

Name: _____

LA Initials:

Lab 10: t-tests

Learning Objectives

Getting started

Before you begin the lab activities, review the steps for setting up a project for a lab activity:

- Weekly Lab Setup

Then, complete the following steps to make sure you are fully set up.

1. **Get the Lab Worksheet.**

Pick up a physical copy of the lab worksheet, or print one if you are working outside of class.
Download Lab Worksheet (PDF, if needed)

2. **Open Posit Cloud and create Lab 10 in Your Workspace.**

3. Create a new R script and save it as **lab-10-script.R**.

4. Add the following code to your R script and save the script. Upon save, you should see a message in the source prompting you to install the package. Click “Install”. Then run the line of code.

```
# Load packages -----  
  
library(tidyverse)  
library(haven)
```

i Checkpoint

At this point, you should have:

- These instructions open in a web browser.
- Your Lab 10 project open in Posit Cloud in another browser window.
- The required packages installed and loaded without errors
- The Lab 10 worksheet in front of you.

Do not continue until all of the above steps are working correctly.

Overview

In this lab, you will analyze the relationship between two categorical variables using **contingency analysis**. A contingency table summarizes the joint distribution of two categorical variables and allows us to estimate measures of association such as **relative risk (RR)** and **odds ratio (OR)**. These measures describe the magnitude and direction of an association between two binary variables. We will also conduct a **chi-square test of independence** to determine whether the observed association is statistically significant.

You will use data from the National Health and Nutrition Examination Survey (NHANES), a large, nationally representative health survey conducted in the United States. For this lab, you will examine the relationship between sleep duration and depressive symptoms among adults.

Prerequisites

In your R script, load the following packages:

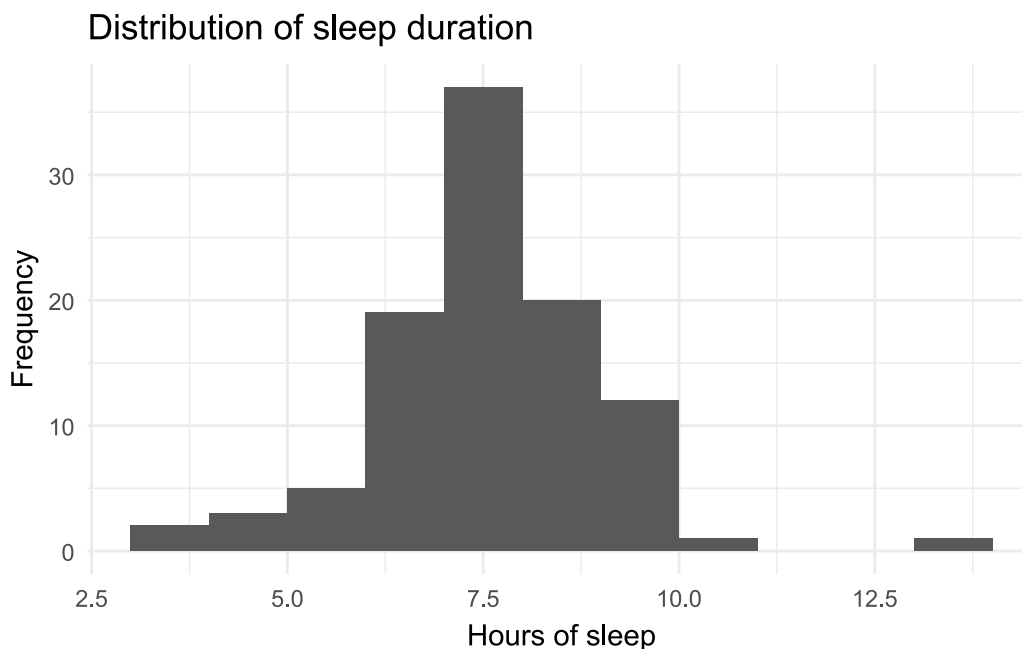
Sleep duration: one-sample test

In this activity, you will:

