Equality

In this application exercise, we'll do inference on two population proportions.

Packages

We'll use the **tidyverse** and **tidymodels** packages.

```
library(tidyverse)
library(tidymodels)
```

Data

A September 16-19, 2023, asked North Carolina voters, among other issues, about issues of equality and women's progress. Specifically, one of the questions asked:

> Which of these two statements come closest to your own views—even if neither is exactly right?
>
> - The country has made most of the changes needed to give women equal rights with men.
>
> - The country needs to continue to make changes to give women equal rights to men.

The data from this survey can be found in your `data` folder: `equality.csv`.

Hypotheses

Exercise 1

The two populations of interest in this survey are 18-24 year olds and 25+ year olds. State the hypotheses for evaluating whether there is a discernible difference between the proportions of those who think "The country needs to continue to make changes to give women equal rights to men." (`need more changes`) in the two age groups.

Add response here.

Exploratory analysis

Load the data.

```
equality <- read_csv("https://sta199-s24.github.io/data/equality.csv")
```

Exercise 2

What proportion of 18-24 year olds think "The country needs to continue to make changes to give women equal rights to men"? What proportion of 25+ year olds? Calculate and visualize these proportions.

```
# add code here
```

Exercise 3

Calculate the observed sample statistic, i.e., the difference between the proportions of respondents who think "The country needs to continue to make changes to give women equal rights to men" between the two age groups.

```
# add code here
```

Testing

Exercise 4

What is the parameter of interest?

Add response here.

Exercise 5

Explain how you can set up a simulation for this hypothesis test.

Add response here.

Exercise 6

Conduct the hypothesis test using randomization and visualize and report the p-value.

```
# add code here
```

Exercise 7

What is the conclusion of the hypothesis test?

Add response here.

Exercise 8

Interpret the p-value in the context of the data and the hypotheses.

Add response here.

Estimation

Exercise 9

Estimate the difference in population proportions of 18-24 year old NC voters and 25+ year old NC voters using a 95% bootstrap interval.

```
# add code here
```

Exercise 10

Interpret the confidence interval in context of the data.

Add response here.

Exercise 11

Describe how the simulation scheme for bootstrapping is different than that for the hypothesis test.

Add response here.

Conceptual

Exercise 12

What is p vs. $\hat{p}$ vs. p-value. Explain generically as well as in the context of these data and research question.

Add response here.

Exercise 13

What is bootstrap distribution vs. null distribution? Explain generically as well as in the context of these data and research question.

Add response here.