Week 3 Lab (Solutions).

Christopher A Oldnall

Package and Data Loading

In this lab, we will use the familiar NHAHES dataset to further practice running hypothesis tests. For this we will initially load in the data and packages necessary. This is the NHANES data set and also tidyverse and BSDA.

Exercise 1

Explore the relationship between education level and home ownership. Are people with different levels of education all equally likely to own their homes? Use the Education and HomeOwn variables from the NHANES dataset.

```
##                 Own Rent Other
## 8th Grade       226  213     9
## 9 - 11th Grade  478  369    30
## High School    1005  468    27
## Some College   1458  731    68
## College Grad   1605  454    31
```

```
##
##  Pearson's Chi-squared test
##
## data:  contingency_table
## X-squared = 221.47, df = 8, p-value < 2.2e-16
```

Here we see that there is evidence to reject the null hypothesis that there is no relationship between the Education and HomeOwn variables ($\chi^2 = 221.470, p < 0.005$). This allows us to infer that there is an association between education and home ownership and indeed conclude that people across different levels of education are not equally likely to become home owners.

Exercise 2

The minimum recommended amount of sleep for adults is 7 hours. Use the appropriate hypothesis test to check whether the average sleep time is significantly different from 7 hours, based on data from NHANES.

```
##
##  One Sample t-test
##
## data:  NHANES$SleepHrsNight
## t = -4.7388, df = 7754, p-value = 2.189e-06
## alternative hypothesis: true mean is not equal to 7
## 95 percent confidence interval:
##  6.897552 6.957509
```

```
## sample estimates:
## mean of x
##  6.927531
```

Using a one sample t-test, we can test whether the true mean of the amount of sleep for adults using the SleepHrsNight variable in the NHANES data set. Here we see that the $\bar{x} = 6.928$ with CI $= [6.896, 6.958]$. The test provides evidence to reject the null that the mean is equal to 7 ($t = -4.739, p < 0.005$). We therefore say there is evidence the average time asleep is different from 7. We cannot infer which way.

Exercise 3

There is evidence that people who are physically active tend to have better mental health. Test this hypothesis using the NHANES data: - Run a test to check whether people who are physically active, also report fewer days when their mental health was not good. - Write an interpretation of the results. Use the PhysActive and DaysMentHlthBad variables from the NHANES dataset.

```
##
##  Welch Two Sample t-test
##
## data:  DaysMentHlthBad by PhysActive
## t = 9.0843, df = 5895.7, p-value < 2.2e-16
## alternative hypothesis: true difference in means between group No and group Yes is greater than 0
## 95 percent confidence interval:
##  1.395667      Inf
## sample estimates:
##  mean in group No mean in group Yes
##          5.089357          3.385049
```

Here we use a two-sample Welch t-test to determine whether there is evidence that people who are physically active tend to have better mental health. This considers whether the means of days with bad mental health variable when grouped by the 'YES' and 'NO' groups in the physically active variable. We may reject the null hypothesis that the difference between the means of the 'NO' and 'YES' groups is less than or equal to 0 ($t = 9.084, p < 0.005$). This means we can conclude that there is evidence that those who are more physically active report fewer days when their mental health is bad.