

# Practice Session 8

## Introduction to Type 1 and Type 2 Errors

In hypothesis testing, we either reject or fail to reject the null hypothesis. However, we can sometimes make the wrong decision. A Type 1 error results when we reject the null hypothesis, but the null hypothesis is actually true. A Type 2 error occurs when we fail to reject the null hypothesis, but the null hypothesis is actually false. One important fact to note is that the significance level that we use for p-values (typically 0.05) is the likelihood that we make a Type 1 error.

## Type 1, Type 2, or Neither?

For each of the following statements, decide whether a Type 1 or a Type 2 Error was made. If no error was made, indicate that no error was committed.

a.) Shiba ran a hypothesis test with the following hypotheses:

- $H_0 : \mu = 50$
- $H_A : \mu > 50$

She rejected the null hypothesis; the true mean is actually 50.

b.) Joe wanted to test if the mean bag weight of Lays Potato chips is less than 5 oz.

- $H_0 : \mu = 5$
- $H_A : \mu < 5$

He failed to reject the null; the true mean is actually 5 oz.

c.) A test for an infectious disease gives a positive result if a patient is believed to have the disease, and a negative result otherwise. However, the test can sometimes be wrong. Label the following statements as a Type 1 error, Type 2 error, or neither. Assume the null hypothesis is that the patient doesn't have the disease.

- A positive test result; the patient actually has the disease
- A positive test result; the patient doesn't actually have the disease
- A negative test result; the patient actually has the disease
- A negative test result; the patient doesn't actually have the disease

## Single Sample T Test

According to a local statistics instructor, the average score on his first exam of the semester (`exam1`) is historically around 80%. Run a one sample T test to see if this year's students performed significantly better than the historic average. Use the formulas learned in class to perform the test, and do the calculations using R. The data is available in the `StatsGrades` data set from the `Lock5Data` library.

## Two Sample T Test

The same local statistics instructor believes that students perform the same on the first exam (`exam1`) and the second exam (`exam2`). Run a two sample T test to see if this year's students performed differently on the two exams. Use the formulas learned in class to perform the test, and do the calculations using R. The data is available in the `StatsGrades` data set from the `Lock5Data` library.

## Paired T Test

Rerun the two sample T test above, but this time do a paired T test. Note any differences in the results. What do you conclude about student performance on the two exams?

## T vs Normal

Joe was performing a hypothesis test, and mistakenly used the normal distribution instead of the T distribution. Due to the differences between the distributions, is he more likely to make a Type 1 error, or a Type 2 error? Explain.

Now suppose the Joe had meant to use the normal distribution but accidentally used the T distribution. Is he more likely to make a Type 1 error, or a Type error? Explain.