Prof. Ahlstrom
MATH 2310 Sec. 01

Probability and Statistics for the Sciences and Engineering (MATH 2310)
Lab 6: Confidence intervals
20 points

Credit to Dr. McLean Sloughter and Dr. Galen Egan for designing significant portions of this assignment

---

In this lab, you will write code to calculate summary statistics and confidence intervals.

For this lab, we will be working with the same snow geese data that we used in our first lab assignment. The columns are not labeled in this data set. The first variable is just an index of the trial number. The second variable is a description of the type of diet fed to the goose. The third variable is their percent weight change after being allowed to feed for 2.5 hours. The fourth variable is their digestion efficiency. The fifth variable is the amount of acid-detergent fiber in their digestive tract. This file (`Snowgeese.xls`) can be accessed on Canvas. **Submit code and outputs** for each of the following questions.

1. **Weight change**

   (a) Calculate the mean and standard deviation for weight change for geese fed a plant diet. Then, calculate the mean and standard deviation for geese fed a chow diet.

   (b) After separating the sample data into these two groups, we see that neither of the groups has a large enough sample size ($n > 40$) to justify using the large-sample confidence interval formula for μ. *We will calculate this confidence interval anyways*, though, with the understanding that its usage is not completely justified based on the size of our sample data sets.

   Calculate the **99% large-sample confidence interval** for μ for the plant diet group. Also calculate the 99% large-sample confidence interval for the chow diet group.

   (c) Based on these results (disregarding the sample size caveat for now), is there evidence that geese fed a plant diet will, on average, either gain or lose weight? What about for geese fed a chow diet?

   (d) Because we're actually dealing with small sample sizes for these two groups, the large-sample confidence interval calculation is probably not the best approach. If we make the base assumption that the underlying population is normally distributed, though, we can compute a confidence interval using the Student's t distribution instead.

Using your results from part (a), calculate a **99% confidence interval** for the mean weight change for geese fed a plant diet (using the Student's t distribution), and a 99% confidence interval for the mean weight change for geese fed a chow diet.

(e) Based on these results (and given the assumption that the underlying population is normally distributed), is there evidence that geese fed a plant diet will, on average, either gain or lose weight? What about for geese fed a chow diet?

2. **Digestion efficiency**

(a) Calculate the means and standard deviations for digestion efficiency for geese fed a plant diet, and for geese fed a chow diet.

(b) We can use this result to calculate a 99% confidence interval for the **difference** in mean digestion efficiencies between these two groups. Again, we will need to assume that the underlying populations for each of the two groups are normally distributed. We will also need to assume independence between the two distributions. Given these assumptions, we can calculate a confidence interval for the **difference** between the two distributions using the following formula:

$$\overline{x}_1 - \overline{x}_2 \pm \left(t_{\alpha/2,\, d_f}\right) \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}},$$

$$d_f = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2},$$

where Sample 1 has a mean of $\overline{x}_1$, a sample standard deviation of s_1, and a sample size of n_1; and Sample 2 has a mean of $\overline{x}_2$, a sample standard deviation of s_2, and a sample size of n_2.

Write code to implement this formula and calculate the confidence interval for this difference.

(c) Based on your confidence interval, what can you say about how the average digestion efficiency compares between the two diets?