**Comparing Exercise across Gender and Study to Exercise**

Tom Lever

April 23, 2021

Statistics I

Spring 2021

Piedmont Virginia Community College

**Introduction**

Exercise is important for physical, mental, and emotional health and well-being and for happiness and ease of learning. In sections I, II, and III of this paper, the research questions , , and , explicated below, are addressed. This paper provides context and motivation for conducting a study to consider whether, on average, female students at PVCC exercise for a different number of hours per week than male students at PVCC; whether, on average, the number of hours per week that students at PVCC study is greater than the number of hours per week that students at PVCC exercise; and whether there is a correlation between the number of hours per week that students at PVCC study and the number of hours per week that students at PVCC exercise.

|  | Is the mean number of hours exercise per week corresponding to a population of female Statistics-I students at PVCC not equal to the mean number of hours exercise per week corresponding to a population of male Statistics-I students at PVCC? |
| --- | --- |
|  | Is the mean number of hours per week that Statistics-I students at PVCC, in a population , study greater than the mean number of hours per week that the students exercise? |
|  | Is there a correlation between number of hours per week that Statistics-I students at PVCC, in a population , study, and the number of hours per week that the students exercise? |

1. **Is not equal to ?**

**Samples and Datasets**

To address and evaluate a corresponding claim that “ is not equal to ”, two appropriate datasets of hours exercise per week were acquired. The dataset corresponds to a sample of female Statistics-I students at PVCC; the dataset corresponds to a sample of male Statistics-I students at PVCC. and are independent; there is no loss of information if one or both of the datasets are reordered arbitrarily.

Collection of the quantitative data in the datasets was managed by Dr. Irina Timchenko, Ph.D., a Lecturer at PVCC. Dr. Timchenko selected by convenience two clusters of Statistics-I students at PVCC from a much larger population of Statistics-I students at PVCC. Dr. Timchenko requested that each student complete a survey by filling out one row of personal data in a Google Sheet corresponding to the student’s section. Students voluntarily completed the survey; surveys were completed at students’ conveniences. Surveys were completed anonymously. Students’ survey answers may have been affected by students reading responses of other students; only the first student to complete a survey in each section completed the survey blindly. Dr. Timchenko combined the rows in each Google Sheet and performed data cleaning (e.g., by substituting a for a ).

Organization of collected data into the datasets was managed by Mr. Thomas Lever, the author of this paper. Mr. Lever extracted the Gender column of qualitative, categorical, nominal data and the Hours Exercise per Week column of quantitative, continuous, ratio data from Dr. Timchenko’s combined Google Sheet. Mr. Lever changed the nominal values , , and in the Gender column to (male), (female), and (non-binary). Mr. Lever sorted the extracted data first by gender (in alphabetical order) and then by Hours Exercise per Week (in ascending order). Mr. Lever extracted hours exercise per week data corresponding to the male and female genders to the appropriate dataset. The two datasets are shown below. Summary statistics are shown for the datasets.

| | **Hours Exercise per Week** | | | | --- | --- | --- | |  | **DF (h)** | **DM (h)** | | **1** | 0 | 0 | | **2** | 0 | 0 | | **3** | 0 | 0 | | **4** | 0 | 0 | | **5** | 0 | 0 | | **6** | 0 | 1 | | **7** | 1 | 2 | | **8** | 1 | 2 | | **9** | 2 | 2 | | **10** | 2 | 2 | | **11** | 2 | 3 | | **12** | 2 | 3 | | **13** | 2 | 3 | | **14** | 3 | 4 | | **15** | 3 | 4 | | **16** | 3 | 5 | | **17** | 3 | 5 | | **18** | 3 | 6 | | **19** | 4 | 6 | | **20** | 4 | 7 | | **21** | 4 | 7 | | **22** | 4 | 7 | | **23** | 5 | 7 | | **24** | 5 | 7 | | **25** | 5 | 8 | | **26** | 5 | 10 | | **27** | 5 | 15 | | **28** | 5 |  | | **29** | 5 |  | | **30** | 5 |  | | **31** | 5 |  | | **32** | 6 |  | | **33** | 7 |  | | **34** | 8 |  | | **35** | 10 |  | | **36** | 10 |  | | **37** | 10 |  | | **38** | 30 |  | | | **Summary Statistics** | | | | --- | --- | --- | |  | **DF** | **DM** | | **Number of Data (unitless)** | 38 | 27 | | **Min** | 0 | 0 | | **Q1** | 2 | 2 | | **Median** | 4.0 | 4 | | **Q3** | 5 | 7 | | **Max** | 30 | 15 | | **IQR** | 3.0 | 5.0 | | **1.5 IQR** | 4.5 | 7.5 | | **Lower Fence** | -2.5 | -5.5 | | **Upper Fence** | 9.5 | 14.5 | | **Outlier Indices (unitless)** | 35, 36, 37, 38 | 27 | | **Mean** | 4.4 | 4.3 | | **Mode(s)** | 5 | 0, 7 | | **Range** | 30.0 | 15.0 | | **Sample Standard Deviation** | 5.1 | 3.6 | | **Variance (h2)** | 25.8 | 12.8 | | **Skewness (unitless)** | 3.3 | 0.9 | | **Standard Error of the Mean** | 0.8 | 0.7 |   Values are in hours unless otherwise noted.  Whole numbers are chosen.  Numbers with decimals are calculated. |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |

