VIDEO GAME SURVEY DATA PROJECT

Navin Souda
University of California, San Diego
nsouda@ucsd.edu

Evan Velasco
University of California, San Diego
emvelasc@ucsd.edu

Shu-Ho Yang
University of California, San Diego
shy096@ucsd.edu

Yu-Ming Lin
University of California, San Diego
yul857@ucsd.edu

Andrew Chavez
University of California, San Diego
adchavez@ucsd.edu

1 Introduction

In this investigation we want to analyze a sample survey conducted at UC Berkley intended on better understanding the relationship students have with video games. We will begin by discussing previous literature encompassing this topic, then analyze the data collected to determine if it accurately portrays a relationship.
Every year, 3,000-4,000 students enroll in statistics courses at UC Berkeley. Of these students, nearly half of them enroll in order to fulfill their quantitative reasoning GE requirement. In order to better the quality of instruction, the department created a committee comprised of faculty and students to design a series of computer labs meant to extend student learning by having an interactive component. The idea behind this is that by providing an interactive learning environment, students will be provided with an alternative method for learning the course material and may increase their odds of success.
Since some have linked characteristics of labs to be parallel to those of video games, the committee of faculty/students decided it would therefore be useful to survey the student population on the matter. A study was then conducted in which the department asked questions to understand which aspect of video games students liked. The department then developed a questionnaire, selected students to be sampled and then collected the data.

Research shows a general trend of women playing video games for less time than men. A study conducted by Shirley Matile Ogletree and Ryan Drake in the research journal titled "Sex Roles" found that men where "significantly more likely than women to play video games two or more hours a week."[3] The participants of this study were 206 college students (79 men, 127 women) whose ages mostly ranged between 18 - 25 years old. The data was obtained through a questionnaire administered in April of 2007. The study found that 21% of the women and 68% of the men played two or more hours per week. This difference was significant at the 0.05 level of significance. Additionally, women were proportionally more likely than men to play an hour or less a week.Another study conducted by Kristen Lucas and John L. Sherry published in the journal "Communication Research" found that women "...report less frequent play, less motivation to play in social situations, and less orientation to game genres featuring competition and three-dimensional rotation." [1]

2 Data

Out of 314 students in Statistics 2, Section 1, during Fall 1994, 95 students were selected at random to participate in the survey. In order to minimize response bias and honor the individual's privacy, the survey participants were kept anonymous.

It should be noted that the response encoded as 99 signifies that a question was answered improperly. Additionally, those who never played a video game or did not like them at all were asked to skip many of the questions.

After the first survey was conducted, those same students were then asked to partake in a follow up second survey. For each question participants were able to select up to three responses. The students were asked what type of video games they typically enjoyed to play such as action, adventure, simulation sports, or strategy. They were also asked which aspects of video games they enjoyed such as graphics/realism, relaxation, eye/hand coordination, mental challenge, feeling of mastery, and boredom. Additionally, those who responded that they did not like video games were asked what reasons they had such as too much time, frustration, loneliness, too many rules, cost, boredom, and lack of friends in the gaming community. Finally, the participants were asked for general information such as their age and sex. One key detail regarding the follow up survey is that not everyone from the first survey responded to the second questionnaire. The students who did not complete the follow up survey were either those who answered that they have never played video games, or indicated they had no interest in video games. Additionally, because of the small sample size, there are some limitations to this data. The data collected came from one class section, from one school, in one semester. This can therefore mean a student's data may have a small degree of correlation with another student's data. If the sample size was larger, taken from multiple class sections, over multiple semesters/years, then this degree of correlation may correct itself. It is then difficult to generalize the results from this survey to a wider population.

3 Background

For this survey, all of the participants were undergraduates enrolled in an introductory probability and statistics course during the 1994 fall semester. When picking the participants, the committee used a list of students who had taken the second exam in a lecture. Each student was then assigned a number at random from 1 to 314 to uphold anonymity, and to aid in the selection process. Once a number was assigned to each student, a pseudo random number generator selected 95 numbers with the parameters of 1 to 314. The survey was then administered a week after the second exam was taken by the students. To maximize the number of responses, a three-stage system of data collection was implemented. Data collectors attended both discussion sections on dates when the exams were to be handed back in order to make contact with those selected survey participants. If any students could not be reached via the discussion sections, then they were located during lecture. The participants were then briefed on the nature of the survey and reassured their anonymity would be honored. Of the 95 initially selected students, 91 of them successfully completed the survey.

4 Investigation

4.1 Proportion Estimate of Students Who Played Video Games

First, we set all the students who had played a video game in the week prior to the survey equal to "1" in the time column, and the students who had not remain equaling to "0". Then, we can use the equation of confidence interval (CI). To do so, we first calculate the sample mean for the point estimate, which is 0.37363. In calculating the standard deviation, we must keep in mind the variance correction factor induced by our large sample size relative to the population. Thus, we calculate the standard error for the sample proportion to be 0.043. Then, we can use this to calculate two confidence intervals: one using the normal distribution quantiles, and one using the t-distribution quantiles. Finally, we will generate a confidence interval using bootstrap. These are all listed in the table below.

