DATA 505: Lab 4

Basic Statistical Inference

YOUR NAME HERE

2025-09-16

Table of contents

1	Setup & Goals	1
2	The data	2
3	Getting acquainted with the data	2
4	Inference for mean BMI	3
5	Comparing two proportions	3

1 Setup & Goals

Just like in Labs 1 and 2, you will edit a .qmd Quarto script, render the result to a PDF, and submit both on Moodle (if you do not succeed at rendering the PDF, still submit the .qmd script: you may still get full "satisfactory" credit if you have made a reasonable effort).

Goals for Lab 4:

- Use the t.test() function to carry out hypothesis tests and to produce confidence intervals for unknown population means.
- Use the prop.test() function to do hypothesis tests for unknown population proportions.

2 The data

- The data for this lab come from the NHANES package
- $\bullet\,$ This is survey data from the National Health and Nutrition Examination Survey, from 2009-2012

```
# install the MASS package if needed
if(!"NHANES" %in% installed.packages()){
  install.packages("NHANES")
}
# load the NHANES package and verify it is now ready for use
stopifnot(require(NHANES))
```

Loading required package: NHANES

• To load the data from NHANES into your environment, the command is

```
data("NHANES")
```

3 Getting acquainted with the data

• Use the str() function to print a concise summary of the NHANES object

```
# YOUR CODE HERE
```

• Look up the help page for the dataset to understand more about the columns (recall: to access help in R, you can write in the console? followed by the name of the function or object for which you want documentation)

?NHANES

• We will only work with one of the younger age groups. This code will subset the data to restrict attention to the age 18-25 respondents.

```
sample <- NHANES[NHANES$Age >= 18 & NHANES$Age <= 25,]</pre>
```

• Find out how many respondents are left in this subset.

4 Inference for mean BMI

• Produce and interpret a 95% confidence interval for the mean BMI among 18-25 year olds

YOUR CODE HERE

- YOUR INTERPRETATION
- Carry out the following hypothesis test about μ , the mean BMI among 18-25 year olds, and interpret the results.
 - $-H_0: \mu = 27$
 - $H_A : \mu \neq 27$
 - $-\alpha = 0.05$

YOUR CODE HERE

• YOUR INTERPRETATION

5 Comparing two proportions

• Use the table() function to create a two-way frequency table for gender (Gender) versus homeownership (HomeOwn), but follow the pattern here in the lecture example to dichotomize homeownership into owners versus non-owners (HomeOwn == "Own").

YOUR CODE HERE

• Use prop.test() to do a hypothesis test to compare the proportion of homeowners between males and females. Provide a p-value, draw a conclusion, and interpret your results in context.

YOUR CODE HERE

• YOUR INTERPRETATION