

Lesson 13: Inference for Two Means; Independent Samples

Homework

Instructions: You are encouraged to collaborate with other students on the homework, but it is important that you do your own work. Before working with someone else on the assignment, you should attempt each problem on your own.

1. In your own words, explain the difference between dependent and independent samples.

Researcher Neil Melancon studied the effect of including a company's URL (web page location) in a printed advertisement. He randomly assigned participants to one of two groups: a treatment group and a control group. He showed each participant an ad for Levi's (a clothing manufacturer). The ad shown to subjects in the treatment group included the URL of the Levi's company. The ad shown to subjects in the control group did not include the URL. Subjects rated their desire to purchase a Levi's brand product on a scale of 1 to 5, where 5 indicates the greatest desire. Use this information to answer the following question.

2. Which hypothesis test in the list above would be most appropriate for this study?
 - a. Independent sampling. How the researchers select the participants for the treatment group is independent of how they select the participants for the control group.
 - b. Independent sampling. Those in the treatment group don't know who are in the control group, so they must have been chosen independently of one another.
 - c. Dependent sampling. Participants for each group are drawn from the same population, meaning their answers will be distributed the same.
 - d. Dependent sampling. Who is chosen to participate in the treatment group depends on who is chosen to participate in the control group.

The data set ToyotaPrius2005 gives the advertised selling prices for a random sample of used 2005 Toyota Prius cars that were listed for sale on cars.com. You want to determine which variables affect the listing prices of cars. Two of the factors that you will explore are the type of seller listing the car and whether or not the car is certified. Use this information to answer questions 3 through 20.

The first variable you will look at is the type of seller who is listing the car. The two types of sellers are "Dealer" and "Private". You will create a confidence interval for the difference in the mean listing price between dealers and private sellers.

3. Make an appropriate graph comparing the listing prices of cars sold by a dealer versus cars sold by a private seller.
4. Based on your graph in Question 3, does there appear to be a difference in the mean listing prices of cars sold by a dealer compared to cars sold by a private seller?
5. Are the requirements for creating a confidence interval satisfied?
6. Create and interpret a 95% confidence interval for the difference in the mean listing price of cars that are sold by a dealer versus cars that are sold by a private seller. Report your confidence interval to the nearest dollar amount.

7. Does the confidence interval you computed in Question 6 contain the value zero? If so, what does this suggest about the mean prices of cars that are listed for sale by a dealer versus those listed for sale by a private individual?

The next variable you will look at is whether or not the car is certified. The variable “IsCertified” contains the value “Yes” if the seller has inspected the car and offers an extended warranty and “No” otherwise. You will create a confidence interval for the difference in the mean listing prices between certified and non-certified cars.

8. Make an appropriate graph comparing the listing prices of cars that are certified compared to cars that are not certified.
9. Based on your graph in Question 8, does there appear to be a difference in the mean listing prices of certified and non-certified cars?
10. Are the requirements for creating a confidence interval satisfied?
11. Create and interpret a 95% confidence interval for the difference in the mean listing price of cars that are certified compared to cars that are not certified.
12. Does the confidence interval you computed in Question 11 contain the value zero? If so, what does this suggest about the mean prices of cars that are certified versus those that are not certified?

You suspect that cars that are certified have a higher mean listing price than non-certified cars. You will conduct a hypothesis test to see if this is the case. Use a level of significance of $\alpha = 0.05$.

13. State the null and alternative hypotheses.
14. Make an appropriate graph of the data.
15. Verify that the requirements to conduct a hypothesis test are met.
16. Give relevant summary statistics (mean, standard deviation, sample size).
17. Report the test statistic. Give its degrees of freedom.
18. State the P-value.
19. Give the decision rule for this test.
20. Present your conclusion in an English sentence relating the results to the original context.

1033 patients who did not have a regular health care provider and who had an acute medical condition were randomly assigned to either a nurse practitioner or a physician. The patients were given an examination at the beginning of the study and again six months later. The researchers wanted to know if there was a difference in the level of satisfaction of the patients who were treated by a nurse practitioner compared to those treated by a doctor. The data set PatientSatisfactionDoctorVsNurse-Stacked gives data representative of the satisfaction ratings reported by the patients. The satisfaction scores are based on a series of questions using a 5-point scale, where 5 is the most positive response. Conduct a hypothesis test to determine if there is a difference in the mean level of satisfaction for the patients in the two groups. Use a level of significance of $\alpha = 0.05$. Use this information to answer questions 21 through 25.

21. State the null and alternative hypotheses.
22. Report the test statistic. Give its degrees of freedom.
23. State the P-value.

24. Give the decision rule for this test.

25. Present your conclusion in an English sentence relating the results to the original context.

```
## 'geom_smooth()' using formula 'y ~ x'  
## 'geom_smooth()' using formula 'y ~ x'
```