Type of Interval	Lower Bound	Upper Bound
Normal Quantile	0.2893996	0.4578531
Student t Quantile	0.2882518	0.4590010
Bootstrap Percentile	0.2857143	0.4615385

Figure 1: Various interval estimates for the proportion of students who played video games the week prior to the exam.

We can see that these are all fairly close together. The normal approximation gives the smallest CI, the student t the next smallest, and the bootstrap the largest confidence interval.

4.2 Amount of Time Spent Playing Video Games in the Week Prior

Category	Daily	Weekly	Monthly	Semesterly
Count	9	28	18	23

Figure 2: Counts of responses to how frequently students play. There were 13 no responses.

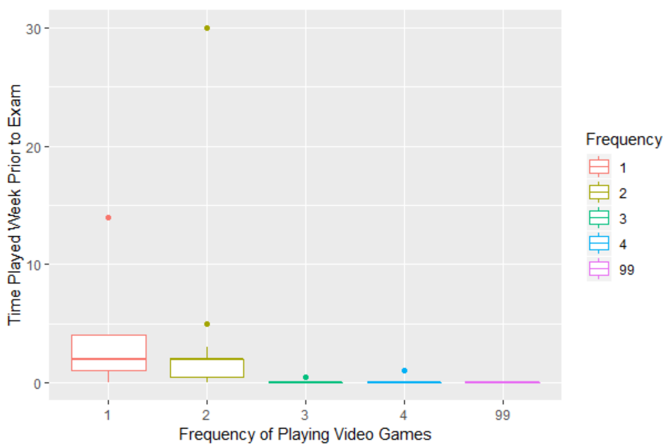

Figure 3: Boxplot showing distribution of time played based on frequency. 1 corresponds to daily, 2 to weekly, 3 to monthly, 4 to semesterly, and 99 for those who did not respond.

Looking at the boxplot above, we see that the median time played in the week prior was 2 hours for both people who claimed to play daily and those who played weekly. For the other groups, including the students who did not respond, the median was 0. In fact, there was only 1 student who said they played on a monthly basis that had played in the week before the test, and they only played half an hour. Similarly, there was one student who said they played semesterly and only played an hour in the week before the test. This makes sense, as we would expect people who play so sporadically to likely not have played in any given week.
When we consider people who play weekly or daily, we start to see more variation. We see a huge outlier in the students who played weekly, where one student claimed to play 30 hours in the week before the exam. We suspect that this was a misreported or misread data point, as one would not expect a student who plays once a week to somehow play 30 hours within a week. It is possible the student meant they played 30 minutes, but in the end we cannot be sure. For both students who played daily and those who played weekly, the median time of play was two hours. Those who played daily however, had a higher lower and upper quartile of time played, which one might expect as they play more often.
Finally, we consider the fact that the survey asks the amount of time played in the week before their exam. It's not unreasonable to believe that students would play less video games in the week before an exam in order to focus more on studying, and thus our data may be affected by that. It's possible if the survey was given at a different time, the reported times played would be higher on average.

4.3 Estimating Time Played

We now want to estimate the average time played the week prior to the exam. We find that the sample mean is 1.243 hours. We then use the bootstrap to make confidence intervals for the mean. One way will be the t-statistic confidence interval with the bootstrap estimate for the standard deviation, the other will simply be the percentile bootstrap confidence interval. Our reasoning for this will be covered in the Theory section below. Thus, performing all computations, our findings for the confidence intervals are shown in the table below. Note that the percentile interval is more conservative (i.e. larger) than the normal quantile interval.

Type of Interval	Lower Bound	Upper Bound
Student t Quantile	0.7272605	1.7584538
Bootstrap Percentile	0.6801099	1.8720604

Figure 4: Various interval estimates for the proportion of students who played video games the week prior to the exam.

4.4 Do Students Enjoy Video Games?

Now that we have investigated how many students played video games in the week prior to the survey and for how long they played, we can ask the question "Do students like video games?" To answer this question we must first classify what it means to "like" or "dislike" video games. A brief overview of student survey results regarding preference of video games is shown below.

Figure 5: Bar Plot Overview of Student's Video Game Preferences

For the purpose of this study, students who responded with "like"=2 or "like"=3 will be classified as enjoying video games and those who answered the other options will be considered to not enjoy video games. In addition to this, those who had never played a video game ("like"=5) will be taken out of the total. The distribution of "like" vs "dislike" is shown below.