**Addressing and Evaluating**

*Classical / Frequentist Independent Samples T-Test, and Requirements*

The only method known to the author for evaluating is a statistical hypothesis test known as a Classical / Frequentist Independent Samples T-Test, which relies on and , the following two hypotheses, and the following four requirements. The *alternative hypothesis* is identical to . The *null hypothesis* is that “ is equal to ”. The first requirement is that and are independent samples. The second requirement is that any sample from is a simple random sample, and that any sample from is a simple random sample. The third requirement is that either

* The size of is greater , and the size of is greater than , or
* The frequency distribution of number of hours exercise per week for is normal, and the frequency distribution of number of hours exercise per week for is normal.

The fourth requirement is that and do not contain any erroneous outliers corresponding to SF and .

*Considering the Requirement of Independent Samples*

Per the section “Samples and Datasets”, and are independent samples. is met.

*Considering the Requirement of Simple Random Samples*

Per the section “Samples and Datasets”, and are not simple random samples; in fact, they are convenience samples. is not met. Therefore, while the steps of an Independent Samples T-Test will be performed, the results of the hypothesis test may be invalid because and do not represent and .

*Reasoning through Requirement of Large Sample Size or Normality, and*

*Describing an Independent Samples T-Test*

Consider a histogram of many bars with heights equal to frequencies and class widths equal to the difference between the midpoint between one mean and the next larger mean and the midpoint between that mean and the next smaller mean.The proportion of means within a certain range is equal to the ratio of the sum of the frequencies for those means and the sum of all frequencies. Because all bars have the same width, the proportion of means within a certain range is also equal to the ratio of the sum of bar areas for those means and the sum of all bar areas. Consider specifically a histogram where the sum of all bar areas is 1. In this case, the proportion of means within a certain range is equal to the sum of bar areas for those means. The proportion of means within a certain range is also equal to the probability that a mean is in that range. Each bar area is a probability. The points at the top center of the histogram bars constitute a distribution and the range of a distribution function. An area under this distribution is a probability, the distribution is a probability density distribution, and the distribution function is a probability density distribution function.

By the Central Limit Theorem, if is greater than or is normal, then the mean number of hours exercise per week for has a corresponding frequency in a normal frequency distribution, and a corresponding probability density in a normal probability density distribution of mean numbers of hours exercise per week for infinitely many theoretical samples of size from . This *sample-mean probability density distribution* is centered around (unknown) and has a standard deviation, and standard deviation of the sample mean , , where (unknown) is the standard deviation for . Note that standard error of the sample mean is .

If is greater than or is normal, then the mean number of hours exercise per week for has a corresponding frequency in a normal frequency distribution of mean numbers of hours exercise per week for many theoretical samples of size from . This samplingdistribution is centered around (unknown) and has a standard deviation, and standard error of the mean , , where (unknown) is the standard deviation for .

Furthermore, if and are greater than or and are normal, then the difference has a corresponding frequency in the frequency distribution of the differences between a random mean number of hours exercise per week from and a random mean number of hours exercise per week from . is normal, is centered around the difference (unknown and, essentially, value of interest), and has a standard deviation

(unknown)

According to an Independent Samples T-Test with null hypothesis ,

If were known, the z*-score* and *test statistic*

could be computed. Because has a corresponding frequency in and is normal, would have a corresponding frequency in the standard normal frequency distribution of *z*-scores corresponding to differences between a random mean number of hours exercise per week from and a random mean number of hours exercise per week from . Then, a probability and p*-value* that a *z*-score is less than or equal to or greater than or equal to , assuming is true, could be computed. If were deemed low (say, is less than the conventional significance level ), would be rejected, the results of the hypothesis test would be statistically significant, and would be supported. If were greater than , then we would fail to reject , the results of the hypothesis test would be statistically insignificant, and we could not support .

Let the number of degrees of freedom for be .

Let be the standard deviation corresponding to the sample of female Statistics-I students at PVCC.

Let the number of degrees of freedom for be .

Let be the standard deviation corresponding to the sample of male Statistics-I students at PVCC.