- read in NHANES sleep data and keep the first 100 observations
- visualize the data using a histogram
- assess the distribution to decide between a t-test and a Wilcoxon test
- test whether mean (or median) sleep duration differs from 8 hours
- interpret the result in plain language

Code

```
slq_data <-  
  read_xpt("https://www.cdc.gov/Nchs/Data/Nhanes/Public/2021/DataFiles/SLQ_L.xpt") |>  
  drop_na(SLD012) |>  
  slice_head(n = 100)  
  
slq_data |>  
  ggplot(aes(x = SLD012)) +  
  geom_histogram(binwidth = 1, boundary = 0) +  
  labs(  
    x = "Hours of sleep",  
    y = "Frequency",  
    title = "Distribution of sleep duration"  
  ) +  
  theme_minimal()
```



Bison body mass: two-sample comparison

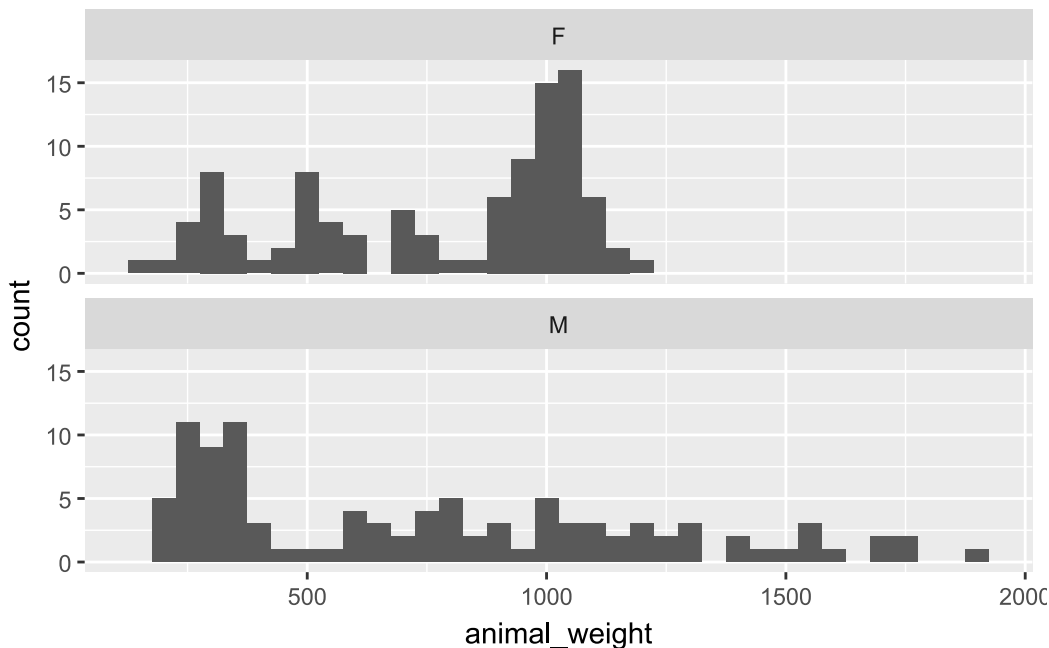
In this activity, you will:

- read in the bison dataset and keep 100 observations for each sex
- visualize body mass distributions for males and females using histograms
- assess the distributions to decide between a t-test and a Mann–Whitney test
- compare body mass between sexes using either `t.test()` or `wilcox.test()`
- interpret the results in terms of differences between groups

Use the following code:

```
bison_data <-
  knz_bison |>
  drop_na(animal_sex, animal_weight) |>
  slice_head(n = 100, by = animal_sex)

bison_data |>
  ggplot(aes(x = animal_weight)) +
  geom_histogram(binwidth = 50) +
  facet_wrap(~animal_sex, ncol = 1)
```



In this activity, you will work with a dataset of shoulder strength measurements.