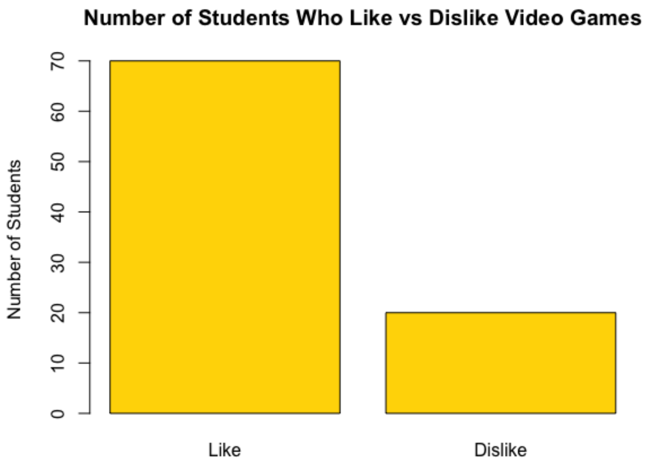

Figure 6: Bar Plot of Students who "Like" vs "Dislike" Video Games

A quick graphical analysis shows us that the proportion of students who like video games is much higher than those who do not like them. The exact proportion of those who like video games is 0.7667 and those who dislike video games is 0.2333. We can construct a 95% confidence interval to determine the accuracy of our claim that there is a qualified majority (greater than 2/3) of students who enjoy video games. The percentile bootstrap confidence interval may be seen below.

Category	Bootstrap
Like Video Games	(0.6854, 0.8539)
Dislike Video Games	(0.1348, 0.3146)

Figure 7: Confidence Interval for Proportion of Students who Like or Dislike Video Games

We will now examine what characteristics cause a student to like or dislike video games. In order to do this we will use a classification and regression tree. Preferences in the tree are categorized such that "like"=1 and "dislike"=0. The results are shown below.

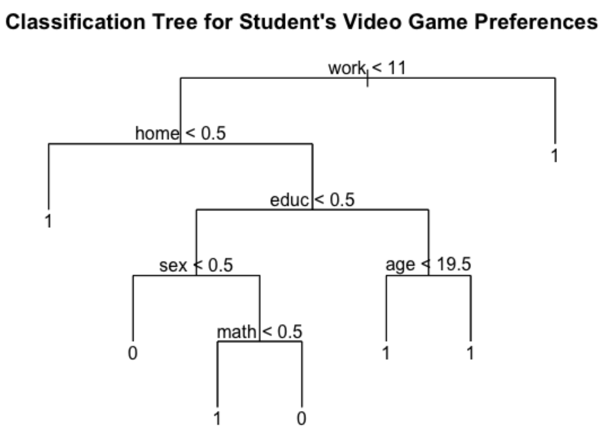

Figure 8: Classification and Regression Tree for Student Video Game Preferences

There are some interesting conclusions we may draw from this classification tree. How much a student works seems to be a large indicator of video game preference, with students working more than 11 hours classified as enjoying video games. Moving down the tree we see that factors such as age do not have a classified difference between younger and older students. When it comes to sex, women are classified as more likely to not enjoy video games. Interestingly enough, men are not automatically put into the like category. It seems that how much male students enjoy math plays a role in their preferences, with those who enjoy math more likely to dislike video games and those who dislike math more likely to dislike video games. So in conclusion, our list of important features would contain the amount a student works per week, whether or not they have a computer at home, and whether or not they think games are educational.

4.5 Comparison of Video Game Appreciation Among Groups

We would like to see if there is a difference between those who like video games and those who do not when considering different groups of people. Specifically, we will assess the proportions of those who like and dislike video games between male and female students, students who work and students who do not, and finally between students who own a PC and those who do not.

4.5.1 Male vs. Female

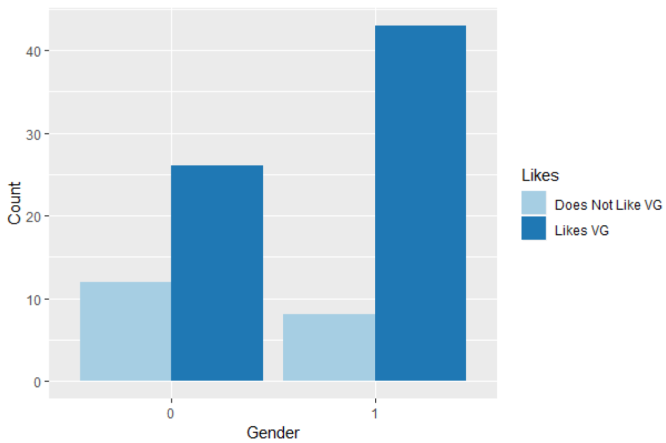

Figure 9: Side-by-side comparison of like vs. dislike between male and female students. 0 represents female, 1 represents male.

Looking at the barplot above, it seems that there is a discrepancy between the proportion of female students that like video games and male students that like video games: specifically, the proportion of male students that like video games is greater. Numerically, we find that out of 51 male students, 43, or 84.3% like video games, whereas out of the 38 female students, only 26, or 68.4% like video games. We can perform multiple hypothesis tests to check this. With Fisher's exact test, we get a p-value of 0.122. It is generally known that Fisher's exact test tends to be conservative, and so we may choose to use a less conservative test. Using a chi-square test for independence (valid because all expected counts are greater than 5), we get a p-value of 0.076, which is still greater than our significance level of 0.05. Thus we say there is not significant evidence that like or dislike of video games is dependent on gender.