Let the pooled standard deviation for and be

Let an approximation of be

Because is unknown, the t*-score* and test statistic

is found instead of . For infinitely large and , and . For infinitely large and , has a corresponding frequency in . For small or , generally speaking,

or . Because of this bias, is considered to have a corresponding frequency not in but in a Student’s *t* distribution centered at with number of degrees of freedom , which might seem like a “horizontally stretched” version of . is symmetric. Then, a probability and *p*-value that a *t*-score is less than or equal to or greater than or equal to , assuming is true, could be computed. If this *p*-value is deemed low (say, is less than the conventional significance level ), then is rejected, the results of the hypothesis test are statistically significant, and is supported. If is greater than , then we fail to reject , the results of the hypothesis test are statistically insignificant, and we cannot support .

*Considering the Requirement of Large Sample Size or Normality*

Per the section “Samples and Datasets”, is greater .

is smaller than .

To evaluate whether is normal, a Shapiro-Wilk Test involving a null hypothesis that “ is normal” and an alternative hypothesis that “ is not normal” was performed using JASP. A CSV file containing the column header “Hours Exercise per Week” and was imported. A “One Sample T-Test” was performed with variable “Hours Exercise per Week”. An Assumption Check for normality was performed. The calculated probability **p** is presented below.

| **Test of Normality (Shapiro-Wilk)** | | | | | |
| --- | --- | --- | --- | --- | --- |
|  | | **W** | | **p** | |
| Hours Exercise per Week |  | 0.910 |  | 0.023 |  |
|  | | | | | |
| *Note.*  Significant results suggest a deviation from normality. | | | | | |

Per Dr. Timchenko, if the probability **p** in the results table is less than a chosen significance level, then the null hypothesis of a Shapiro-Wilk Test is rejected and the alternative hypothesis is supported. The conventional significance level is chosen. Since **p** is less than , is rejected, is supported, and there is sufficient evidence to support the alternative hypothesis that is not normal.

Because is smaller than , and is not normal, is not met. Therefore, while the steps of an Independent Samples T-Test will be performed, the results of the hypothesis test may be invalid.

*Considering the Requirement of Removing Erroneous Outliers*

While the datasets and contain outliers corresponding to and , these outliers are not erroneous. Therefore, is met.

*Performing the Hypothesis Test for the Claim*

is, “Is the mean number of hours exercise per week corresponding to a population of female Statistics-I students at PVCC not equal to the mean number of hours exercise per week corresponding to a population of male Statistics-I students at PVCC?”.

is, “The mean number of hours exercise per week corresponding to a population of female Statistics-I students at PVCC is not equal to the mean number of hours exercise per week corresponding to a population of male Statistics-I students at PVCC”.

The hypothesis test is an Independent Samples T-Test because compares the mean of and , and are independent samples, and and are unknown.

is met because *SF* and *SM* are independent samples.

is not met because *SF* and *SM* are convenience samples. The results of the hypothesis test may be invalid.

is not met because is less than and is not normal. The results of the hypothesis test may be invalid.

is met because the outliers of and corresponding to and are not erroneous.

is, “ is equal to ”.

is .

The relevant statistics relating to *SF* and *SM* are:

| ***SF*** | ***SM*** |
| --- | --- |
|  |  |
|  |  |
|  |  |

per the null hypothesis.

Let be the Student’s *t* Cumulative Distribution Function for a lowest *t*-score , a highest *t*-score , and the number of degrees of freedom for the symmetric Student’s *t* Distribution . Then

This hypothesis test was also performed using JASP. A CSV file containing the column headers “Gender” and “Hours Exercise per Week”, and stacked in the “Hours Exercise per Week” column, and instances of “F” and “M” in the “Gender” column was imported. A Classical / Frequentist Independent Samples T-Test was performed with data variable “Hours Exercise per Week” and grouping variable “Gender”. Some calculated values are presented below.

| **Group Descriptives** | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | **Group** | | **N** | | **Mean** | | **SD** | | **SE** | |
| Hours Exercise per Week |  | F |  | 38 |  | 4.447 |  | 5.076 |  | 0.823 |  |
|  |  | M |  | 27 |  | 4.296 |  | 3.582 |  | 0.689 |  |
|  | | | | | | | | | | | |

| **Independent Samples T-Test** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | | **t** | | **df** | | **p** | |
| Hours Exercise per Week |  | 0.133 |  | 63 |  | 0.895 |  |
|  | | | | | | | |
| *Note.*  Student's t-test. | | | | | | | |

A significance level is chosen.

Because the probability **p** is greater than , we fail to reject .

Because we have failed to reject , at the significance level , the results of our Independent Samples T-Test are statistically insignificant.

Because we have failed to reject , at the significance level , there is insufficient evidence to support .

Because there is insufficient evidence to support , at the significance level , there is insufficient evidence to answer with “Yes”.

Because requirements were not met, these four verdicts may not be valid.

**A Confidence Interval for *D***

*Defining the Practical Confidence Interval*

A confidence interval for parameter , based on the sets of sample statistics for <sample of female PVCC students> and <sample of male PVCC students>, is a range of differences , symmetric about difference , such that the probability that is within this range is equal to a confidence level . A proportion of differences equal to will have confidence intervals for actually containing , and a proportion of differences equal to will differ from by no more than half the width of the confidence interval.

