Stat 218 - Day 16
Hypothesis Testing: *t*-test

Recall from yesterday that we form a confidence interval for the difference in population means μ_A - μ_B by: $\left(\bar{y}_A - \bar{y}_B\right) \pm t^* \sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}$, where the degrees of freedom is the smaller of n_A - 1 and $n_B - 1$.

Also recall that when drawing conclusions from a study, one of the most important considerations is the manner in which the data were collected. We can draw a cause-and-effect conclusion from a **randomized comparative experiment** but not from an **observational study**.

Statistical inference means to *infer* something about a population parameter based on a sample statistic. There are two major types of statistical inference:

- A **confidence interval** estimates the value of a population parameter with an interval of plausible values.
- A **hypothesis test** assesses the evidence provided by the data against a particular conjecture/hypothesis about the population parameter(s).

The specific hypothesis testing procedure that we will study this week is called a two-sample *t*-test for comparing means. Before we get to its details, though, we will start with a silly example that illustrates the reasoning process behind hypothesis testing.

Example: Dice Rolling
Describe the reasoning process that went through your mind as the dice rolls were called out.

Four essential components of every hypothesis test are:

- **Null hypothesis**, denoted by H_0, typically a statement of "no effect" or "no difference"
    - In this case: H_0: $\mu_A = \mu_B$
- **Alternative hypothesis**, denoted by H_A, also called the research hypothesis
    - In this case: H_A: $\mu_A \neq \mu_B$
- **Test statistic**, which is a measure of how consistent the observed sample data are with the null hypothesis
    - In this case: $t_s = \frac{\left(\bar{y}_A - \bar{y}_B\right)}{\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}}$

- ***P*-value**, which is the probability of obtaining such an extreme test statistic if the null hypothesis were true
    - The smaller the P-value, the stronger the evidence against the null hypothesis
    - *P*-values less than .05 are generally considered to be moderately convincing evidence against H_0; *P*-values less than .01 are generally considered to be strongly convincing evidence against H_0.
    - In this case: *P*-value = area in tails of *t*-distribution beyond $-t_s$ and $+ t_s$; use df = min (n_A-1, n_B-1)

Two optional components are:

- **Significance level**, denoted by α, which is a threshold for determining how small the *P*-value must be to provide convincing evidence against H_0.
- **Test decision**
    - Reject H_0 if *P*-value ≤ α (difference between group means is "statistically significant")
    - Fail to reject H_0 if *P*-value > α (difference is not "statistically significant")

Example: Fish Oil Diet

A group of fourteen men were randomly divided into two groups: one group went on a diet of fish oil and the other on a diet of regular oil for two weeks. The reduction in diastolic blood pressure (in millimeters of mercury) was recorded for each subject (a negative value means that the person experienced an increase in blood pressure). The data turned out as follows:

Fish oil diet:	8, 12, 10, 14, 2, 0, 0	mean: 6.57	std dev: 5.86
Regular oil diet:	-6, 0, 1, 2, -3, -4, 2	mean: -1.14	std dev: 3.18

(a) Is this an observational study or an experiment? Explain briefly.

(b) Identify the explanatory and response variables.

(c) State the null and alternative hypotheses, in symbols and in words.

(d) Calculate the value of the test statistic.

(e) Determine the *P*-value of the test, as accurately as possible from the *t*-table.

(f) What would the test decision be at the α=.05 significance level? Relate this to the context of this study.

(g) How about at the α=.01 significance level? Relate this to the context of this study.

(h) How would the test statistic and *P*-value have changed if the sample sizes were larger and all else remained the same? Explain, both intuitively and algebraically.

(i) What scope of conclusion can you draw from this study? Explain why.