4.5.2 Work

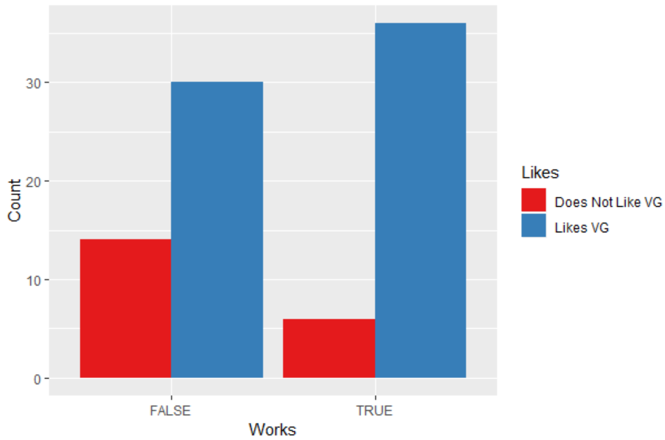

Figure 10: Side-by-side comparison of like vs. dislike between working and non-working students

Once again, the barplot (Figure 6) seems to indicate that students who work tend to enjoy video games at a higher rate than those who do not. We find that out of 44 students who do not work, 30 - that is, 68.2%, like to play video games. On the other hand, 36 out of 42 (85.7%) working students like video games. Once again deferring to statistical tests, however, we find that Fisher's test gives us a p-value of 0.074, and the less conservative chi-square test gives us a p-value of 0.054, meaning we do not have significant evidence that like or dislike of video games depends on whether or not a student works as well.

4.5.3 Owning a PC

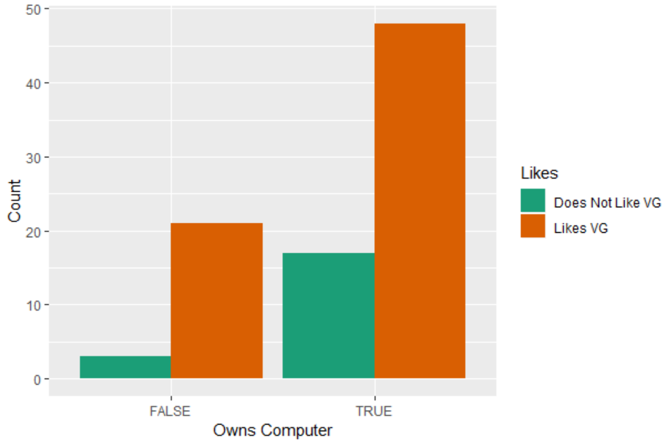

Figure 11: Side-by-side comparison of like vs. dislike between students who don't own a PC and students who do.

First, it is important to note that there is a large difference between the number of students who own a PC and those who do not - 65 and 24 respectively. Out of the 65 who do own a PC, 48, or 73.8% said they like video games. Of the 24 who do not own a PC, 21, or 87.5% like playing video games. To test whether like and dislike for video games differs based on computer ownership, we once again turn to Fisher's exact test and the chi-square test. Fisher's test gives us a p-value of .2537, where as the chi-square test gives a p-value of 0.1708. Thus, just as we saw in the previous cases, we do not have statistically significant evidence to claim that like or dislike of video games depends on whether or not one owns a computer.

4.5.4 Conclusions

We can see that while upon first glance, especially graphically, it seems that there are in fact differences among groups that affect whether or not one likes video games. We find that a higher proportion of men like video games than women, a higher proportion of working students like video games than those who do not, and, somewhat surprisingly, a higher proportion of students who do not own computers like video games compared to those who do not.
Taking a closer look, however, and specifically using hypothesis testing, we find that none of these differences are statistically significant. In the case of PC ownership, this may have something to do with the large difference in group sizes. For the others, the proportions are simply too close to make a strong claim.

4.6 Grade Distribution

Now we will look at the distribution of grades in the class and see whether it matches the desired distribution of grades, specifically 20% A's, 30% B's, 40% C's and 10% D's or lower. Using our sample size of 91, we display the expected and observed counts in the table below. Note that all expected counts are greater than 5, so we may use a chi-square test for goodness of fit. Specifically, we test H_0: The grades follow the desired distribution vs H_1: They do not follow the desired distribution. Our chi-square statistics comes out to 62.6. This has 4 - 1 = 3 degrees of freedom. This is much greater than the 95% quantile of the corresponding chi-square distribution, which is 7.81, so we reject the null, and decide the desired distribution is not followed.

Grade	A	B	C	D/F
Observed	31	52	8	0
Expected	18.2	27.3	36.4	9.1

Figure 12: Confidence Interval for Proportion of Students who Like or Dislike Video Games

4.7 Time Women Spend Playing Video Games

Previous research has seen a variety of results regarding women's video game playing habits. In this section we will be examining the amount of time women in the survey spent playing video games. The minimum amount of time women spent playing video games was 0.0 hours, while the max was 14.0 hours. The mean for time women spent playing video games was 0.75 hours. A histogram showing the overall density of the data is shown below.