Let confidence level

where is the number of regions considered in determining .

The z-score and test statistic

has a corresponding probability density in the standard normal probability density distribution . This standard normal distribution is centered at and has a standard deviation of .

For *N*, there exists a critical value such that the probability of a -score being between

and is equal to . This critical value is given by

In math,

The difference has a corresponding frequency in the normal frequency distribution . Because is normal, the probability of a difference being between and is .

In math,

Because is unknown and is, in fact, the value of interest, is approximated with .

Let be the margin of error for and .

The ideal confidence interval for is given by

The *t*-score and test statistic

is utilized instead of . is considered to have a corresponding frequency in the Student’s *t* distribution centered at with number of degrees of freedom .

For the Student’s t distribution , there exists a critical value such that the probability of a *t*-score *t* being between and is equal to . This critical value is given by

Let be an approximation of the margin of error .

The practical confidence interval for for confidence level is given by

*Calculating the Practical Confidence Interval*

A significance level was chosen during the Independent Samples T-Test of . A significance level is chosen.

The practical confidence interval for *D* for confidence level *L* is given by

This confidence interval calculation was also performed using JASP. A CSV file containing the column headers “Gender” and “Hours Exercise per Week”, and stacked in the “Hours Exercise per Week” column, and instances of “F” and “M” in the “Gender” column was imported. A Classical / Frequentist Independent Samples T-Test was performed with data variable “Hours Exercise per Week” and grouping variable “Gender”. A confidence interval for for a confidence level of percent was displayed. Some calculated values are presented below.

| **Independent Samples T-Test** | | | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | | | | | **95% CI for Mean Difference** | | | |
|  | | **t** | | **df** | | **p** | | **Mean Difference** | | **SE Difference** | | **Lower** | | **Upper** | |
| Hours Exercise per Week |  | 0.133 |  | 63 |  | 0.895 |  | 0.151 |  | 1.138 |  | -2.122 |  | 2.424 |  |
|  | | | | | | | | | | | | | | | |
| *Note.*  Student's t-test. | | | | | | | | | | | | | | | |

Since the practical confidence interval for for a confidence level of percent contains , at the significance level , there is no significant difference between and .

Because there is no significant difference between and , at the significance level , there is insufficient evidence to support .

Because there is insufficient evidence to support , at the significance level , there is insufficient evidence to answer with “Yes”.

These three reasonings confirm the results of the Independent Samples T-Test of *C*. These reasonings may not be valid.

1. **Is greater than ?**

**Samples and Datasets**

To address and evaluate a corresponding claim that “ is greater than ”, a dataset corresponding to a sample of Statistics-I students at PVCC was acquired. Each matched pair in corresponds to a unique student in and is of the form , where indicates the number of hours per week the th student in studies, and represents the number of hours per week that student exercises.

Mr. Lever extracted the “Average study time per week (in hrs)” and “Average number of hours per week you spend participating in sports or exercising” columns of quantitative, continuous, ratio data from Dr. Timchenko’s combined Google Sheet to dataset . Dataset is shown below.

| | **: Hours Study and Exercise per Week** | | | | | | | --- | --- | --- | --- | --- | --- | | ***i*** | **(h)** | **(h)** | ***i*** | **(h)** | **(h)** | | **1** | 3 | 7 | **34** | 2 | 1 | | **2** | 25 | 5 | **35** | 5 | 3 | | **3** | 15 | 10 | **36** | 6 | 4 | | **4** | 1 | 3 | **37** | 3 | 4 | | **5** | 12 | 0 | **38** | 20 | 0 | | **6** | 3 | 5 | **39** | 28 | 10 | | **7** | 15 | 5 | **40** | 30 | 30 | | **8** | 21 | 3 | **41** | 20 | 0 | | **9** | 19 | 5 | **42** | 20 | 0 | | **10** | 8 | 2 | **43** | 20 | 3 | | **11** | 8 | 3 | **44** | 15 | 1 | | **12** | 10 | 0 | **45** | 20 | 0 | | **13** | 8 | 5 | **46** | 15 | 7 | | **14** | 6 | 8 | **47** | 15 | 3 | | **15** | 15 | 6 | **48** | 20 | 10 | | **16** | 5 | 3 | **49** | 15 | 0 | | **17** | 4 | 5 | **50** | 15 | 5 | | **18** | 5 | 4 | **51** | 20 | 10 | | **19** | 6 | 4 | **52** | 20 | 15 | | **20** | 6 | 4 | **53** | 19 | 5 | | **21** | 5 | 10 | **54** | 20 | 7 | | **22** | 5 | 5 | **55** | 12 | 5 | | **23** | 5 | 0 | **56** | 4 | 5 | | **24** | 5 | 2 | **57** | 6 | 7 | | **25** | 4 | 3 | **58** | 25 | 2 | | **26** | 5 | 7 | **59** | 5 | 4 | | **27** | 2 | 6 | **60** | 6 | 2 | | **28** | 4 | 0 | **61** | 4 | 2 | | **29** | 5 | 1 | **62** | 15 | 6 | | **30** | 4 | 7 | **63** | 6 | 2 | | **31** | 1 | 0 | **64** | 4 | 2 | | **32** | 7 | 2 | **65** | 8 | 2 | | **33** | 5 | 0 | **66** | 6 | 8 | | | **: Summary Statistics** | | | --- | --- | | **Number of Pairs (unitless)** | 66 | | **Mean of Differences** | 6.152 | | **Sample Standard Deviation**  **of Differences** | 7.275 | | **Pearson’s Linear Correlation Coefficient (unitless)** | 0.354 |   Values are in hours unless otherwise noted.  Whole numbers are chosen.  Numbers with decimals are calculated. |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |

**Addressing and Evaluating**

*Classical / Frequentist Paired Samples T-Test, and Requirements*

The only method known to the author for evaluating is a statistical hypothesis test known as a Classical / Frequentist Paired Samples T-Test, which relies on , the following two hypotheses, and the following three requirements. The alternative hypothesis is identical to . The null hypothesis is that “ is equal to ”. The first requirement is that , associated with , consists of matched pairs of data. The second requirement is that any sample from is a simple random sample. The third requirement is that either

* The size of is greater , or
* The frequency distribution of differences is normal.

The fourth requirement is that does not contain any erroneous outliers corresponding to .

*Considering the Requirement of Matched Pairs*

consists of matched pairs, each corresponding to a unique student in and belonging to a tuple . Reordering one column of without reordering the other column of would corrupt the pairings.

*Considering the Requirement of Simple Random Samples*

is not a simple random sample; in fact, is a convenience sample. Therefore, while the steps of a Paired Samples T-Test will be performed, the results of the hypothesis test may be invalid because does not represent .

*Reasoning through Requirement of Large Sample Size or Normality, and*

*Describing a Paired Samples T-Test*

By the Central Limit Theorem, if is greater than or is normal, then , the mean of all differences for , has a corresponding frequency in the frequency distribution of the means for many random samples from . is normal, is centered around the difference (unknown and, essentially, value of interest), and has a standard deviation, and “standard error of the mean ” (unknown)

where is the standard deviation of differences for .

According to a Paired Samples T-Test with null hypothesis ,

If were known, the *z-*score and test statistic

could be computed. Because has a corresponding frequency in and is normal, would have a corresponding frequency in the standard normal frequency distribution of *z*-scores corresponding to means . Then, a probability and *p*-score that a *z*-score is greater than or equal to , assuming is true, could be computed. If this probability is deemed low (say, is less than the conventional significance level ), then is rejected, the results of the hypothesis test are statistically significant, and is supported. If is greater than , then we fail to reject , the results of the hypothesis test are statistically insignificant, and we cannot support .

Let an approximation of be

where is the standard deviation of the differences for .

Because is unknown, the *t-score* and test statistic

is found instead of . For infinitely large , . For infinitely large , has a corresponding frequency in . For small , generally speaking, . Because of this bias, is considered to have a corresponding frequency not in but in a Student’s *t* distribution *T* centered at with number of degrees of freedom , which might seem like a “horizontally stretched” version of . Then, a probability that a *t*-score is greater than or equal to , assuming is true, could be computed. If this probability is deemed low (say, is less than the conventional significance level ), then is rejected, the results of the hypothesis test are statistically significant, and is supported. If is greater than , then we fail to reject , the results of the hypothesis test are statistically insignificant, and we cannot support .

*Considering the Requirement of Large Sample Size or Normality*

Per the section “Samples and Datasets”, is greater than .

*Considering the Requirement of Removing Erroneous Outliers*

While the dataset contains outliers, these outliers are not erroneous. Therefore, is met.

*Performing the Hypothesis Test for the Claim*

Research question is, “Is the mean number of hours per week that Statistics-I students at PVCC, in a population , study greater than the mean number of hours per week that the students exercise?”.

Claim is, “The mean number of hours per week that Statistics-I students at PVCC, in population , study is greater than the mean number of hours per week that the students exercise”.

The hypothesis test is a Paired Samples T-Test because compares two parameters of , consists of matched pairs of data, and is unknown.

Requirement is met because consists of matched pairs, each corresponding to a unique student in and of the form .

Requirement is not met because is a convenience sample. The results of the hypothesis test may be invalid.

Requirement is met because is greater than .

Requirement is met because the outliers of are not erroneous.

Null hypothesis is, “ is equal to ”.

Alternative hypothesis is .

The relevant statistics relating to are:

|  |
| --- |
|  |
|  |

Difference per the null hypothesis.

