A Statistical Analysis of Students' Responses in a Longwood University Survey

Delaney McMahon

MATH 171, Spring 2020, Dr. Steven Hoehner

Introduction

This study was conducted to compare female and male students at Longwood University. A total of 349 students that took MATH 171 and MATH 301 at Longwood University during the Spring 2019 semester were surveyed to collect information such as their approval of the current president and their body mass indexes (BMI). A simple random sample (SRS) of 40 was created from the population. Using the SRS, histograms and boxplots were created to analyze the responses to represent the whole population. The mean BMI of each gender was evaluated more to investigate whether there was a difference in BMI between male and female statistics students. After the hypothesis test, it was determined that we do not have strong enough evidence to determine if there is a difference in BMI of female and male students because the test statistic is less than the z-critical value, therefore we fail to reject the H_0.

Data Collection

In this study, females and males that were currently taking MATH 171 or MATH 301 in Spring 2019 were surveyed about their presidential approval of approve, disapprove, or unsure, being a qualitative discrete variable. The population in this study is all Longwood University students that are taking MATH 171 or MATH 301 in the Spring 2019 semester. Their body mass index (BMI) was also surveyed, being a quantitative continuous variable. A total of 219 females and 130 males were surveyed. After the survey results were obtained, asking for their presidential approval and BMI, a simple random sample (SRS) of 40 females and 40 males were obtained from all survey results. The data analysis tool was used to collect the SRS. All female values were selected for the input range. The sampling method was changed to random and entered 40 for the number of samples. The cell next to the survey answers was selected as the output range. The same steps were used to obtain an SRS for the male surveys; the data analysis tool was used, and all male results were selected for the input range and the sampling method was changed to random and 40 samples was entered. It is reasonable to use simple random samples to approximate the populations of all male and female MATH 171 and MATH 301 students for Spring 2019 because each sample from the population has an equal chance of being selected and as a result it represents the whole population by doing so.

Data Description

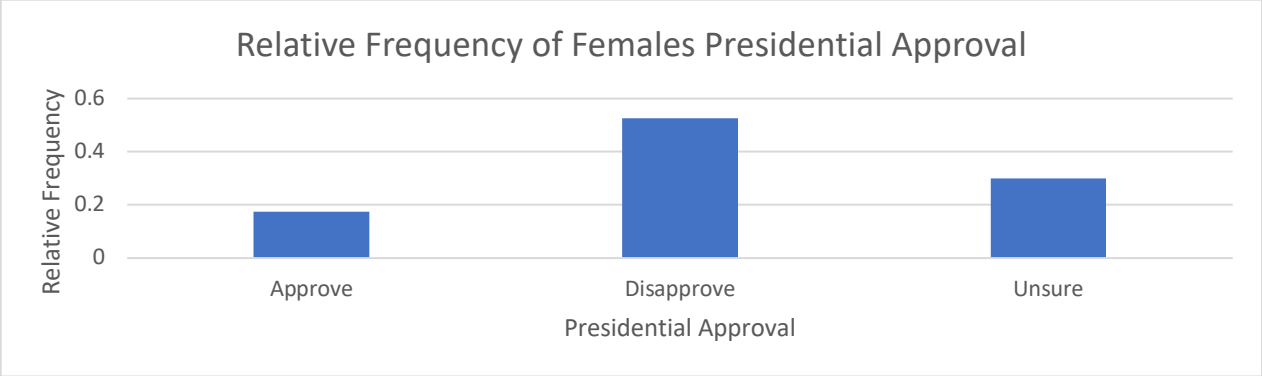
The mode is disapproving for presidential approval for females and approve for presidential approval for males. Out of all females taking MATH 171 and MATH 301 during the Spring 2019 semester, 52.5% does not approve of the current president (Graph 1). Compared to the male students, only 27.5% disapprove of the current president. The mode is approving of the current president for males. Out of all males 37.5% approve of the mode, of approving of the current president (Graph 2). Between the two genders there is an overall difference in the results. More females disapprove or are unsure of the president than approve. While more males approve or unsure than disapprove of the president (Graph 1 and 2).