Figure 13: Histogram of Women Time

Looking over the graph we can see that our data is heavily skewed right. The graph shows us that the majority of women appear to play video games less than two hours during the week prior to the survey. However, this is just a visual observation from one sample. To gain a better understanding of the range of values the mean time could take, we construct a 95% confidence interval using bootstrap. The results of the confidence interval are displayed below.

Type	Confidence Interval
Bootstrap Percentile	(0.1641, 1.7105)

Figure 14: Confidence Interval for Proportion of Students who Like or Dislike Video Games

We can see that our findings are in line with the previous research on the topic. While we are not specifically testing a hypothesis that the average amount of time women play is less than 2 hours, we can see that the upper bound on our confidence interval is below 2. For this reason we may believe that women on average seem to be playing video games for less than 2 hours a week.

5 Theory

5.1 Estimators and Confidence Intervals for Survey Data

When utilizing survey data, we shouldn't immediately assume that we use the same estimators for sample mean and sample variance, as these estimators and the confidence intervals that follow from them rely on the data to be i.i.d. While with a theoretically infinite population size, the probability that a certain unit gets picked does not change enough to make a difference, with a finite population size and a sample size large enough, the difference does matter. It is quite clear that units that are picked in the sample later have a higher chance of being chosen - the size of the population reduces by 1 each time another unit is chosen for the sample. Thus, we must reconsider our estimators.
We will be considering the expected value and variance of the sample mean. Let the population size be N, sample size be n, population average be μ, and let $I(j)$ be the j-th number chosen from a list of 1,...,N without replacement. Then x_i is the i-th member of our population, and $x_{I(j)}$ for $1 \leq j \leq n$ refers to the j-th member of our sample, which was unit $I(j)$ of the population. Then we have

$$\begin{aligned}
E[x_{I(j)}] &= \sum_{i=1}^{N} x_i P(i = I(j)) \\
&= \sum_{i=1}^{N} x_i \frac{1}{N} \\
&= \mu
\end{aligned}$$

Now let $\bar{x} = \frac{1}{n}\sum_{j=1}^{n} x_{I(j)}$. Then it follows that

$$\begin{aligned} E[\bar{x}] &= E[\frac{1}{n}\sum_{j=1}^{n} x_{I(j)}] \\ &= \frac{1}{n}E[\sum_{j=1}^{n} x_{I(j)}] \\ &= n\mu\frac{1}{n} \\ &= \mu \end{aligned}$$

Thus, our usual $\bar{x}$ is an unbiased estimator for the sample mean. We will find, however, that the same does not hold for the variance.
Let σ^2 be the population variance. Then $Var(x_{I(j)}) = E[(x_{I(j)} - \mu)^2] = \frac{1}{N}\sum(x_i - \mu)^2 = \sigma^2$. Now, consider Var($\bar{x}$).

$$\begin{aligned} Var[\bar{x}] &= Var[\frac{1}{n}\sum_{j=1}^{n} x_{I(j)}] \\ &= \frac{1}{n^2}Var[\sum_{j=1}^{n} x_{I(j)}] \end{aligned}$$

Due to the dependence of the $x_{I(j)}$, we get

$$\begin{aligned} &= \frac{1}{n^2}(\sum_{j=1}^{n} Var[x_{I(j)}] + \sum_{j=1, j\neq k}^{n} Cov[x_{I(j)}, x_{I(k)}] \\ &= \frac{1}{n} + \frac{n-1}{n}Cov[x_{I(1)}, x_{I(2)}] \end{aligned}$$

Using the knowledge that $Cov[x_{I(1)}, X_{I(2)}] = -\frac{\sigma^2}{N-1}$, our final value for $Var[\bar{x}] = \frac{1}{n}\sigma^2\frac{N-n}{N-1}$. Note that as N approaches infinity, this value approaches the usual value for the variance of $\bar{x}$, but otherwise it is less than the usual variance. Then, we can derive the unbiased estimator for $Var(\bar{x})$, $\frac{s^2}{n}\frac{N-n}{N}$, where s^2 is the usual estimator for variance. The variance estimator is slightly different for proportions - the n in the denominator is replaced by an $n-1$, and $s^2 = \bar{x}(1-\bar{x})$. Now, we simply build confidence intervals for our sample mean or sample proportion as usual, but using these smaller variance estimates, giving us a tighter confidence interval.
Note that when creating a confidence interval from a known distribution, we have more than one option: we can either use quantiles from a t-distribution with n-1 degrees of freedom, or we can use a normal distribution. In general, the t-distribution may be preferred when we have to estimate the variance, as we do in our case. However for large values of n, the t distribution converges to the normal, so either may be appropriate.

5.2 Bootstrap for Survey Data

Bootstrap for survey data is similar to the bootstrap as usual, but with a key difference in the building of the bootstrap samples. We take our initial sample data, counting the number of each value of the variable we are looking at (for example, time spent playing video games in a week). We then multiply these counts by N/n, rounding where necessary, to make a bootstrap population. Then we compute, say, bootstrap mean values by taking samples of size n from this bootstrap population, and computing the means of those. Then we can compute either percentile confidence intervals using just the percentiles of these bootstrap means, or we can use the computed standard error of these bootstrap means as well as quantiles from a known distribution (e.g. normal or Student t distribution) to form confidence intervals as usual.

