IS THERE AN ASSOCIATION BETWEEN DOING MATH HOMEWORK BECAUSE IT IS FUN AND AVERAGE MATH CONFIDENCE IN HIGH SCHOOL STUDENTS?

MATIAS RELYEA AND KYLE HOLLARS

Abstract. In this paper, we use data obtained from a survey of high school students to determine the link between students having fun on math homework and their confidence in solving math problems.

Introduction and the sample

The data obtained for this paper is from a survey regarding high school math anxiety conducted by a graduate student from Appalachian State University at Watauga High School located in Boone, North Carolina. Questions on the survey varied between four different categories: motivation to do math homework, participation in class, confidence in ability to solve math problems, and anxiety levels surrounding math assignments. We selected a criterion within the "motivation to do math homework category" and decided to conduct an independence test with confidence in ability to solve math problems.

For this survey, there were 19 respondents. All 19 respondents volunteered to participate in the survey, and varied across all grade levels. Furthermore, all 19 respondents responded to this survey whilst in "School's Out", an after-school program at Watauga High School. Questions on the survey asked students to rate their confidence with regard to a statement on a scale of either 1 to 4 or 1 to 5, where a lower number corresponded to being less confident and a higher number corresponded to being more confident.

As the explanatory variable, we used a question that asked volunteers to rate their degree of confidence in relation to the statement "I do my math homework because it is fun" on a scale of 1 to 4. The data that we chose to consider as the response variable was the average degree of confidence that respondents had in solving several math problems (questions of which involved numerical analysis, geometry, and basic algebra) on a scale of 1 to 5, with the final score for each respondent being the average of all confidence scores on all math problems. For the purposes of distributing the data more evenly for the response variable, we partitioned the average confidence scores into two groups: the first group had an average confidence score greater than 1 and less than or equal to 3, and the second group had an average confidence score greater than 3 and less than or equal to 5. The specific math questions, along with all other questions that we did not consider, can be found in the linked data sheet.

Date: May 2023.

Setting up the significance test

Hypotheses and conditions. These are our null and alternative hypotheses.

> H_0 : There is no association between doing math homework because it's fun and average math confidence.
> H_a : There is an association between doing math homework because it's fun and average math confidence.

We will be using an alpha level of $\alpha = 0.05$ for this significance test. As can be seen below, several of the entries in the expected counts matrix do not satisfy large counts, namely, some entries are not > 5. As this experiment involved only 19 volunteers, we cannot generalize to a larger population. The data is inherently categorical. We discuss several of these problems and our solutions to them in the final section.

The data is organized as such:

Conf. by Avg. Conf.	Average confidence $\in [1, 3]$	Average confidence $\in (3, 5]$
Confidence $= 1$	2	5
Confidence $= 2$	0	6
Confidence $= 3$	0	5
Confidence $= 4$	1	0

χ^2 test for independence. Here we have the data arranged in a rectangular matrix:

$$\begin{bmatrix} 2 & 5 \\ 0 & 6 \\ 0 & 5 \\ 1 & 0 \end{bmatrix}$$

The horizontal rows represent individual confidence scores for the explanatory variable "I do math homework because it is fun" from 1 to 4. The vertical columns represent two partitions of the average confidence scores of the response variable mentioned above. We calculate the expected counts to be

$$\begin{bmatrix} 1.11 & 5.89 \\ 0.95 & 5.05 \\ 0.79 & 4.21 \\ 0.16 & 0.84 \end{bmatrix}$$

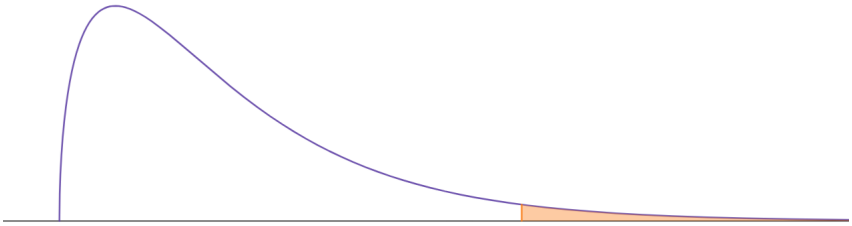
With the expected counts we can calculate the χ^2 test statistics. Thus we have

$$\begin{aligned} \chi^2 &= \frac{(2-1.11)^2}{1.11} + \frac{(5-5.89)^2}{5.89} + \frac{(0-0.95)^2}{0.95} + \cdots + \frac{(0-0.84)^2}{0.84} \\ &= 8.256 \end{aligned}$$

Since we have a 4×2 matrix, we calculate

$$\text{df} = (4-1)(2-1) = (3)(1) = 3.$$

Now we graph the function with the specified values and calculate the p-value.



Then we have

$$p = \chi^2\text{cdf}(8.256, \infty, 3) = 0.0410069.$$

Conclusion. Since $p = 0.0410069 < \alpha = 0.05$, we do have convincing evidence that there is an association between the average confidence score toward the statement "I do homework because it is fun" and their confidence toward solving math problems. With this p-value, we thus reject H_0. Thus we conclude that there is likely an association between "I do math because it is fun" and confidence score in relation to solving math problems.

Discussion

Although we were able to perform a χ^2 significance test on this data, not all of the conditions were met. The 10% condition was met, nineteen students is less than ten percent of the school's population. This data-set was both a convenience sample and was obtained through a voluntary survey. All of these were students who attended School's Out, and chose to fill out a survey about their experience taking math classes. This introduces many possible confounding variables: students that stay after school may tend to be younger, as they are not able to drive themselves home, or they may be older students that choose to stay at school, implying that they enjoy school related activities. On the other hand, School's Out offers math tutoring, so the surveyed students may have greater levels of math anxiety than the true average student, leading them to seek help with their math homework. Additionally, the survey was completely voluntary. Students that were too busy doing homework, students that were engaged in games or other activities, and students that simply did not care about math education may have been less likely to fill out the survey. Kyle Hollars, as a participant in this survey, noticed that the students in the focus group following the survey tended to be those more interested in learning math, having taken advanced math courses. This could lead to a possibly skewed data-set.

The greatest problem with the data, though, was the population size. There were only 19 respondents, meaning that even a random sample would not necessarily be normal by the central limit theorem. The expected counts for each possible response were less than 5, and many of them were less than 1. There just simply wasn't enough data collected to draw a meaningful conclusion. Any conclusions drawn from this data, even with a sufficiently large sample size, wouldn't be able to be generalized to a larger population due to the convenience sample, response bias, and lack of random sampling.

References

Thank you to Maddie Chang from Appalachian State University for allowing us to use the data from her survey. The spreadsheet is linked here.