Lab 12

Jared Dyreson

TR @ 11:30 - 14:15

1. Write the null hypothesis for this test of independence.

- **Someone's personality is strictly agnostic from their major, contributing in no way, shape or form**

2. If the sample size assumptions are met (all expected counts > 5), what would be the sampling distribution of your test statistic (i.e., what is the type of distribution and the degrees of freedom)?

- **The sampling distribution of my test statistic would be χ^2 and the degrees of freedom is $(4-1)(3-1) = 6$**

3. If the null hypothesis is true, calculate the expected number of analysts who are Computer Science majors. If your number is not an integer, round it to at least one decimal place.

- $\frac{20}{48} \times 7 = 2.916666667$**, where 20 out of the 48 students are computer science majors and there are 7 total diplomats in the class**

4. If the null hypothesis is true, calculate the Pearson residual and contribution to the chi-squared statistic for analysts who are Computer Science majors.

- $\frac{O-E}{\sqrt{E}} = \frac{5-2.92}{\sqrt{2.92}} = 1.217$

5. To obtain the p-value, can we use the sampling distribution from Question #2, or do we have to simulate a sampling distribution? Explain your reasoning. (HINT: look at your answer to Question #3)

```
majors <- read.csv("~/Downloads/majors.csv")
majors.table <- xtabs(~ Major + Personality.Type, data = majors)
```

- **We would need to simulate because the result from question 3 is 2.9 which is less than 5. We need to proceed with method two.**

6. Copy the RStudio output below.

```
chisq.test(majors.table, simulate.p.value = TRUE)
```

```
Pearson's Chi-squared test with simulated
p-value (based on 2000 replicates)

data:  majors.table
X-squared = 7.1918, df = NA, p-value = 0.3088
```

7. What is the value of the chi-square test statistic as computed by R?

- **The value of χ^2 is 7.1918**

8. What is the p-value for this test?

- **The p-value is 0.3088**

9. Using a 5% significance level, can you conclude that people's personality type affects their choice of major?

- **Since the p-value greatly exceeds the cut off of 5 percent, we can fail reject the null hypothesis.**

10. Do you believe that your answer (from Question #9) applies to all students at Cal State Fullerton? (HINT: Think about the sample we used and the way we collected the data)

- **This will most likely not represent the population because we only had a population consisting of natural science majors rather than a broader distribution of majors.**