5.3 Justifications for Scenarios 1 and 3

5.3.1 Scenario 1

The normal quantile and t quantile intervals were computed using the normal equations for confidence intervals, and the 0.025 and 0975 quantiles of the respective distributions. The percentile bootstrap interval was used by formulating many bootstrap samples as above, and then finding the 0.025 and 0975 quantiles of the sample means of those samples. We can use the normal and Student quantile intervals because of the Bernoulli approximation to the normal distribution, meaning for large sample size, our proportions (i.e. the sample mean) is normally distributed. Looking at the histogram and quantile-quantile plot, there seems to be some skew, but a Kolmogorov-Smirnov test comparing to normality gives us a p-value of 0.06, implying the difference is not statistically significant. Thus, our quantile confidence intervals are valid.

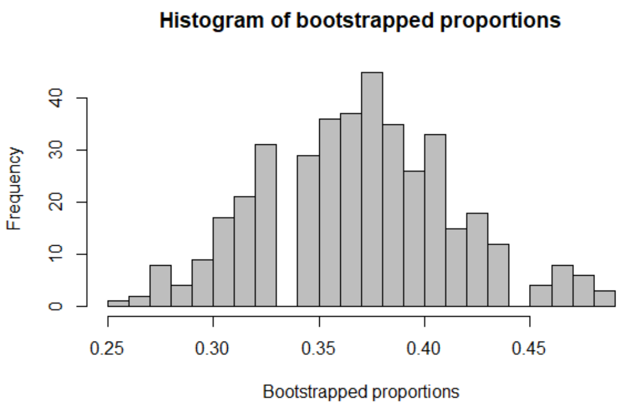

Figure 15: Histogram of bootstrap proportions

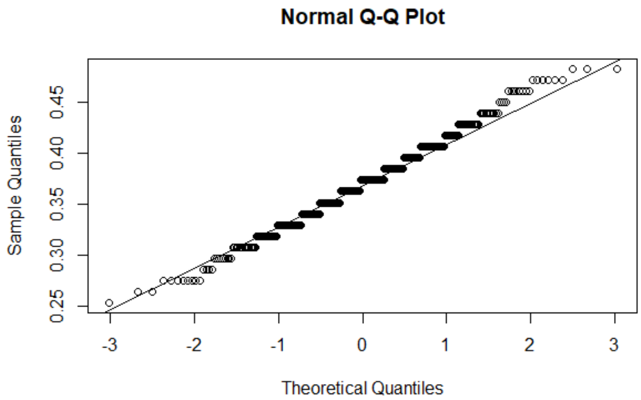

Figure 16: Quantile-quantile plot of bootstrap proportions

5.3.2 Scenario 3

Similarly to the previous part, in order to use the Student quantile interval, our sample mean must have a normal distribution, which we test, like before, using a histogram, a q-q plot, and a Kolmogorov-Smirnov test. The quantile-quantile plot and the histogram seem slightly skewed, but the K-S test tells us that it is not a statistically significant departure from normality, with a p-value of 0.467.

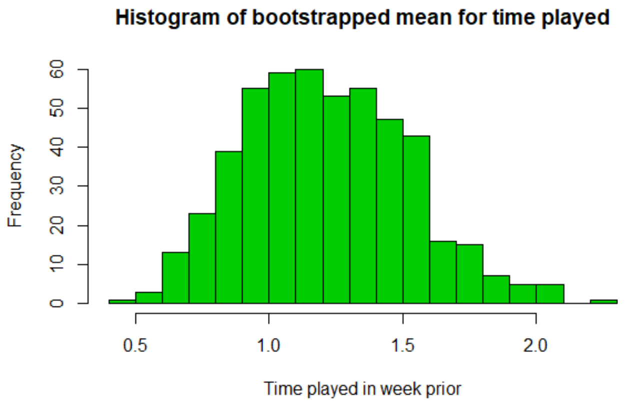

Figure 17: Histogram of bootstrap means for time played

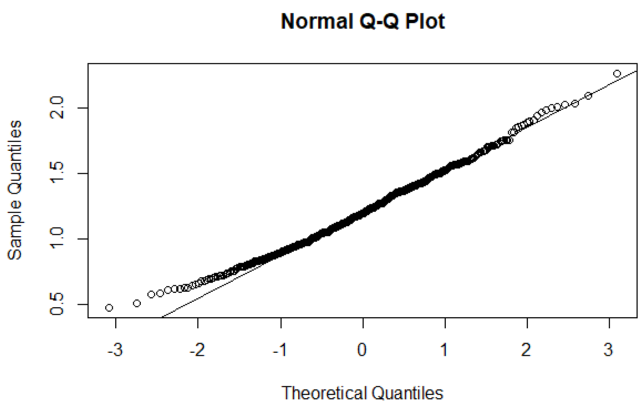

Figure 18: Quantile-quantile plot of bootstrap means to normal distribution

5.4 CART

Classification and regression trees, or CART, is a powerful machine learning tool which helps determine the "most important" variables in a particular data set. In this instance we used CART analysis to determine what variables were the biggest indicators to whether or not a student enjoyed video games. Since the variable that indicates video game preferences is 0-1 binary, we use a classification tree. Had the variable we were analyzing been numerical or continuous, a regression tree would have been better suited. Essentially, a classification tree splits the data based on homogeneity. This involves what many call a "purity criterion," which involves categorization based on similar data in order to make it more "pure." [2] A visual representation of the CART process created by Majid (2013) is shown below.