The relative frequency of male and female BMIs was surveyed. For females, the most frequent class was 18.6-22 kg/m^2. For males, 22.1-25.5 kg/m^2 was the most frequent class for BMI. The most frequented class of BMI was lower in females than in males in the SRSs. The BMIs of females are more distributed than the males (Graph 3). The male histogram shows that they have a more centralized class towards the class of 22.1-25.5, while the female BMIs are spread out across all classes almost evenly in the larger BMI classes (Graph 4). The shape of the female histogram is skewed right while the shape of the male BMI histogram is symmetrical (Graph 3 and 4). Based on the box plots, the mean for the female BMIs was 25.52 and the mean male BMI was 24.52. The minimum for female BMIs was 17.3 and the maximum was 36.7. For males BMI, the minimum was 16.8 and the maximum was 32.6. The median for the female BMIs was 24.45 and the males BMI median was 23.95. There were no outliers for the female BMI results. The upper fence of the female BMI was 40.9 and the maximum value in the SRS was 36.7 and the lower fence was 9.3 and the minimum was 16.8, so there were no outliers. The same went for the male BMIs. There were no outliers because the upper fence for males BMI was 32.6 and the maximum was 32.6, so that BMI value could be considered an outlier. The lower fence for male BMI values was 16.3 and the lowest BMI was 16.8, so therefore there were no outliers (Graph 5 and 6). The upper fence was calculated by using the third quartile (Q3) value and adding it to the interquartile range (IQR) value that was multiplied by 1.5. The lower fence was calculated by adding the first quartile (Q1) value to the IQR that was multiplied by 1.5. The IQR was calculated by subtracting Q1 from Q3. A box plot was generated from the mean, standard deviation, and the five-number summary including the median, maximum, minimum, Q1 and Q3 (Table 1 and 2).

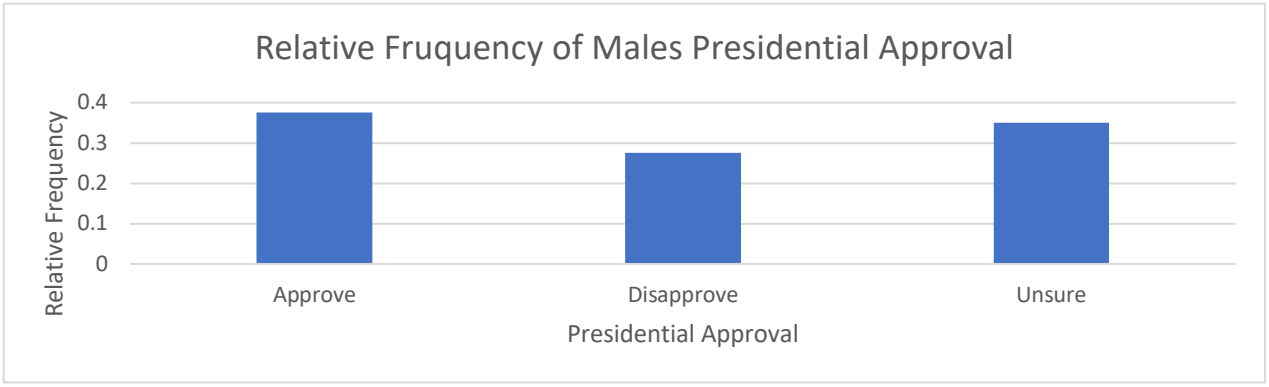
Data Analysis

The required conditions on the sample are satisfied to run the hypothesis test because the two samples were randomly and independently selected, and the populations are normal (greater than or equal to 30). A two tailed test was used to determine the difference in BMI between male and female statistics at Longwood University. The hypotheses to determine the difference in BMI of the two genders are H_0: $\mu_1 - \mu_2 = 0$ and H_a: $\mu_1 - \mu_2 \neq 0$. The mean for female BMI was 25.5 kg/m^2 and the variance was 31.9. The mean for male BMI was 24.5 kg/m^2 and the variance was 14.9. The z-critical value is 1.96 The test statistic was z* = $\frac{(25.5-24.5)-0}{\sqrt{\frac{31.9}{40}+\frac{14.9}{40}}} = 0.92$. The degrees of freedom are unable to be calculated because it is associated with t-distributions, not z-distributions which was used in this study. The p-value is 0.356. Since z* > $z_{\alpha/2}$, we fail to reject the H_0. This means that we do not have strong enough evidence to conclude that there is a difference in the mean BMI between females and males at Longwood University. The result is not statistically significant because the p-value is higher than α (0.05). With 95% confidence, the difference in the two-population means is between -1.122 and 3.1223. The confidence interval that was found supports the result of the hypothesis test because the z* value of 0.92 is within the values of -1.122 and 3.1223.

