Gender and Academic Dishonesty in High School Students

Note: the High School name has been redacted for blind judging.

American Statistical Association (ASA) Statistics Project Competition

May 31, 2018

1. Introduction

Academic dishonesty has been a serious problem in classrooms everywhere. The causes and any possible solutions to the problem need to be studied from all perspectives. Before a detailed analysis of academic dishonesty at my school, [High School] in Massachusetts, I decided to consult the Student Handbook for a clear definition. Under the "Academic Dishonesty" section of the handbook, the following behaviors are examples of cheating:

1. Homework/lab reports: A student is cheating when he/she attempts to copy or borrow another student's homework.
2. Quizzes/tests/mid-year and final exams: A student is cheating when he/she attempts to gain any information from another student or from any unauthorized materials.
3. Written projects/research reports: A student is cheating when he/she uses anyone else's words or ideas without documentation (plagiarism).
4. Grades: A student is cheating when he/she changes a grade or answer on a paper/test or quiz.

Students are also to be considered as having cheated if they provide the information or materials for other students for uses as identified above. ("BHS Student Handbook", 30)

I was curious about the proportion of males versus females who had committed one of the cheating actions as mentioned above. I have heard of different stereotypes of male and female genders being sneaky and dishonest. As such, one could assume that one gender would have a higher proportion of people committing academic dishonesty. However, I was wondering whether either assumption had supporting evidence, if either of them had some at all. This became the topic of my investigation.

2. Statistical Question

Question:

Is there a significant difference between the proportion of males and the proportion of females

who have committed academic dishonesty at [High School]?

Parameters:

p_m = proportion of males who have committed an act of academic dishonesty.

p_f = proportion of females who have committed an act of academic dishonesty.

Hypotheses:

H_o: $p_m - p_f = 0$

  The proportion of males and the proportion of females who have committed academic

dishonesty <u>are</u> the same.

H_a: $p_m - p_f \neq 0$

  The proportion of males and the proportion of females who have committed academic

dishonesty <u>are not</u> the same.

3. Data Collection

At the time when the survey was conducted (late May, 2018), the senior (12th grade) class had already graduated and left the high school. As such, I was only able to obtain data from the 9th, 10th, and 11th grade classes.

[High School] has 42 sections of math classes for 9th, 10th, and 11th graders. I decided to list out all the 42 sections and assigned each section a number from 00 to 41. I used a random number table to pick out 10 sections of classes from the 42, and my statistics teacher assisted me in requesting each of the sections' teachers to distribute the survey to his or her students. In this way cluster sampling was conducted, as the population was split into section clusters. A simple random sample of the clusters was selected and the students in the selected 10 clusters were chosen to be in the sample.

The survey was sent out via Google Forms. Especially given the nature of the questions (asking whether the students had committed academic dishonesty or not), I wanted to reduce response bias. To do this, I made sure that no user-specific data (such as email or name) were collected; on the survey itself (Appendix A) I explicitly stated the confidentiality of the survey.

I decided to ask the gender of the student (Male, Female, Other), as one question; also, I decided to have the academic dishonesty question as another. Since I wanted to decrease response bias, I set three choices for the latter question— "yes", "no", and "not sure". This way, I could record a definite "yes" as the student having committed an act of academic dishonesty.

Data, included in Appendix B, were recorded from 5/25/2018 12:01 AM to 5/29/2018 11:59 PM.

4. <u>Data Display</u>

Results of the survey (copy is in Appendix A), conducted between 5/25/2018 12:01 AM and

5/29/2018 11:59 PM.

(Survey Question: Have you ever ACTUALLY committed an act of academic dishonesty, as listed

in the [High School] Student Handbook?)

	Male Students	Female Students
Sample size	$n_m = 54$	$n_f = 67$
Number of students responding "yes" to question	$x_m = 14$ students responded "yes"	$x_f = 13$ students responded "yes"
Proportion	$\hat{p}_m = \frac{14}{54} = 0.259$	$\hat{p}_f = \frac{13}{67} = 0.194$

Raw data can be found in Appendix B.