- each row represents an individual
- variables include **Left** and **Right** shoulder strength measurements
- individuals are classified as **Swimmer** or **Control (non-swimmer)**
- the goal is to compare strength **within individuals**, not between individuals

You will:

- calculate **strong arm** and **weak arm** for each individual
- note that comparing **left vs right directly is misleading** because people are left- or right-handed
- instead, compute **strong arm - weak arm** to create a meaningful paired comparison
- perform **paired t-tests** comparing strong vs weak arm strength
- run separate tests for: - swimmers - non-swimmers
- optionally, test **left vs right in non-swimmers** to confirm the difference is near zero

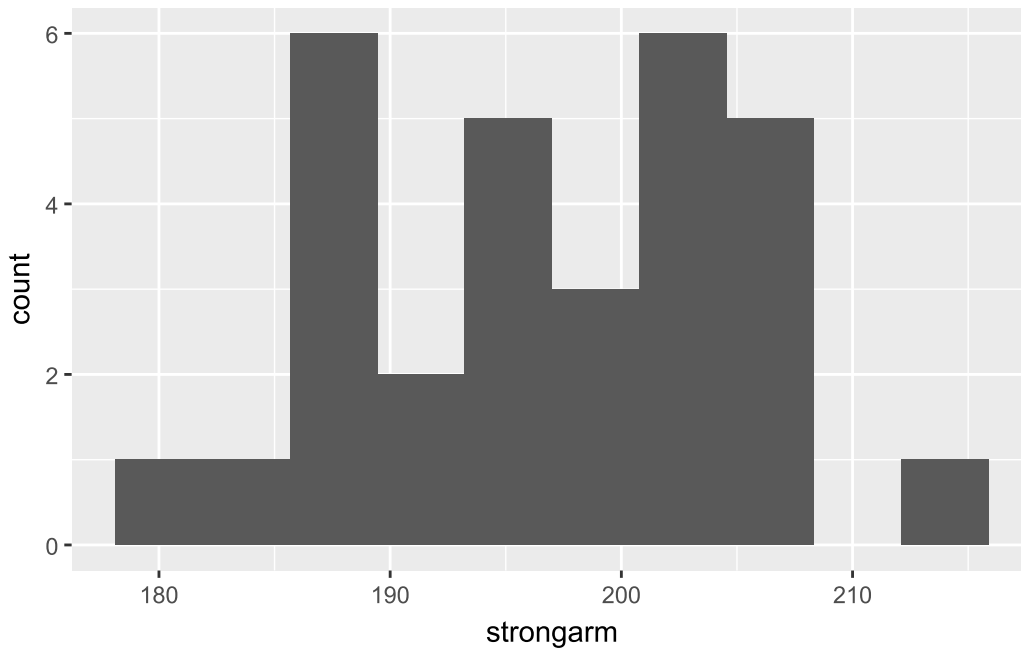
- interpret results in terms of within-individual differences

