

Problem set: Independent and Dependent t-tests

Pablo E. Gutiérrez-Fonseca

2023-08-09

For our next foray into inferential statistics ***we will be practicing both independent and dependent t-tests***. As usual, submission is in blackboard. However, We encourage you to complete the work outside of blackboard and recording all answers either on paper or in word (in case BB decides to balk). We also encourage you to work in groups to discuss these questions. However, you must ultimately do the work on your own and turn in your answers and interpretations independently. HAVE FUN!

Bird_Migration.csv contains data on the species, gender, pre-migration weight and post-migration weight for 28 individual birds measured before and after their annual migration. You are being set loose on this data with little guidance (let's get ready for the next exam). But we will walk through the steps for inferential statistics:

Part I.

Consider the data on bird migration in **Bird_Migration.csv**. We have 28 individuals that were captured and banded in Pennsylvania just before the start of their annual migration. We are interested to see if there is a significant difference in size between males and females. Complete each of the following steps to address this question:

We will start you off with the following:

A statement of research hypothesis:

There is a difference in the pre-weights of male and female songbirds.

Set the level of risk associated with the null hypothesis (significance level):

use a p-value threshold of 0.05

1. What type of analysis do you think you should run to address this hypothesis?
2. Analyze data distribution. Is your pre-weight data normally distributed?
3. Test to see if you meet the assumption of equal variance for this test. Based on the results of this test for unequal variance, which test should you use?
4. Run this analysis. What is the test statistic (obtained) value for the test you selected?
5. Based on the nature of the hypothesis (one or two tailed), the 0.05 significance level, and the type of test your ran (parametric, non-parametric, unequal variance), is there a significant difference between Male and Female songbird pre-weights?
6. Write a one paragraph summary.

Part II.

Again consider the data on bird migration in **Bird_Migration.csv**. This time we are interested to see if there is a significant difference in size between the two species instead of gender. Import the data set into R and complete each of the following steps to address this question:

Again, We will start you off with:

A statement of research hypothesis:

There is a difference in the pre-weights of Hylocichla and Progne species.

Set the level of risk associated with the null hypothesis (significance level):

use a p-value threshold of 0.05

And you already tested for normality above in number 1. Use that same result here.

7. Select the appropriate test (i.e. what statistical analysis will you use to test this research hypothesis?) and then conduct this analysis in R.
8. What is the test statistic calculated and reported in the R analysis?
9. Based on the nature of the hypothesis and type of test to run, report the appropriate p-value for this test. (report this to 4 decimal places)
10. Based on this p-value do you accept the null hypothesis?
11. Write a one paragraph summary of your statistical results and your conclusion based on these results. If your results are significant, don't forget to describe the nature of any differences and include a discussion of how meaningful results are.

Part III.

Same data set, new question. Consider the data on bird migration in **Bird Migration Data.csv**. Birds were banded, sexed and weighed in Pennsylvania at the start of the migratory season. One month later, 14 of each species were re-captured in the Yucatan. We want to test to see if migration has a negative impact on their vitality using the amount of weight lost as a proxy. Complete each of the following steps to address this question:

Again, I will start you off with:

A statement of research hypothesis:

Songbirds lose body mass during migration (quantified as a comparison of pre-and post-weights).

Set the level of risk associated with the null hypothesis (significance level):

use a p-value threshold of 0.05.

12. Analyze data distribution to meet assumptions. Report the p-value of your goodness of fit test (be sure that you are testing the correct data for normality based on the test you need to run to test the stated hypothesis).
13. Select the appropriate test (i.e. what statistical analysis will you use to test this research hypothesis?).
14. What is the appropriate test statistic (obtained) value for the test you selected.
15. Report the appropriate p-value considering the type of test you are running (left, right or two tailed).
16. Do you accept the null hypothesis?
17. What are your conclusions (i.e. what does this mean in terms of our original research hypothesis)? Write a one paragraph summary of your statistical analysis and conclusions.

Part IV.

Import your **Bird_Migration.csv** into R. Now we will do this test again, but this time, we want to see if changes in body mass during migration is differentially significant in both males and females. Basically, we will do the same test you chose above, but separately on males and females:

Again, I will start you off with:

A statement of our TWO research hypotheses:

Females lose body mass during migration (measured as change in body weight). AND

Males lose body mass during migration (measured as change in body weight).

Set the level of risk associated with the null hypothesis (significance level):

use a p-value threshold of 0.05.

And you already tested for normality above. Use that same result here.

18. What are your conclusions (i.e. what does this mean in terms of our original research hypothesis)?
Write a one paragraph summary reporting the results of both tests and what this means about males and females during migration.