Let be the Student’s *t* Cumulative Distribution Function for a lowest *t*-score , a highest *t*-score , and the number of degrees of freedom for the symmetric Student’s *t* Distribution . Then the probability that a *t*-score is greater than or equal to , assuming is true,

This hypothesis test was also performed using JASP. A CSV file containing the columns of data “Hours Study per Week” and “Hours Exercise per Week” was imported. A Classical / Frequentist Paired Samples T-Test was performed with a variable pair of “Hours Study per Week” and “Hours Exercise per Week”. Some calculated values are presented below.

| **Paired Samples T-Test** | | | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Measure 1** | |  | | **Measure 2** | | **t** | | **df** | | **p** | | **Mean Difference** | | **SE Difference** | |
| Hours Study per Week |  | - |  | Hours Exercise per Week |  | 6.869 |  | 65 |  | < .001 |  | 6.152 |  | 0.896 |  |
|  | | | | | | | | | | | | | | | |
| *Note.*  Student's t-test. | | | | | | | | | | | | | | | |

A significance level is chosen.

Because the probability **p** that a *t*-score is greater than , assuming is true, is less than , we reject .

Because we have rejected , at the significance level , the results of our Independent Samples T-Test are statistically significant.

Because we have rejected , at the significance level , there is sufficient evidence to support .

Because there is sufficient evidence to support , at the significance level , there is sufficient evidence to answer with “Yes”.

These verdicts may not be valid.

**A Confidence Interval for**

*Defining the Practical Confidence Interval*

The z-score and test statistic

has a corresponding frequency in the standard normal frequency distribution . This standard normal distribution is centered at and have a standard deviation of .

Let confidence level

where is the number of regions considered in determining the probability .

For *N*, there exists a critical value such that the probability of a -score *z* being between

and is equal to . This critical value is given by

The mean has a corresponding frequency in the normal frequency distribution . Because is normal, the probability of a mean being between and is .

Because is unknown and is, in fact, the value of interest, is approximated with .

Let be the margin of error for .

The ideal confidence interval for for confidence level is given by

Since is unknown, let an approximation of be

The *t*-score and test statistic

is utilized instead of . is considered to have a corresponding frequency in a Student’s *t* distribution centered at with number of degrees of freedom .

For , there exists a critical value such that the probability of a *t*-score *t* being between and is equal to . This critical value is given by

Let be an approximation of the margin of error .

The practical confidence interval for for confidence level is given by

*Calculating the Practical Confidence Interval*

A significance level was chosen during the Paired Samples T-Test of . A significance level is chosen.

This hypothesis test was also performed using JASP. A CSV file containing the columns of data “Hours Study per Week” and “Hours Exercise per Week” was imported. A Classical / Frequentist Paired Samples T-Test was performed with a variable pair of “Hours Study per Week” and “Hours Exercise per Week”. A confidence interval for for a confidence level of percent was displayed. Some calculated values are presented below.

| **Paired Samples T-Test** | | | | | | | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | | | | | | | | | **90% CI for Mean Difference** | | | |
| **Measure 1** | |  | | **Measure 2** | | **t** | | **df** | | **p** | | **Mean Difference** | | **SE Difference** | | **Lower** | | **Upper** | |
| Hours Study per Week |  | - |  | Hours Exercise per Week |  | 6.869 |  | 65 |  | < .001 |  | 6.152 |  | 0.896 |  | 4.657 |  | 7.646 |  |
|  | | | | | | | | | | | | | | | | | | | |
| *Note.*  Student's t-test. | | | | | | | | | | | | | | | | | | | |

Since the practical confidence interval for for a confidence level of does not contain , at the significance level , there is a significant difference between and .

Because there is a significant difference between and , at the significance level , there is sufficient evidence to support .

Because there is insufficient evidence to support , at the significance level , there is sufficient evidence to answer with “Yes”.

These reasonings confirm the results of the Independent Samples T-Test of . These reasonings may not be valid.

1. **Is there a correlation between and ?**

To address and evaluate a corresponding claim that “there a correlation between and ”, the dataset from Section II is used.

**Addressing and Evaluating**

*Pearson’s Linear Correlation Coefficient*

Pearson’s Linear Correlation Coefficient corresponding to *S*

Pearson’s Linear Correlation Coefficient corresponding to , , is of the same form as , but the summations for are over the indices of all matched pairs for .

*Classical / Frequentist Correlation Analysis, and Requirements*

The only method known to the author for evaluating is a Classical / Frequentist Correlation Analysis, which relies on , the following two hypotheses, and the following four requirements. The alternative hypothesis corresponds to the claim and is . The null hypothesis is . The first requirement is that , associated with , consists of matched pairs of data , which as pairs are independent. The second requirement is that any sample from is a simple random sample. The third requirement is that a scatterplot of versus shows a linear trend. The fourth requirement is that does not contain any erroneous outliers corresponding to .

