



Note # 9: Statistical Inference for Numerical Variables

Problem 1. A researcher is interested in knowing the average amount of time individuals in Cincinnati, Ohio spend commuting. He takes a random sample of 100 Cincinnati residents and asks them what their typical commute time is. He finds the average length of time for the residents in his sample is 32 minutes, with a standard deviation of 12 minutes. Create a 98% confidence interval for the average commute time of all Cincinnati residents.

- a. Construct the 98% confidence interval.
- b. Interpret your confidence interval from part a.



Problem 2. A researcher is interested in knowing the average amount of time individuals in Cincinnati, Ohio spend commuting. He takes a random sample of 100 Cincinnati residents and asks them what their typical commute time is. He finds the average length of time for the residents in his sample is 32 minutes, with a standard deviation of 12 minutes. The researcher believes the average commute time of all Cincinnati residents is greater than 30 minutes. Conduct a hypothesis test at the 0.02 significance level to test this.

- a. What are the hypotheses?
- b. What is the significance level?
- c. What is the value of the test statistic?
- d. What is the p-value?
- e. What is the correct decision?
- f. What is the appropriate conclusion/interpretation?
- g. Does the hypothesis test agree with the confidence interval from question 1?



Problem 3. The Human Toxome Project (HTP) is working to understand the scope of industrial pollution in the human body. Industrial chemicals may enter the body through pollution or as ingredients in consumer products. In October 2008, scientists at HTP tested cord blood samples for 20 newborn infants in the United States. The cord blood of the "In utero/newborn" group was tested for 430 industrial compounds, pollutants, and other chemicals, including chemicals linked to brain and nervous system toxicity, immune system toxicity, reproductive toxicity, and fertility problems. The researchers recorded the number of these targeted chemicals that were found in each infant's cord blood. The average number of targeted chemicals found was 127.45, with a standard deviation of 25.965. Create a 90% confidence interval for the average number of targeted chemicals in all infants' cord blood.

- a. Construct the 90% confidence interval.
- b. Interpret your confidence interval from part a.



Problem 4. The Human Toxome Project (HTP) is working to understand the scope of industrial pollution in the human body. Industrial chemicals may enter the body through pollution or as ingredients in consumer products. In October 2008, scientists at HTP tested cord blood samples for 20 newborn infants in the United States. The cord blood of the "In utero/newborn" group was tested for 430 industrial compounds, pollutants, and other chemicals, including chemicals linked to brain and nervous system toxicity, immune system toxicity, reproductive toxicity, and fertility problems. The researchers recorded the number of these targeted chemicals that were found in each infant's cord blood. The average number of targeted chemicals found was 127.45, with a standard deviation of 25.965. In 2000, the average number of targeted chemicals found in infant's cord blood was 120. Conduct a hypothesis test at the 0.10 significance level to test if the average number has changed.

- a. What are the hypotheses?
- b. What is the significance level?
- c. What is the value of the test statistic?
- d. What is the p-value?
- e. What is the correct decision?
- f. What is the appropriate conclusion/interpretation?
- g. Does the hypothesis test agree with the confidence interval from question 3?
- h. How would the hypothesis test have changed if we wanted to see if the number of targeted chemicals had increased, instead of seeing if it had changed?



Problem 5. Suppose we want to determine if there is a difference between the price of a textbook on Amazon and the price of the same textbook at the UCLA bookstore. We sampled 201 UCLA courses. Of those, 68 required books could be found on Amazon. For each of these 68 books, we recorded the price of the book at the UCLA bookstore, the price of the book on Amazon, and the difference between the two (UCLA bookstore minus Amazon). The average difference in our sample was \$3.58, with a standard deviation of \$13.42. Create a 96% confidence interval for the average difference between the UCLA bookstore price and the Amazon price.

- a. Construct the 96% confidence interval.
- b. Interpret your confidence interval from part a.



Problem 6. Suppose we want to determine if there is a difference between the price of a textbook on Amazon and the price of the same textbook at the UCLA bookstore. We sampled 201 UCLA courses. Of those, 68 required books could be found on Amazon. For each of these 68 books, we recorded the price of the book at the UCLA bookstore, the price of the book on Amazon, and the difference between the two (UCLA bookstore minus Amazon). The average difference in our sample was \$3.58, with a standard deviation of \$13.42. Conduct a hypothesis test at the 0.04 significance level to test this.

- a. What are the hypotheses?
- b. What is the significance level?
- c. What is the value of the test statistic?
- d. What is the p-value?
- e. What is the correct decision?
- f. What is the appropriate conclusion/interpretation?
- g. Does the hypothesis test agree with the confidence interval from question 5?



Problem 7. A study was conducted to investigate the effectiveness of hypnotism in reducing pain. Eight randomly selected individuals were asked to rank their pain both before and after being hypnotized. A lower score indicates less pain. The average difference (after-before) for the sample was -3.13 , with a standard deviation of 2.91 . Create a 90% confidence interval for the average difference in pain levels.

- a. Construct the 90% confidence interval.
- b. Interpret your confidence interval from part a.



Problem 8. A study was conducted to investigate the effectiveness of hypnotism in reducing pain. Eight randomly selected individuals were asked to rank their pain both before and after being hypnotized. A lower score indicates less pain. The average difference (after-before) for the sample was -3.13 , with a standard deviation of 2.91 . Conduct a hypothesis test at the 0.04 significance level to test if hypnotism changes the pain level.

- a. What are the hypotheses?
- b. What is the significance level?
- c. What is the value of the test statistic?
- d. What is the p-value?
- e. What is the correct decision?
- f. What is the appropriate conclusion/interpretation?
- g. Does the hypothesis test agree with the confidence interval from question 7?
- h. How would the hypothesis test have changed if we wanted to determine if hypnotism reduced pain (instead of if it changes pain levels)?