Code

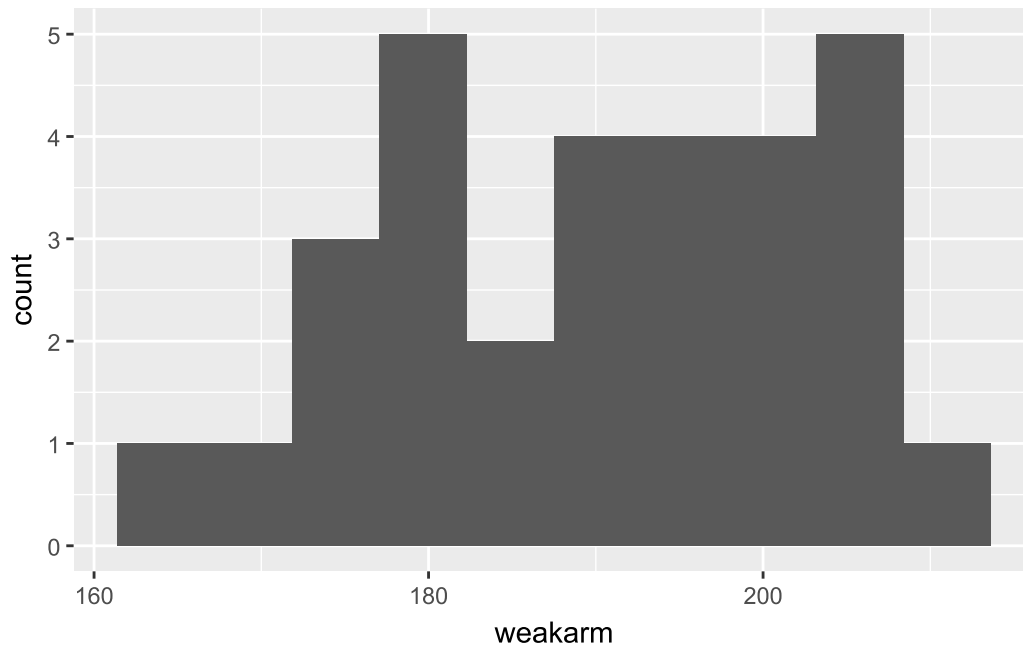
```
library(PairedData)
data("Shoulder")

shoulder_data <-
  Shoulder |>
  as_tibble() |>
  rowwise() |>
  mutate(
    strongarm = max(Left, Right),
    weakarm = min(Left, Right)
  ) |>
  ungroup()

shoulder_data |>
  ggplot(aes(x = strongarm)) +
  geom_histogram(bins = 10)
```



```
shoulder_data |>
  ggplot(aes(x = weakarm)) +
  geom_histogram(bins = 10)
```



```
swimmers <-
  shoulder_data |>
  filter(Group == "Swimmer") |>
  print()
```

```
# A tibble: 15 × 6
  Subject Group   Left Right strongarm weakarm
  <chr>   <fct>   <dbl> <dbl>   <dbl>   <dbl>
1 S1     Swimmer  193   192     193     192
2 S2     Swimmer  208   207     208     207
3 S3     Swimmer  198   198     198     198
4 S4     Swimmer  201   203     203     201
5 S5     Swimmer  196   194     196     194
6 S6     Swimmer  196   193     196     193
7 S7     Swimmer  211   214     214     211
8 S8     Swimmer  206   207     207     206
9 S9     Swimmer  197   195     197     195
10 S10    Swimmer  204   198     204     198
11 S11    Swimmer  197   198     198     197
12 S12    Swimmer  205   206     206     205
13 S13    Swimmer  207   200     207     200
14 S14    Swimmer  204   204     204     204
15 S15    Swimmer  204   205     205     204
```

```
nonswimmers <-
  shoulder_data |>
  filter(Group == "Control") |>
  print()
```

```
# A tibble: 15 × 6
  Subject Group   Left Right strongarm weakarm
  <chr>   <fct>   <dbl> <dbl>   <dbl>   <dbl>
1 S16     Control  184   197     197     184
2 S17     Control  172   187     187     172
3 S18     Control  178   180     180     178
4 S19     Control  186   175     186     175
5 S20     Control  194   192     194     192
6 S21     Control  188   189     189     188
7 S22     Control  164   185     185     164
8 S23     Control  202   168     202     168
9 S24     Control  182   204     204     182
10 S25     Control  186   181     186     181
11 S26     Control  188   182     188     182
12 S27     Control  186   172     186     172
13 S28     Control  192   183     192     183
14 S29     Control  204   189     204     189
15 S30     Control  178   198     198     178
```

Wrap-up and Submission

1. Make sure your script is saved in your project on Posit Cloud.
2. Ask the instructor to explain anything you aren't sure about.
3. **Show your handout to a Learning Assistant for a completion grade before you leave lab.** You may do this as soon as you finish. Keep the handout for yourself.