*Considering the Requirement of Matched Pairs*

consists of matched pairs, each corresponding to a unique student in and belonging to a tuple . These pairs are independent; the pairs may be reordered without changing the correlation analysis.

*Considering the Requirement of Simple Random Samples*

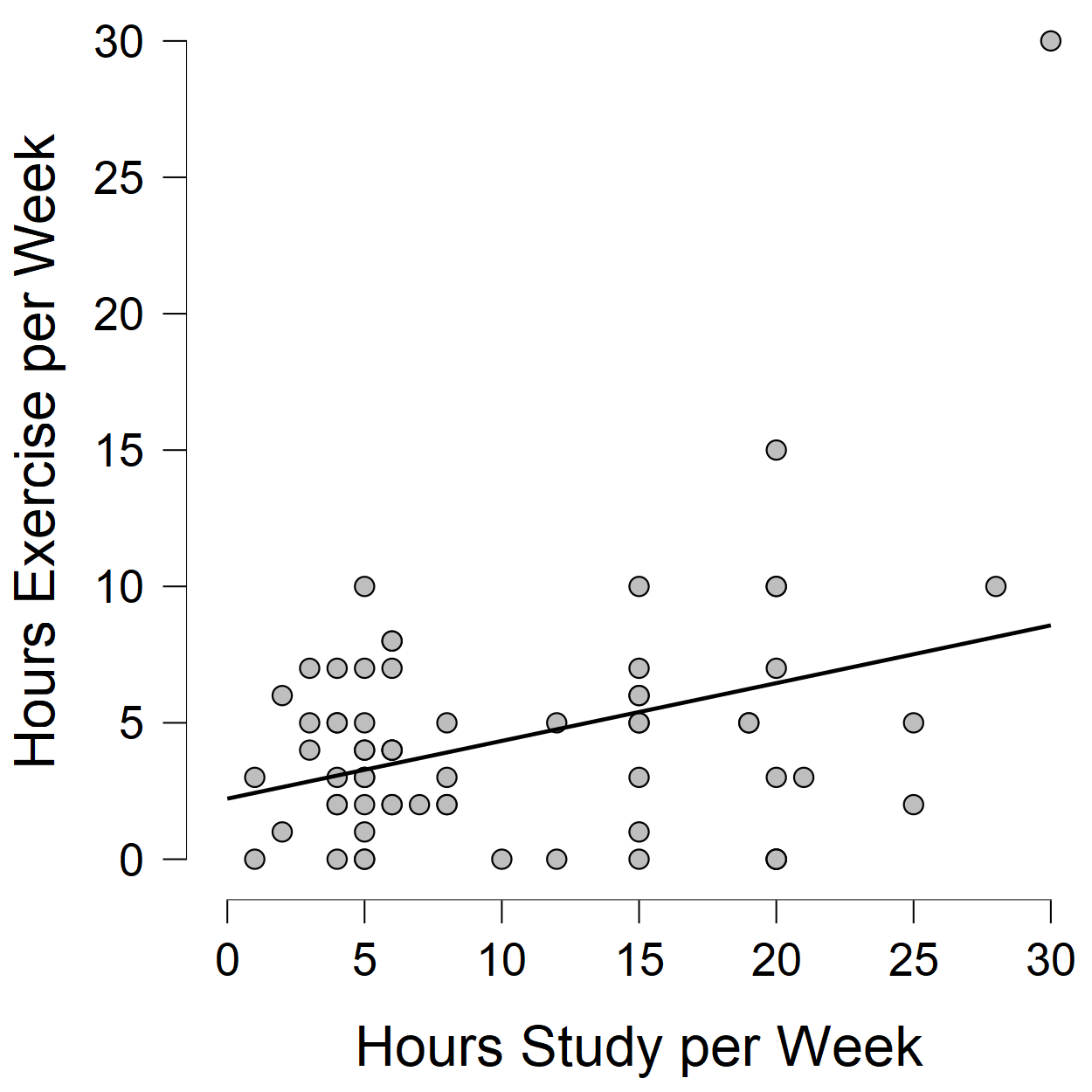
is not a simple random sample; in fact, is a convenience sample. Therefore, while the steps of a Correlation Analysis will be performed, the results of the analysis may be invalid because does not represent *P*.

*Considering the Requirement of a Linear Trend*

The below scatterplot was produced by JASP. A CSV file containing the columns of data “Hours Study per Week” and “Hours Exercise per Week” was imported. A “Classical / Frequentist Correlation Analysis” was performed with variables “Hours Study per Week” and “Hours Exercise per Week”. The title for the scatterplot was reordered to present the dependent variable “Hours Exercise per Week” first.

The below scatterplot shows a linear trend; hours exercise per week seem to gradually increase with increasing .

**Hours Exercise per Week vs. Hours Study per Week**



*Considering the Requirement of Removing Erroneous Outliers*

While the dataset *A* contains outliers, these outliers are not erroneous. Therefore, is met.