5. Data Analysis

Parameters: Population is all students at [High School].

p_m = proportion of males who have committed an act of academic dishonesty.

p_f = proportion of females who have committed an act of academic dishonesty.

H_o: p_m - p_f = 0 $\qquad\qquad$ H_a: p_m - $p_f \neq 0$ $\qquad\qquad$ $\alpha = 0.05$

We will conduct a 2-sample z test for proportions.

Conditions:

Randomization: Cluster sampling was conducted, as the population was split into section clusters. A simple random sample of the clusters was selected and the students in the selected 10 clusters were chosen to be in the sample. However, we will still proceed with caution.

Normal: The statistic based on Normal calculations is reasonable because all appropriate counts of successes and failures are greater than 10.

$n_m = 54$	$n_f = 67$
$\hat{p}_c = \frac{14+13}{54+67} = 0.223$	$\hat{q}_c = 1 - \hat{p}_c = 0.777$

$n_m \cdot \hat{p}_c = 54(0.223) = 12.042 > 10$

$n_m \cdot \hat{q}_c = 54(0.777) = 41.958 > 10$

$n_f \cdot \hat{p}_c = 67(0.223) = 14.941 > 10$

$n_f \cdot \hat{q}_c = 67(0.777) = 52.059 > 10$

Independent: The two samples are independent because the proportion of females that have committed an act of academic dishonesty should not affect the proportion of males that have committed an act of academic dishonesty. Each student can be assumed to be independent of each other.

Calculations:

$$z = \frac{(\hat{p}_m - \hat{p}_f) - (p_{om} - p_{of})}{\sqrt{\hat{p}_c \hat{q}_c \left(\frac{1}{n_m} + \frac{1}{n_f}\right)}} \qquad = \frac{(0.259 - 0.194) - 0}{\sqrt{0.223 \cdot 0.777 \cdot (\frac{1}{54} + \frac{1}{67})}} = 0.854$$

P-Value = 0.392

Raw data can be found in Appendix B.

Interpretation:

At the $\alpha = 0.05$ significance level we fail to reject H_o: $p_m - p_f = 0$, that there is no significant difference between the proportion of males that have committed an act of academic dishonesty and the proportion of females that have committed an act of academic dishonesty; this is because the p-value 0.392 is greater than $\alpha = 0.05$. Therefore, it appears that there is a lack of evidence supporting that there is a significant difference in the male and the female student proportions who have committed academic dishonesty at [High School].

6. <u>Conclusion</u>

In this report, I attempted to see if there were data supporting whether there was a significant difference between the proportion of males and the proportion of females who have committed academic dishonesty. In our 2-sample z test for proportions, our determined P-Value was 0.392, which is far greater than the significance level $\alpha = 0.05$. As such, we fail to reject H_o: $p_m - p_f = 0$, that there is no significant difference between the proportion of males that have committed an act of academic dishonesty and the proportion of females that have committed an act of academic dishonesty. Since we have failed to reject H_o, we do not have any evidence supporting that there is a difference in the male and female student proportions who have committed academic dishonesty at [High School].

7. <u>Reflection on Process</u>

First and foremost, our method of sampling was done via survey. While surveys do have its merits of convenience, as was in this study, they are subject to heavy response bias. The mere nature of this survey, as it probed students history on academic dishonesty, affected getting extremely accurate data. There was likely some nonresponse bias: some of the selected sections' teachers did not send out the survey. Some students may have failed to fill out the survey entirely, either due to absence, hesitation, or apathy. However, we did get over a hundred student responses, with significant number of responses from both male and female genders.

Also, the survey question "Have you ever ACTUALLY committed an act of academic dishonesty, as listed in the [High School] Student Handbook?" could have included a more explicit definition of academic dishonesty, like quoting the definition from the handbook.

