ST 516: Foundations of Data Analytics

Case Study Example—Computer Instruction

Second Example: Computer Instruction

Taken from Chapter 10 of {*An Introduction to Statistical Methods and Data Analysis*}

> *An educational researcher designs a study to compare the effectiveness of teaching English to non-English speaking people by using a computer software program and by the traditional classroom system. She randomly assigns 125 students from a class of 300 to the computer instruction, and the remaining 175 students to the traditional instruction. At the end of a 6-month period all 300 students are given an exam, with pass/fail results: under computer instruction, 94 passed, 31 failed, and under traditional instruction 113 passed, 62 failed. Is one method better than the other?*

Think About an Approach

Some things to consider:

1. The data are proprotions, and we have been asked a question about comparing proportions.
2. The sample sizes are large, $n_1 = 125$ and $n_2 = 175$.
3. The 300 students were presumably not randomly selected, but rather self-selected into the course.
4. The treatments were randomly assigned by the researchers.
5. Students in the traditional instruction may have results that are not statistically independent due to the potential for their interactions.

Perform an Analysis

```
ys <- c(94,113)
ns <- c(125,175)
prop.test(ys,ns)
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  ys out of ns
## X-squared = 3.3701, df = 1, p-value = 0.06639
## alternative hypothesis: two.sided
## 95 percent confidence interval:
##  -0.004268342  0.216839771
## sample estimates:
##    prop 1    prop 2
## 0.7520000 0.6457143
```

Consider Alternatives

The sample sizes are fairly large here, so the difference in proportions test is valid since the Central Limit Theorem applies.

We should be a little concerned that the pass/fail observations in the traditional instruction group are probably not independent. We make the assumption of independence when we use the differnce in proportions test. The analysis we just performed may be over-stating a difference between the two groups.

Write a Summary

There is only weak evidence ($p = 0.066$ from a two-sample test, $n_1 = 125, n_2 = 175$) of a difference in the proportion of students passing under the two different interventions. A 95% confidence interval for the difference in proportion passing under the computer instruction and the proportion passing under the traditional instruction is (-0.004,0.217)—students appear to do better, on average, under the computer instruction than under the traditional instruction since a large portion of the confidence interval captures positive differnces. This evidence is rather weak, and it is made even weaker when we consider that the responses in the traditional instruction group are not independent, and therefore likely to be even more variable than has been accounted for by this two-sample test procedure.