Let be the strength of the linear relationship between all and .

Let be the percent of variation in all explained by the linear relationship between all and .

Let be the percent of variation in all not explained by the linear relationship between all and .

Let be the weakness of the linear relationship between all and .

Let the number of degrees of freedom .

Let some

per the null hypothesis.

The test statistic for Pearson’s Linear Correlation Coefficient

is considered to have a corresponding frequency in a Student’s *t* distribution centered at with number of degrees of freedom .

Let be the Student’s *t* Cumulative Distribution Function for a lowest *t*-score , a highest *t*-score , and the number of degrees of freedom for the symmetric Student’s *t* Distribution . Then the probability that a *t*-score is less than or equal to or greater than or equal to , assuming is true,

This hypothesis test was also performed using JASP. A CSV file containing the columns of data “Hours Study per Week” and “Hours Exercise per Week” was imported. A Classical / Frequentist Correlation Analysis was performed with variables “Hours Study per Week” and “Hours Exercise per Week”. Some calculated values are presented below.

| **Pearson's Correlations** | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | |  | |  | | **n** | | **Pearson's r** | | **p** | |
| Hours Study per Week |  | - |  | Hours Exercise per Week |  | 66 |  | 0.354 |  | 0.004 |  |
|  | | | | | | | | | | | |

A significance level is chosen.

Because the probability **p** that a *t*-score is less than or equal to or greater than or equal to , assuming is true, is less than , we reject .

Because we have rejected , at the significance level , the results of our Correlation Analysis are statistically significant.

Because we have rejected , at the significance level , there is sufficient evidence to support .

Because there is sufficient evidence to support , at the significance level , there is sufficient evidence to answer with “Yes”.

However, because the coefficient of determination , only percent of variation in can be explained by the linear relationship between and . Because , the line of best fit presented in the scatterplot above should not be used for prediction.

**Conclusion**

In section I of this paper, a research question of “Do female Statistics-I students at Piedmont Virginia Community College (PVCC) exercise for a different number of hours per week than male Statistics-I students at PVCC?” was addressed. To address the research question, a claim that “female Statistics-I students at PVCC exercise for a different number of hours per week than male Statistics-I students at PVCC” was evaluated. A dataset corresponding to a sample of female Statistics-I students at PVCC was acquired; a dataset corresponding to a sample of male Statistics-I students at PVCC was acquired. A Classical / Frequentist Independent Samples T-Test, which relied on and , an alternative hypothesis equivalent to and a contradictory null hypothesis , and four requirements– which were not met –was conducted. While the results of the hypothesis test may not be valid because the requirements were not met, the results addressed the research question by suggesting that, at the significance level , there is insufficient evidence to answer with “Yes”. A practical confidence interval for the difference between the mean number of hours exercise per week for and the mean number of hours exercise per week for , for a confidence level of percent, was calculated. Reasonings based on the confidence interval were consistent with the results of the Independent Samples T-Test of .

In section II of this paper, a research question of “Is the mean number of hours per week that Statistics-I students at PVCC, in population , study greater than the mean number of hours per week that the students exercise?” was addressed. To addressed this research question, a claim that “The mean number of hours per week that Statistics-I students at PVCC, in population , study is greater than the mean number of hours per week that the students exercise” was evaluated. A dataset corresponding to a sample of Statistics-I students at PVCC was acquired. A Classical / Frequentist Paired Samples T-Test, which relied on , an alternative hypothesis equivalent to and a contradictory null hypothesis , and four requirements– which were not met –was conducted. While the results of the hypothesis test may not be valid because the requirements were not met, the results addressed the research question by suggesting that, at the significance level , there is sufficient evidence to answer with “Yes”. A practical confidence interval for the mean , for a confidence level of percent, was calculated. Reasonings based on the confidence interval were consistent with the results of the Paired Samples T-Test of .

In section III of this paper, a research of “Is there a correlation between number of hours per week that Statistics-I students at PVCC, in a population , study, and the number of hours per week that the students exercise?” was addressed. To address this research question, a claim that “There is a correlation between number of hours per week that Statistics-I students at PVCC, in a population , study, and the number of hours per week that the students exercise” was evaluated. The dataset from section II was used. A Classical / Frequentist Correlation Analysis, which relied on , an alternative hypothesis equivalent to and a contradictory null hypothesis , and four requirements– which were not met –was conducted. While the results of the hypothesis test may not be valid because the requirements were not met, the results addressed the research question by suggesting that, at the significance level , there is sufficient evidence to answer with “Yes”.

This paper provides context and motivation for conducting a study to consider whether, on average, female students at PVCC exercise for a different number of hours per week than male students at PVCC; whether, on average, the number of hours per week that students at PVCC study is greater than the number of hours per week that students at PVCC exercise; and whether there is a correlation between the number of hours per week that students at PVCC study and the number of hours per week that students at PVCC exercise.