We easily could have improved the study by including seniors in the survey. The survey was conducted after the senior class had left; conducting the survey prior to their departure would be more effective. Furthermore, the survey study was only representative of [High School] students; its student body was the population. As such, we think further study on other high schools can conclude more broadly. Also, larger sample size has more evident conclusions.

Some other interesting ideas came up while I was conducting this study. It would be interesting to observe the detailed correlation between grade levels instead of only genders, and proportion of each grade where students have committed academic dishonesty. I would use a Chi-Square Test for Homogeneity of Population to compare proportions across multiple grade levels. It would also be interesting to study if course load or sleep time amount also had correlations. Future studies could include other proportion tests between other variables.

8. <u>References</u>

Works Cited

"Student Handbook 2016-17," *[removed] High School*, 2016. Web. Accessed 30 May 2018.

Yates, Daniel S., et al. *The Practice of Statistics, Third Edition: TI-83/84/89 Graphing Calculator Enhanced*. New York: W.H. Freeman, 2008. Print.

9. <u>Acknowledgements</u>

<div align="center">Appendix A: Survey Questions</div>

Survey Questions:(Note: survey was conducted via Google Forms. No emails or other user-specific information was collected.)

<u>AP Statistics Anonymous Survey</u>

Hello! This survey is part of a final project for the AP Statistics class at [High School]. For all questions, please answer honestly. Your response will be completely anonymous. Thank you for your time!

Gender:

    a. Male

    b. Female

    c. Other: _____

Anonymous Academic Dishonesty Question - *For the purposes of this survey, we will use the definition of "academic dishonesty" as listed in the [High School] Student Handbook.*

Have you ever ACTUALLY committed an act of academic dishonesty, as listed in the [High School] Student Handbook?

Please answer honestly. Your response will be completely anonymous.

    a. Yes

    b. No

    c. Not sure

<div align="center">*<survey end>*</div>

Appendix B: Raw Survey Data:

Data was recorded from 5/25/2018 12:01 AM to 5/29/2018 11:59 PM.

Timestamp	Gender:	Have you ever ACTUALLY committed an act of academic dishonesty, as listed in the Belmont High School Student Handbook?
5/25/2018 9:12:18	Male	Yes
5/25/2018 9:12:27	Male	No
5/25/2018 9:13:21	Male	Yes
5/25/2018 9:13:24	Male	Not sure
5/25/2018 9:13:29	Male	Not sure
5/25/2018 9:13:31	Male	Not sure
5/25/2018 9:14:03	Male	Not sure
5/25/2018 9:14:04	Male	Not sure
5/25/2018 9:14:14	Male	No
5/25/2018 9:16:06	Male	No
5/25/2018 9:55:22	Male	Yes
5/25/2018 9:55:43	Female	Yes
5/25/2018 9:56:00	Female	Not sure
5/25/2018 9:56:05	Male	No
5/25/2018 9:56:05	Female	No
5/25/2018 9:56:08	Male	Yes
5/25/2018 9:56:19	Male	No
5/25/2018 9:56:37	Male	No
5/25/2018 9:56:44	Male	No
5/25/2018 9:57:04	Female	No
5/25/2018 9:59:12	Female	Not sure
5/25/2018 11:02:08	Female	Yes
5/25/2018 11:07:05	Male	No
5/25/2018 11:09:30	Female	No
5/25/2018 11:18:53	Female	Not sure
5/25/2018 11:57:33	Male	Not sure
5/25/2018 11:58:26	Female	No
5/25/2018 12:15:16	Male	No

