Analysis of a Video Game Survey

Ryan Cooper, Duy Dao

May 18, 2017

1   Introduction

Every year, around 1,500-2,000 students enroll in an introductory statistics course at UC Berkeley. The UC Berkeley administration is designing a series of computer labs for the students to learn concepts from their statistics class through a virtual learning environment. The purpose of this is to provide an alternative method for students to learn the material. These labs are heavily linked to video games. To develop improved teaching methods though the use of video games, a study was conducted by students in an advanced statistics class.

Out of 314 students in UC Berkeley's 1994 Statistics 2 class, 95 students were randomly selected for the survey and 91 students completed the survey. We first conduct exploratory data analysis on the sample to understand what fraction of students play video games and to what extent they do. To better develop the virtual environment for the students, we next analyze the reasons for why students like and dislike playing. We further compare the differences between the students who like to play with the students who don't. Lastly, we make recommendations based on our findings for how to structure the games to be integrated into a statistics lab.

2   Analysis

The survey can be divided into three different components. The first part of the survey asks the students how many hours they spent playing video games in the week prior to the survey, how frequently they play video games and to what extent they like to play video games. The second part of the survey asks the students what type of video games they like and reasons why they like and dislike video games. The last part of the survey asks the students basic identifying information like gender, age, if they own a computer, etc.

2.1   Time Spent Playing Video Games Prior to the Exam

In this section we look at the time played by the students in the week prior to the exam and compare it with the reported frequency. As a first step, we estimate the proportion of students that played video games in the week prior to the survey. The point estimate for the proportion is 0.374. In this survey the sampling fraction (sample size divided by

population size) is 0.29 and therefore we cannot ignore the finite population correction factor. The estimated standard error in this case is 0.043. The approximate 95% confidence interval obtained from the Central Limit Theorem is (0.289, 0.458). Due to simple random sampling the measurements are dependent, so we simulate the sample statistic to check the approximate normality assumption used in creating the confidence interval. The simulated distribution of the proportion is obtained by creating a bootstrap population of size 314 and recording the proportions for a large number of simple random samples from the bootstrap population. Figure 1 shows the Normal-quantile plot of the simulated distribution of the proportion against the standard normal distribution. The approximate normality is acceptable and justifies the approximate 95% confidence interval.

Next we consider estimating the population average for time played in the week prior to the exam. The point estimate for the average is 1.243 with finite population corrected standard error 0.334. Similar to the proportion interval, we first create a bootstrap population and estimate the average times on a large number of simple random samples from the bootstrap population. Figure 2 shows the normal-quantile plot for this distribution. We accept the approximate normality and the resulting 95% confidence interval (0.589, 1.897).

Lastly we investigate how the time spent playing video games in the week prior to the survey compares with the reported frequency of video game playing. The possibly choices for frequency were daily, weekly, monthly, and semesterly with the number of students in each respective category being 9, 28, 18 and 23. Table 1 shows the distribution of time for each frequency. Note that the outliers in both daily and weekly players make the mean time an unreliable estimator for the center of each distribution. We see that daily players play 1 to 4 hours per week, weekly players play 0.5 to 2 hours per week. And among the monthly and semesterly players, as one would expect, very few played in the week prior to the survey. Note that the estimated time played in the prior week might be lower due to the fact that there was an exam for the class during that week.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Daily	0	1.0	2	4.44400	4	14.0
Weekly	0	0.5	2	2.53900	2	30.0
Monthly	0	0.0	0	0.05556	0	0.5
Semesterly	0	0.0	0	0.04348	0	1.0

Table 1: Summary of the time distribution for each frequency

2.2   Favorable and Unfavorable Video Game Attributes

From the survey we estimate the proportion of students that like playing video games to be 0.767 with a 95% confidence interval of (0.693, 0.84). We conclude that the majority of the students in the class enjoy playing video games.

Those students who answered that they either liked video games very much, somewhat, or not really were asked to select up to three reasons why they like video games. The available choice were graphics, relaxation, eye/hand coordination, mental challenge, feeling of mastery, and bored. Table 2 shows the percentage of students that selected the given
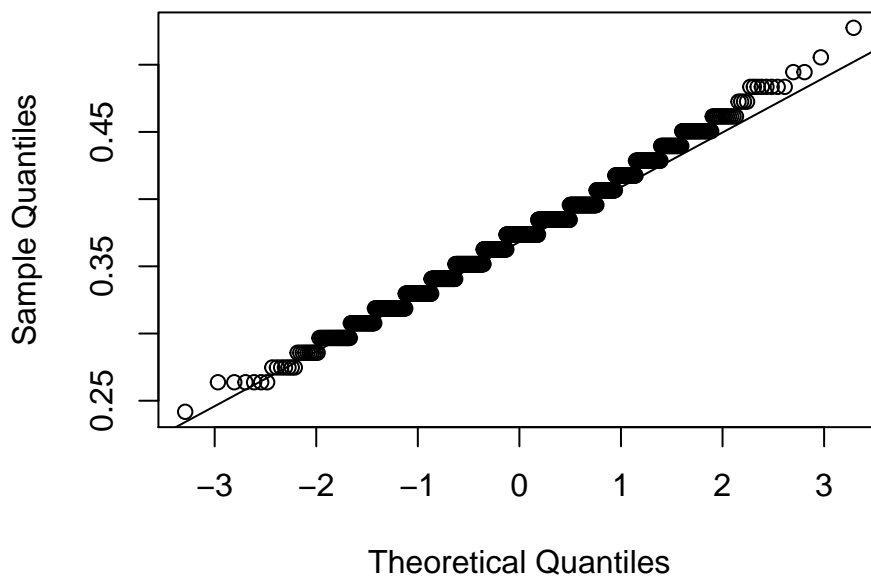
Figure 1: Normal-quantile plot of the bootstrap proportion distribution.

choices. The reasons that received the majority of the vote were relaxation, mastery and boredom. Note that mastery, bored, graphic and challenge each received roughly the same number of votes. Hence we cannot infer which would rank highest among the population.