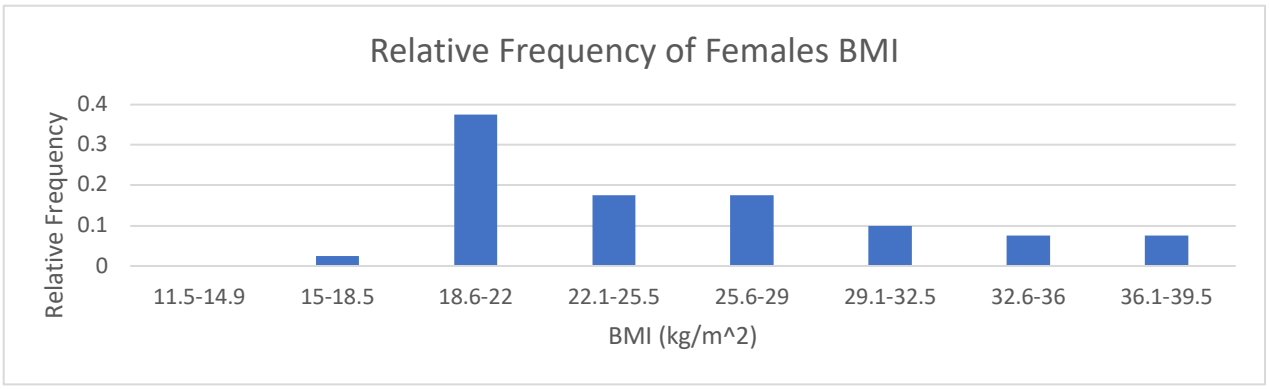
Appendix



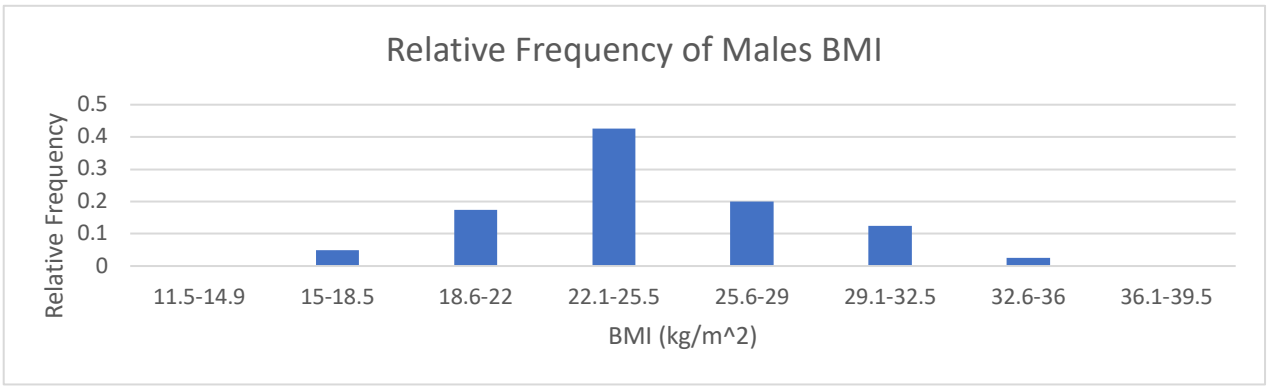
Graph 1. Relative Frequency of Female SRS of Presidential Approval of MATH 171 and MATH 301 Students at Longwood University Spring 2019.



