Introductory Statistics – Day 25

Paired Means Test

4.4 – Paired Tests Find the data set *textbooks.xlsx* on the Moodle page. This data set is from OpenIntro Stats. In this data set we have 73 textbooks that are sold in the UCLA bookstore with their course abbreviations, course numbers, ISBN numbers, UCLA bookstore price, and the price of the same book if purchased on Amazon.com.

- Discuss: What do you notice? What do you wonder?
- Write a research question for this data set.
- What would your null and alternative hypotheses be?
- Would your hypothesis test require a t or z distribution? Why?

1. Discuss which student-generated questions are appropriate.
2. Perform the most appropriate hypothesis test and write the conclusion in plain language.
3. Create a 95% confidence interval for the average price difference between books at the UCLA bookstore and books on Amazon.com.

True / False: Air quality measurements were collected in a random sample of 25 country capitals in 2013, and then again in the same cities in 2014. We would like to use these data to compare air quality between the two years.

(a) We should use a t test.

(b) We should use a p test.

(c) We should use a z test.

(d) We cannot use a statistical test on this data.

True / False: Air quality measurements were collected in a random sample of 25 country capitals in 2013, and then again in the same cities in 2014. We would like to use these data to compare air quality between the two years.

(a) We should use a one-tailed paired t test.

(b) We should use a one-tailed non-paired t test.

(c) We should use a two-tailed paired t test.

(d) We should use a two-tailed non-paired t test.

True / False: In a paired analysis we first take the difference of each pair of observations, and then we do inference on these differences.
(a) True and I am very confident
(b) True and I am not very confident
(c) False and I am very confident
(d) False and I am not very confident

True / False: Two data sets of different sizes cannot be analyzed as paired data.
(a) True and I am very confident
(b) True and I am not very confident
(c) False and I am very confident
(d) False and I am not very confident

True / False: Each observation in one data set has a natural correspondence with exactly one observation from the other data set in a paired t test.
(a) True and I am very confident
(b) True and I am not very confident
(c) False and I am very confident
(d) False and I am not very confident

True / False: Each observation in one data set has a natural correspondence with the mean from the other data set in a paired t test.
(a) True and I am very confident
(b) True and I am not very confident
(c) False and I am very confident
(d) False and I am not very confident

True / False: Each observation is subtracted from the average of the other data set's observations in a paired t test.
(a) True and I am very confident
(b) True and I am not very confident
(c) False and I am very confident
(d) False and I am not very confident

Is there strong evidence of climate change? Let's consider a small scale example, comparing how temperatures have changed in the US from 1968 to 2008. The daily high temperatures on Jan. 1 was collected in 1968 and 2008 for 100 randomly selected location in the continental US. Then the difference between the two readings (2008 temp minus 1968 temp) was calculated for each of the 100 different locations. The average of these 100 values was 1.1 degrees with a standard deviation of 4.9 degrees.

1. Is this a case for a paired t test?
2. Write appropriate hypotheses.
3. Conduct your hypothesis test. Label all important pieces in your Excel sheet.
4. What do you conclude?

Follow-up. Another researcher attempts to duplicate this study using 40 randomly selected US cities. Conveniently, they happen to get the same average difference of 1.1 degrees with a standard deviation of 4.9 degrees.

1. Explain why this second p-value will not be significant by explaining what changes in the formulas used to calculate the p-value.

2. Explain why two studies with the same standard deviation and the same average difference could give you different conclusions.

3. If both of these studies and p-values were presented to you, what would you conclude about evidence for climate change?