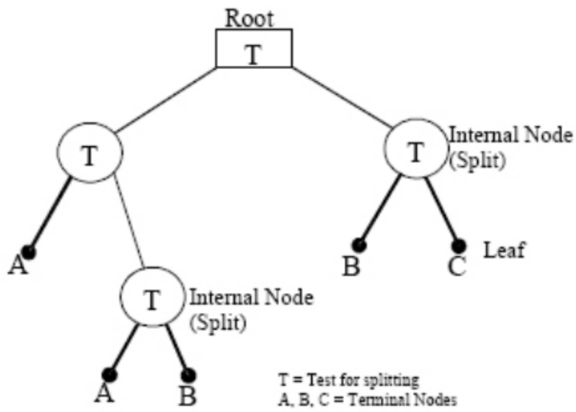

Figure 19: Structure of CART

The premise of CART investigation is quite simple. Given factors $x_1, x_2, x_3, ..., x_n$ in domain x we want to predict the outcome of interest, Y [2]. Figure 12 shows the general picture of cart with the domain of the factors associated with the Y value in descending order. Each time the data splits is called a node, with the final terminal node often being referred to as a leaf. Each leaf implies that that after this split, further splitting of the data does not explain enough of the variance to be relevant in describing Y. Splits are chosen by whichever variable would increase purity the most. Thus the splits towards the top of the tree, and those that lead directly to a classification are variables that one might consider most important. [2]

5.5 Comparing Likes/Dislikes

In Section 4.4, we compare the counts and proportions of those who like video games and those who do not within specific groups. We split these groups up by gender, status of employment, and computer ownership. To compare these groups graphically, a side-by-side barplot was used. Using a side-by-side

barplot, we can visually compare the ratio of likes to dislikes between different categories, as well as compare the total number of people in each category.
We then moved on to hypothesis tests to see whether or not these counts implied a significant difference in those who liked or disliked video games. As stated earlier, all of these tests indicated that there were not statistically significant differences in liking video games when controlling for gender, employment or computer ownership. We used two different types of hypothesis tests: Fisher's exact test and the chi-square test for independence.
Fisher's exact test tests for an association between two categorical variables-that is, it tests whether the state of one of the variables affects the counts of the other. The test determines an exact p-value by fixing the marginals of the table (i.e. the row and column counts) and checking, out of all possible tables with those marginals, how many had counts that were at least as extreme as the ones shown in our data. Fisher derived that this followed a hypergeometric distribution, from which all calculations follow. Unlike the chi-square test, which provides only an approximation of a p-value that may be unreliable at small sample sizes or with unbalanced categories, Fisher's test gives us an exact value[4]. On the other hand, Fisher's test, while giving an exact p-value, may not actually reject at the nominal (0.05) level, because of the discrete nature of the count data. That is, looking at the proportion of tables at least as extreme may never be able to give you a p-value of 0.05 depending on the counts in the table. Thus, it is considered conservative and other methods, such as the chi-square test are considered.
Thus, we will now consider the chi-square test for independence. This test has one main assumption to be valid, and that is that expected count values are greater than 5 for all bins of our data. In our case, the expected counts are determined by $\frac{\text{(Row size}*\text{column size)}}{\text{(Total size of sample)}}$. Note that while expected counts were not shown for brevity, all of our expected counts in Scenario 5 satisfied this condition. One example is shown in the tables below (Figure 20 and 21), when splitting between male and female students. The chi-square test statistic is calculated by

$$\sum_{i=1}^{l}\sum_{j=1}^{m}\frac{(E_{i,j}-O_{i,j})^2}{E_{i,j}}$$

where $E_{i,j}$ are the expected counts as calculated above, and $O_{i,j}$ are the observed counts from our data. This statistic, as stated previously, has asymptotically a chi-square distribution, so the p-values we get are approximations. We then reject for large values of our test statistic, which give small values of p.

Category	Male	Female	Row Total
Likes	12	8	20
Dislikes	26	43	69
Column Total	38	51	89

Figure 20: Actual counts for 2x2 contingency table showing Male and Female students who like or dislike video games

Category	Male	Female	Row Total
Likes	8.54	11.46	20
Dislikes	29.46	39.54	69
Column Total	38	51	89

Figure 21: Expected counts for 2x2 contingency table showing Male and Female students who like or dislike video games

5.6 Women's Video Game Play Time

In Section 4.6 we examined the video game play time of women. We estimated the amount of time they spent playing video games through using a confidence interval. To obtain this interval we used the percentile bootstrap. The decision to use percentile bootstrap was made in order to ensure the accuracy of our findings. The "time" data for women was not normally distributed. A quick look at the skewness and kurtosis of the data tells us that we were not working with a normal distribution. The "time" category for women had a skewness of 4.62 and a kurtosis of 25.08. Comparing these to the expected results from a normal distribution where skewness is 0 and kurtosis is 3, we see a large difference. To further confirm this conclusion we used a Kolmogorov-Smirnov test to see if our data differentiated significantly from a normal distribution. The test resulted in a p-value of 2.614e-05, which causes us to reject the null hypothesis that our data is normally distributed at the 0.05 level of significance. Finally, we used a Q-Q plot to graphically show the difference.