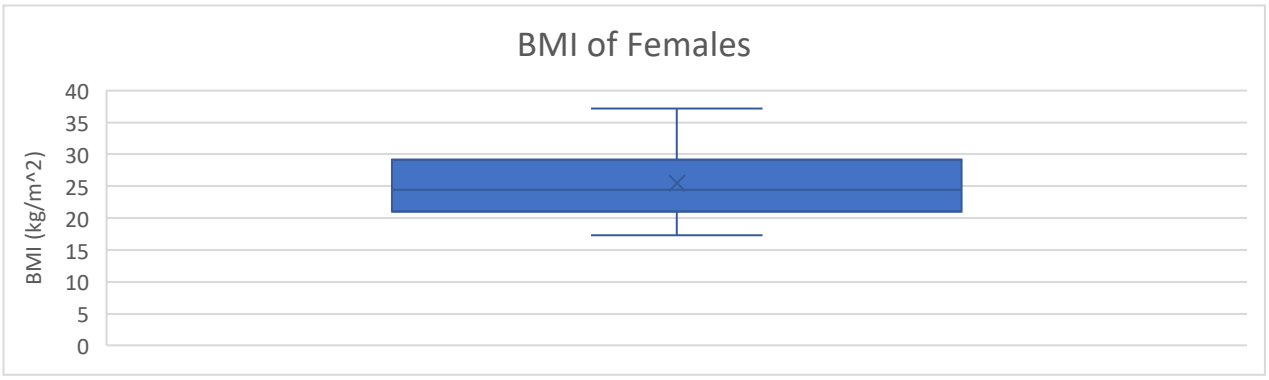
Graph 2. Relative Frequency of Male SRS of Presidential Approval of MATH 171 and MATH 301 Students at Longwood University Spring 2019.



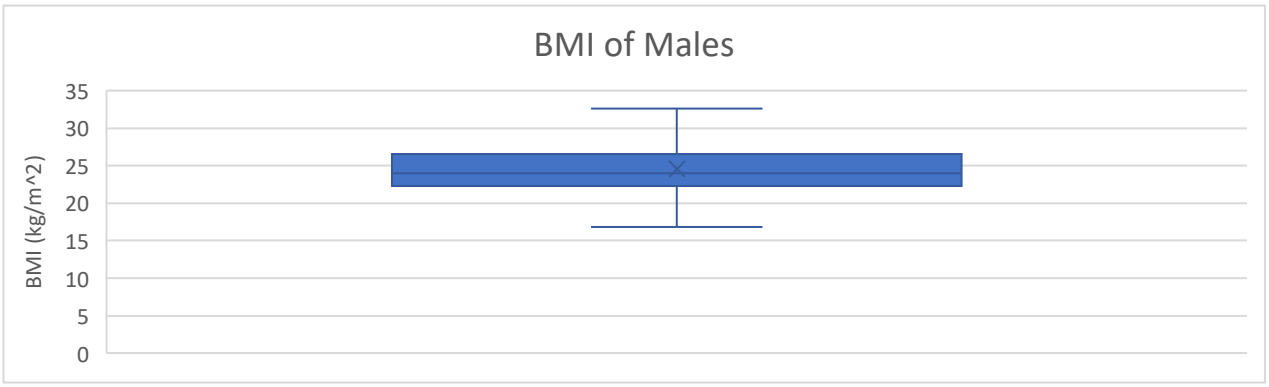
Graph 3. Relative Frequency of BMIs of Female SRS of MATH 171 and MATH 301 Students at Longwood University Spring 2019.



Graph 4. Relative Frequency of BMIs of Male SRS of MATH 171 and MATH 301 Students at Longwood University Spring 2019.



Graph 5. Box Plot of BMIs of Female SRS of MATH 171 and MATH 301 Students at Longwood University Spring 2019.



Graph 6. Box Plot of BMIs of Male SRS of MATH 171 and MATH 301 Students at Longwood University Spring 2019.

Table 1. Five Number Summary of Female SRS of MATH 171 and MATH 301 Students at Longwood University Spring 2019.

Minimum	Q1	Median	Q3	Maximum
17.3	21.1	23.4	28.5	36.7

Table 2. Five Number Summary of Male SRS of MATH 171 and MATH 301 Students at Longwood University Spring 2019.

Minimum	Q1	Median	Q3	Maximum
16.8	22.375	23.95	26.45	32.6