5/25/2018 12:15:30	Male	No
5/25/2018 12:17:04	Female	Yes
5/25/2018 12:17:25	Female	No
5/25/2018 12:24:45	Female	No
5/25/2018 12:42:56	Male	Not sure
5/25/2018 12:43:02	Female	No
5/25/2018 12:43:07	Female	No
5/25/2018 12:43:07	Male	No
5/25/2018 12:43:10	Male	No
5/25/2018 12:43:12	Female	No
5/25/2018 12:43:13	Female	Yes
5/25/2018 12:43:16	Female	Not sure
5/25/2018 12:43:16	Female	No
5/25/2018 12:43:20	Male	No
5/25/2018 12:43:23	Female	No
5/25/2018 12:43:26	Female	Not sure
5/25/2018 12:43:32	Female	Not sure
5/25/2018 12:43:32	Male	Not sure
5/25/2018 12:43:40	Female	Yes
5/25/2018 12:43:49	Male	Yes
5/25/2018 12:44:05	Female	No
5/25/2018 12:44:06	Male	No
5/25/2018 12:44:08	Female	No
5/25/2018 12:44:09	Male	Yes
5/25/2018 12:44:10	Female	Yes
5/25/2018 12:44:12	Male	No
5/25/2018 12:44:17	Male	Yes
5/25/2018 12:44:17	Male	Yes
5/25/2018 12:44:17	Female	No
5/25/2018 12:44:17	Female	No
5/25/2018 12:44:56	Male	No
5/25/2018 13:17:31	Male	Not sure

5/25/2018 13:18:02	Female	Yes
5/25/2018 13:18:08	Male	Yes
5/25/2018 13:18:30	Male	Yes
5/25/2018 13:21:15	Female	Not sure
5/25/2018 13:25:32	Female	No
5/25/2018 13:36:26	Male	No
5/25/2018 13:36:37	Male	Yes
5/25/2018 13:36:48	Female	No
5/25/2018 13:36:58	Female	No
5/25/2018 13:37:03	Female	Yes
5/25/2018 13:37:03	Female	Yes
5/25/2018 13:37:08	Female	No
5/25/2018 13:37:19	Female	No
5/25/2018 13:37:24	Female	No
5/25/2018 13:37:25	Female	No
5/25/2018 13:37:31	Female	No
5/25/2018 13:38:10	Whatever Gender's More Convenient	Not sure
5/25/2018 13:38:54	Male	No
5/25/2018 13:38:55	Female	No
5/25/2018 13:39:03	Female	No
5/25/2018 13:39:46	Male	No
5/25/2018 13:39:59	Female	No
5/25/2018 13:40:42	Female	Not sure
5/25/2018 13:40:48	Male	Not sure
5/25/2018 13:43:03	Male	Yes
5/25/2018 16:27:11	Female	No
5/25/2018 20:13:45	Female	No
5/25/2018 23:28:33	Female	No
5/26/2018 20:48:25	Male	Not sure
5/28/2018 11:37:22	Female	No

5/28/2018 13:53:23	Male	No
5/28/2018 13:54:15	Female	Not sure
5/28/2018 15:14:29	Male	No
5/28/2018 20:41:23	Female	No
5/28/2018 20:50:38	Female	No
5/28/2018 23:56:49	Female	No
5/29/2018 7:58:03	Male	Yes
5/29/2018 7:58:07	Female	No
5/29/2018 7:58:07	Female	Not sure
5/29/2018 7:58:08	Female	Not sure
5/29/2018 7:59:11	Female	Not sure
5/29/2018 7:59:19	Female	No
5/29/2018 9:55:07	Male	No
5/29/2018 10:07:15	Female	Yes
5/29/2018 13:56:25	Male	Not sure
5/29/2018 13:58:04	Female	No
5/29/2018 13:59:51	Female	Yes
5/29/2018 14:02:32	Female	No
5/29/2018 14:07:45	Female	No
5/29/2018 14:10:21	Male	No
5/29/2018 14:13:48	Male	Not sure
5/29/2018 14:26:29	Female	No
5/29/2018 14:29:11	Male	No
5/29/2018 14:37:48	Female	Not sure
5/29/2018 14:39:22	Male	Not sure
5/29/2018 15:32:22	Female	Yes
5/29/2018 16:40:00	Female	Yes
5/29/2018 17:31:40	Female	No
5/29/2018 18:29:24	Male	No
5/29/2018 18:39:28	Male	Yes
5/29/2018 20:22:09	Male	Not sure
5/29/2018 20:40:16	Female	Not sure