intervention. He then gives each subject a dose of melatonin every night for a week. He records the time each subject spent sleeping on the 7th day of the experiment. Research has shown that the population of differences is approximately normal. Below is the data:

Subject ID number	1	2	3	4	5	6	7	8	9	10
Before Melatonin	4.2	3.8	5.1	4.3	4	3.1	2.9	3.4	4.6	5.8
After Melatonin	4.3	4	5	5	6	3.2	3.4	2	4.1	6.3

Below is Theo's solution. Spot any errors. Explain how you would improve/correct each error.

Step 1: $n = 10 \geq 10$ so the sampling distribution is normal.

Step 2: $\bar{x}_{\text{before}} = 4.12 \text{ hours}, \bar{x}_{\text{after}} = 4.33 \text{ hours}$, so the difference is $\bar{x} = 0.21 \text{ hours}$. The standard deviation is $s = s_{\text{after}} - s_{\text{before}} = 0.408 \text{ hours}$. To calculate the critical value, I entered tdist(10).inversecdf(0.95) into desmos and found that $T_c = 1.81$.

Step 3: $E = 1.81 \cdot \frac{0.408}{\sqrt{9}} = 0.246$

Step 4: $(0.21 - 0.246), (0.21 + 0.246) = (-0.036, 0.456)$

Step 5: The true population mean is between -0.036 and 0.456.

Interpretation: Since 0 is contained in the interval, melatonin does not have an effect on sleep for individuals with insomnia.

4. Complete problem 3 correctly.

5. An educational website offers a practice program for the Law School Admissions Test (LSAT). Users of the program take a pretest and posttest. Here are the scores and gains for a random sample of 6 users:

User ID	1	2	3	4	5	6
Pretest	140	152	153	159	150	146
Posttest	150	155	160	159	148	151
Difference (Posttest - Pretest)						

Assume the population of differences in test scores is approximately normal. Test the company's claim that their program improves a customer's average score at a 5% level of significance.

6. Mariah wants to know if growth-minded language influences performance on mathematics exams. She randomly selects 35 students and gives each student two assessments. One assessment has growth-minded language throughout, and the other does not. She records the scores, finds the difference between the assessment scores (with growth-minded language minus without growth-minded language) for each student. She tests the claim that using growth-minded language to frame questions on an assessment improves the average performance on mathematics exams at a 5% level of significance. She calculates the test statistic to be T=2.87 and the corresponding P-value to be 0.0044. Improve the following conclusions.
    a. Since the P-value is less than the level of significance, we reject the null hypothesis and support the alternative hypothesis. Therefore, we have proved that using growth-minded language to frame questions on an assessment improves performance on mathematics exams.
    b. Since the P-value is less than the level of significance, the null hypothesis is false and we accept the alternative hypothesis. Therefore, the data support the claim that the difference in sample means is positive.
7. Theo wants to know if using melatonin helps with insomnia. He randomly selects 10 individuals with insomnia. He records how long each individual sleeps without any intervention. He then gives each subject a dose of melatonin every night for a week. He records the time each subject spent sleeping on the 7th day of the experiment. Research has shown that the population of differences is approximately normal. Below is the data:

Subject ID number	1	2	3	4	5	6	7	8	9	10
Before Melatonin	4.2	3.8	5.1	4.3	4	3.1	2.9	3.4	4.6	5.8
After Melatonin	4.3	4	5	5	6	3.2	3.4	2	4.1	6.3
Difference (After-Before)	0.1	0.2	-0.1	0.7	2	0.1	0.5	-1.4	-0.5	0.5

Below is Theo's solution. Spot any errors. Explain how you would improve/correct each error.

Step 1: Let μ represent the population mean difference in amount of sleep after and before using melatonin.

$H_0 : \bar{x} = 0.21$

$H_a : \bar{x} > 0.21$

Step 2: The sample is large enough so it's normal. $\bar{x} = 0.21, s = 0.871, n = 10$

Step 3: $Z = \frac{0.21}{\frac{0.871}{\sqrt{10}}} \approx 0.76$. The P-value was calculated using normaldist() in desmos and inputting the max as 0.76. The P-value is 0.7764.

Step 4: The P-value is 0.7764 which is greater than the level of significance 5%. Therefore, we accept the null hypothesis and reject the alternative hypothesis. The data support the claim that the true mean difference in sleep after and before taking melatonin is equal to 0.21 hours. Therefore, melatonin increases sleep.

8. Complete problem 7 correctly.