EDUC 641 Assignment 02 Key

1. Read in the dataset

#### 2. Observed relationship between treat and absenteeism (20% point)

2.1. How many students in treatment group were chronically absent/not chronically absent? How many students in control group were chronically absent/not chronically absent? Provide your response and create a contingency table (frequencies in cell) to demonstrate your answer.

* 86/440 students in treatment group were chronically absent/not chronically absent
* 86/330 students in control group were chronically absent/not chronically absent

Text

Description automatically generated with medium confidence

2.2. What are the percentages of chronically absent and not chronically absent students in treatment group? How about in control group? Provide your response and create a contingency table (percentages in cell) to demonstrate your answer.

* 16%/84% students in treatment group were chronically absent/not chronically absent
* 21%/79% students in control group were chronically absent/not chronically absent



2.3. Create a figure to visualize the relationship between treat and absenteeism. Make sure to label the x- and y-axis.

Chart, bar chart

Description automatically generated

2.4. Based on your observations from 2.1 to 2.3, do you think there is a relationship between absenteeism and treat (e.g., students whose teacher were in treatment group were less likely absent)? Why or why not?

* The table in 2.2. shows that there were 16% students in treatment group who were chronically absent, less than the 21% in control group. However, this difference is not visually observable in figure in 2.3 so we are skeptical about drawing any conclusions here.

#### 3. Chi-Square Test of Categorical Association

3.1. What are the percentages of chronically absent and not chronically absent students in the full sample? Suppose that there’s NO relationship between absenteeism and treat, redo the table in 2.1 using R code or by hand.

* In the full sample, about 18% were chronically absent and 82% were not.



3.2. Calculate the Chi-Squared statistic using the 2.1 and 3.1 tables. Use one sentence to describe the statistical meaning of this statistic. Do you think whether or not this statistic is resulted from sampling idiosyncrasy? Explain.

* The Chi-Squared value of 2.884 measures the net discrepancy between the observed and expected frequencies. It is not but close to zero so it’s possible that it’s a product of sampling idiosyncrasy.

3.3. Our first research question of interest is about whether there was a relationship between the two variables, treat and absenteeism. What is your null hypothesis for this research question?

* There was no relationship between student absenteeism and treatment condition.

3.4. Perform a Pearson’s Chi-squared test (with Yates’ continuity correction) in R to test the relationship between absenteeism and treat. Write 3-4 sentences to interpret your results and answer the research question: Were the students less likely to be chronically absent if their teachers participated the consultancy PD intervention?

* In the population in which the study was conducted, there was no relationship between student absenteeism and consultancy PD intervention (*χ2* = 2.627, *p* = 0.105). In other words, students were not less likely or more likely to be chronically absent if their teachers participated in the consultancy PD intervention.

#### 4. Sub-sample comparisons

4.1. We are also interested in whether absenteeism and treat are related for female students. Perform a Pearson’s Chi-squared test (with Yates’ continuity correction) in R to investigate the relationship between absenteeism and treat for female students. Write 2-3 sentences to state your null hypothesis, interpret your results, and answer the research question: Were the female students less likely to be chronically absent if there teachers participated in the consultancy PD intervention?

* In the population in which the study was conducted, there was a significant relationship between female students’ absenteeism and consultancy PD intervention (*χ2* = 7.335, *p* < 0.05). In other words, female students were less likely to be chronically absent if their teachers participated in the consultancy PD intervention.

4.2. Do the same for male students.

* In the population in which the study was conducted, there was no relationship between male students’ absenteeism and consultancy PD intervention (*χ2* = 0.178, *p* = 0.673). In other words, male students were not less likely or more likely to be chronically absent if their teachers participated in the consultancy PD intervention.