Quantitative Methods

Computer Lab 2 - Statistical Inference

Work in small groups of 2 to 3 students to complete the exercises below. Feel free to ask each other for help and discuss as much as you want during the computer lab. Remember to reflect on what you are doing and what the output tells you. It's better to take your time than to rush through the exercises. If you get stuck, you are free to ask the teacher for help. You can use `JAMOVI` or `R` to complete all the exercises.

About the data

The East African long-eared elephant shrew (Elephantulus rufescens) is a small mammal found in Africa. The name of the animal is rather misleading:

> "Although the common name of Elephantulus rufescens is 'elephant shrew', it is not a shrew nor is it related to elephants. It gets its name from its long mobile snout, which it can move around rather like an elephant's trunk. It uses its snout to search for worms, ants, termites and other inverterbrates. Its legs are long and thin; its hindlimbs are longer than its forelimbs, allowing it to jump and hop. It has a long tail, and large eyes and ears. It also has long, soft fur; the upper parts are sandy brown, buffy gray or buffy orange and the underparts are white, or grayish" [1]

In this computer lab, you will investigate a (fictional but reasonable) dataset regarding this animal. The dataset consists of 75 observations and contains the following four variables:

- **Sex**: The sex of the elephant shrew. (Male/Female)
- **BodyMass**: The body mass of the elephant shrew. (Grams)
- **Captured**: Which time of the day the elephant shrew was captured. (Night/Day)
- **EarLength**: The length of the elephant shrew's right ear. (Millimeters)

The data file, `ElephantShrew.xlsx`, can be found in the folder "Computer Labs" in Studium.

[1] `https://animaldiversity.org/accounts/elephantulus_rufescens/`, accessed 2 Feb 2024.

Exercises

You can use `JAMOVI` or `R` to carry out all the exercises.

1. Download the `ElephantShrew.xlsx` file from Studium and store it in an easily accessible place on the computer.

2. Confirm that the data has been correctly imported.

3. What is the measurement level for the four variables in the dataset, and why does it matter what the measurement levels are?

4. You will now create a 95% confidence interval for the average body mass of elephant shrews in the population, both manually and with the help of a built-in function in `JAMOVI` or `R`. We start by manually creating it. The following steps will help you create it:

   (a) Is the population standard deviation known for the body mass of elephant shrews?

   (b) Do we know if body mass is normally distributed in the population or not?

   (c) What is the sample size here?

   (d) What does the central limit theorem tell us about the sampling distribution of the mean?

   (e) Using the reasoning from the points above, should the confidence interval be based on a normal distribution or a t-distribution?

   (f) If we base the confidence interval on the t-distribution, how many degrees of freedom does it have?

   (g) To calculate the confidence interval, we need the sample mean, the sample standard deviation, the sample size, and the $t_{\alpha/2}$ value. Calculate all those things.

   (h) Use the numbers calculated in the point above to construct the confidence interval. If you did everything correctly, the 95% confidence interval should be approximately (47.98; 51.06) grams.

5. Interpret the confidence interval you calculated.

6. Would the confidence interval be wider or narrower if you increased the confidence level? Make sure that you understand why.

7. Create a 99% confidence interval for the population average body mass of elephant shrew. Did your results correspond to what you concluded in the previous question?

8. Perform a one-sample t-test to investigate if the average body mass of an elephant shrew is greater than 48 in the population. Do the following:

   (a) What are the null and alternative hypotheses you are testing? Write them down and make sure that you define all notation used and that you use the correct choice of $=$, $>$, and $<$ in the alternative hypothesis.

   (b) Choose a suitable significance level. You may pick any reasonable level here.

   (c) Argue why the one-sample t-test is reasonable here.

   (d) Perform the test by running `JAMOVI`. Perform one-sided test by set "less" or "greater" on the alternative hypotheses.

   (e) What was the p-value of the test?

   (f) Can you reject the null hypothesis at your chosen significance level?

   (g) Interpret the results of the test.

9. Perform a two-sample t-test in `JAMOVI`. You will now investigate if there is a difference in average body mass of male and female elephant shrews.

10. Finally, perform a chi-square test to see if there is an association between the sex of the elephant shrew and if they were captured during the day or the night. Again, make sure that you do the steps of hypothesis testing.