M440B HW 10

Erick Castillo

3/26/2021

I will once again define a function that I might have to use throughout the assignment:

```
pears_chi <- function(obs, exp){
  return(sum((obs - exp)^2/exp))
}
```

Problem 1: Below is the data provided in the book.

```
##   diab.obs norm.obs tot.gen.obs
## 1       12        4          16
## 2       39       49          88
## 3       51       53         104
```

I will begin this problem by performing a χ^2 test of homogeneity. Below is the output for the expected values, the χ^2 statistic, and the corresponding p-value:

```
##    diab.exp  norm.exp tot.gen.exp
## 1  7.846154  8.153846          16
## 2 43.153846 44.846154          88
## 3 51.000000 53.000000         104

## [1] 5.099788

## [1] 0.02392877
```

The above χ^2 statistic and p-value indicate that there appears to be an association between the genetics of an individual and whether or not they get diabetes.

It was further asked to perform a Fisher's Exact Test. Utilizing the same function found on section 13.2, I can find a p-value considering cases where $n_{11} \geq 12$. This results in the following sum:

$$\frac{\binom{16}{12}\binom{88}{39}}{\binom{104}{51}} + ... + \frac{\binom{16}{16}\binom{88}{35}}{\binom{104}{51}} \approx 0.02245$$

Notice the question asks whether the relative frequencies are significantly different. This prompts me to multiply the above p-value by 2. That is

$$P(n_{11}) \approx 0.0449$$

This p-value indicates there is decent evidence to show that this type of diabetes appears to be caused by the genetics of an individual.

Problem 19: The question asks to use the Fisher's Exact Test on the data provided and to draw a conclusion from the resulting p-value.

```
##   wait together wait alone total of rows
## 1            12          5            17
## 2             4          9            13
## 3            16         14            30
```

Similar to problem 1, the formula from Section 13.2 can be used to calculate the p-values for such a test. The $P(n_{11} \geq 12)$ is calculated as follows:

$$\frac{\binom{17}{12}\binom{13}{4}}{\binom{30}{16}} + ... + \frac{\binom{17}{16}\binom{13}{0}}{\binom{30}{16}} \approx 0.035474$$

I would then multiply the above p-value by 2 because the question asks "if there is a significant difference" indicating a two-tail test. This results with

$$P(n_{11}) \approx 0.071$$

.

This p-value indicates that there is weak evidence to suggest people with high anxiety would want to wait with another person prior to a shock.

Problem 21: The question asks what are the relevant odds ratio and its estimate for the table in Problem 1. Looking back at the table from earlier, we can calculate the odds ratio with the following formula:

$$\hat{\Delta} = \frac{n_{00}n_{11}}{n_{01}n_{10}} = \frac{(12)(49)}{(39)(4)} \approx 3.792$$

According to the table, the odds of people contracting diabetes are increased by a factor of about 3.8 if they have the Bb or bb genes present in their DNA.

Problem 22: There are a total of four 2×2 tables in this question to calculate odds ratios for and to interpret. All four tables have to do with whether or not patients were advised to stop smoking or not. The first table compares these occurrences between males and females. The odds ratio is calculated as follows:

$$\frac{(48)(136)}{(80)(47)} \approx 1.74$$

This can be interpreted as, the odds of a person being advised to stop smoking increased by a factor of 1.7 if they were a male.

The next table considers the differences in advising when looking at different ethnic groups. The odds ratio for this table is calculated as follows:

$$\frac{(26)(149)}{(102)(34)} \approx 1.12$$

That is, the odds of a person being advised to stop smoking increased by a factor of 1.1 if they were white.

Now we consider the table that accounts if whether advise was given based on the physician's sex. The odds ratio is calculated to be:

$$\frac{(78)(89)}{(50)(94)} \approx 1.477$$

This means that the odds of a physician advising a patient to stop smoking increased by a factor of 1.5 if the doctor was a male.

The final table considers if whether the physician smokes. Calculating the odds ratio the way the table is presented yields an odds ratio less than one. Switching the order of the rows yields the following OR:

$$\frac{(115)(37)}{(13)(146)} \approx 2.242$$

The odds of a physician advising a patient to stop smoking increased by a factor of 2.2 if the physician was a non-smoker.

Problem 27: From the tables provided, I have gathered the following information.

When calculating percentages based on the defendant's race:

- $\frac{37}{621} \approx 5.96\%$ of white defendants received the death penalty.
- $\frac{62}{838} \approx 7.4\%$ of non-white defendants received the death penalty.

When calculating percentages based on the victim's race:

- $\frac{66}{825} \approx 8\%$ of defendants that committed crimes against white victims receive the death penalty.
- $\frac{33}{696} \approx 4.74\%$ of defendants that committed crimes against non-white victims receive the death penalty.

What these percentages seem to indicate is that non-white defendants are more likely to receive the death penalty than white defendants. Furthermore, crimes committed against non-white victims had less death penalty sentences than crimes against white defendants. The judicial system appears to favor white individuals.