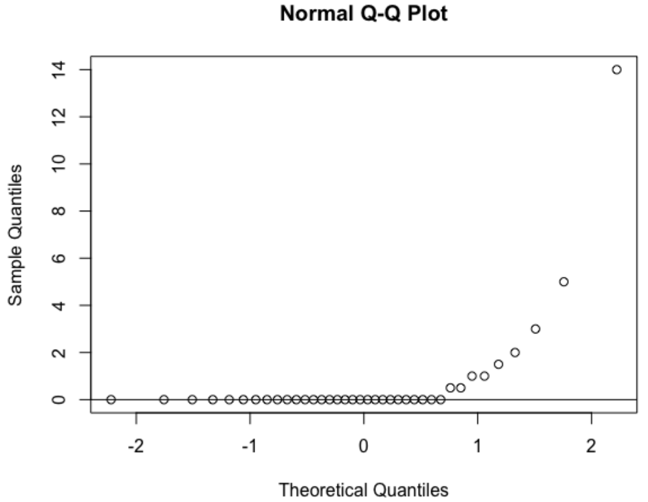

Figure 22: Quantile-quantile plot of "time" variable for women compared to a normal distribution

Due to this, we could not use a t statistic to test a hypothesis regarding women's video game play time. The next most viable option would be to use bootstrap to create a t statistic and conduct a hypothesis test using this data. Unfortunately, we run into a similar issue. In order to properly conduct this test we need our bootstrap means to be normally distributed. Examining the distribution of our bootstrap means shows us that they are not in fact normal. We once again look at the skewness and kurtosis. Our bootstrap had skewness of 0.83 and kurtosis of 3.85, indicating that it deviates from a normal distribution. The Kolmogorov-Smirnov test gives us a p-value of 0.028, showing up this data is not normal. Finally, the Q-Q plot confirms our findings.

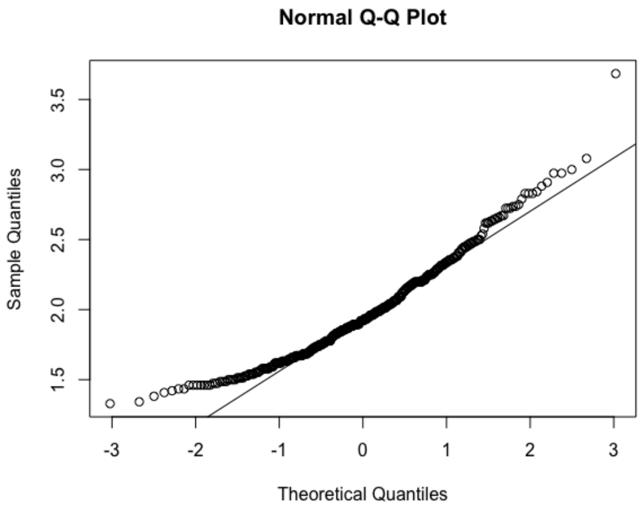

Figure 23: Quantile-quantile plot of "time" bootstrap means to normal distribution

Due to this, we could not use a t statistic to test a hypothesis regarding women's video game play time. Due to the non-normal nature of our data, the best test we can use is a percentile bootstrap to find a confidence interval. We do this by first creating an empty list to which we will store our bootstrap means. The "time" data for women is then re-sampled with replacement. We take the mean of this resampled data, and then continue with the next iteration of the bootstrap. After this has concluded, we use the 0.025 and 0.975 quantile points to find a 95% confidence interval for our data.

References

[1] K. Lucas and J. L. Sherry. Sex differences in video game play:: A communication-based explanation. *Communication Research*, 31(5):499–523, 2004. doi: 10.1177/0093650204267930. URL `https://doi.org/10.1177/0093650204267930`.

[2] J. Morgan. Classification and regression tree analysis. May 2014. URL `https://www.bu.edu/sph/files/2014/05/MorganCART.pdf`.

[3] S. M. Ogletree and R. Drake. College students' video game participation and perceptions: Gender differences and implications. *Sex Roles*, 56 (7):537–542, Apr 2007. ISSN 1573-2762. doi: 10.1007/s11199-007-9193-5. URL `https://doi.org/10.1007/s11199-007-9193-5`.

[4] F. Yates. Tests of significance for 2 × 2 contingency tables. *Journal of the Royal Statistical Society. Series A (General)*, 147(3):426–463, 1984. ISSN 00359238. URL `http://www.jstor.org/stable/2981577`.