Next all students were asked to select up to three reasons why they disliked video games. The possible choices were: too much time, frustrating, lonely, too many rules, cost to much, boring, friends don't play, it's pointless. Table 3 shows the percentage of students that selected the given choices. From the responses, the main reasons that students dislike playing video games are that it takes too much time, it costs to much and it's pointless.

2.3   Comparison Between Students that Like Video Games and Students Who Don't

The last part of the survey asks for general information about the respondent, e.g. age and gender. In this section we use these general questions to differentiate those students that like video games and those that don't. To facilitate this process we use a decision tree. More specifically, we classify whether a student likes video games based on their gender, how many hours they work, if they own a computer and if they expect an A in the course. Figure 3 shows the results of the decision tree. The majority of students that like video games fall into one of three types: students who work more than 5 hours (11% misclassification), students who work at most 5 hours and don't own a computer (8% misclassification), and students who work less than 5 hours, own a computer and are male (27% misclassification). The
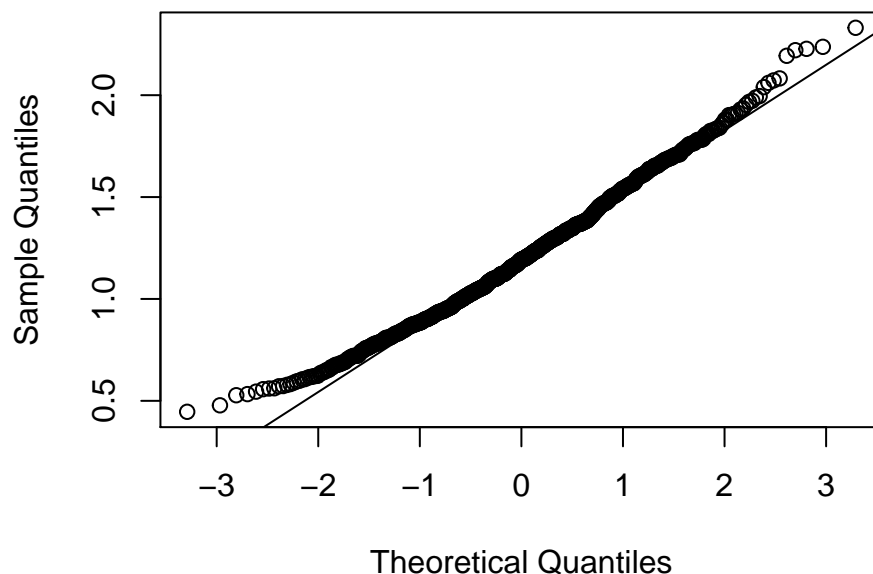
Figure 2: Normal-quantile plot of the bootstrap average time distribution.

only category where the majority of students disliked video games was among students that worked less than 5 hours, owned a computer, were female and expected an A. However, the misclassification rate in this category is high, so this classification is unreliable.

3   Conclusion

After analyzing the data collected from the survey, our team has emerged with suggestions in order to assist the designers of the new computer lab for statistics courses.

   One of the main reasons that students disliked video games was that they are too time consuming. Among the students that play video games on a daily or weekly basis, we found that the median time played in the week prior to the survey was 2 hours. We suggest that the statistics course should have a shortened lab time in order to keep all students engaged in the video game. To match the average amount of time spent playing in the week prior to the survey, the lab time should be in the range of half an hour to an hour and a half. The shortened lab duration will allow these students to feel less overwhelmed with the time commitment. Moreover, the lab should not be too intense since the most important reason students played video games was to relax. The game should also come equipped with multiple levels, in order for the students to feel a sense of accomplishment or mastery in the lab.

	Proportion	95% CI
relax	0.637	(0.554, 0.721)
master	0.275	(0.197, 0.352)
bored	0.264	(0.187, 0.34)
graphic	0.253	(0.177, 0.328)
challenge	0.231	(0.157, 0.304)
coord	0.044	(0.008, 0.08)

Table 2: Reasons why students like video games. Each respondent was asked to select up to three reasons.

	Proportion	95% CI
time	0.462	(0.375, 0.548)
cost	0.385	(0.3, 0.469)
point	0.319	(0.238, 0.4)
frust	0.253	(0.177, 0.328)
rules	0.187	(0.119, 0.255)
boring	0.154	(0.091, 0.217)
lonely	0.044	(0.008, 0.08)
friends	0.022	(0, 0.048)

Table 3: Reasons why students dislike video games. Each student was asked to select up to 3 reasons.
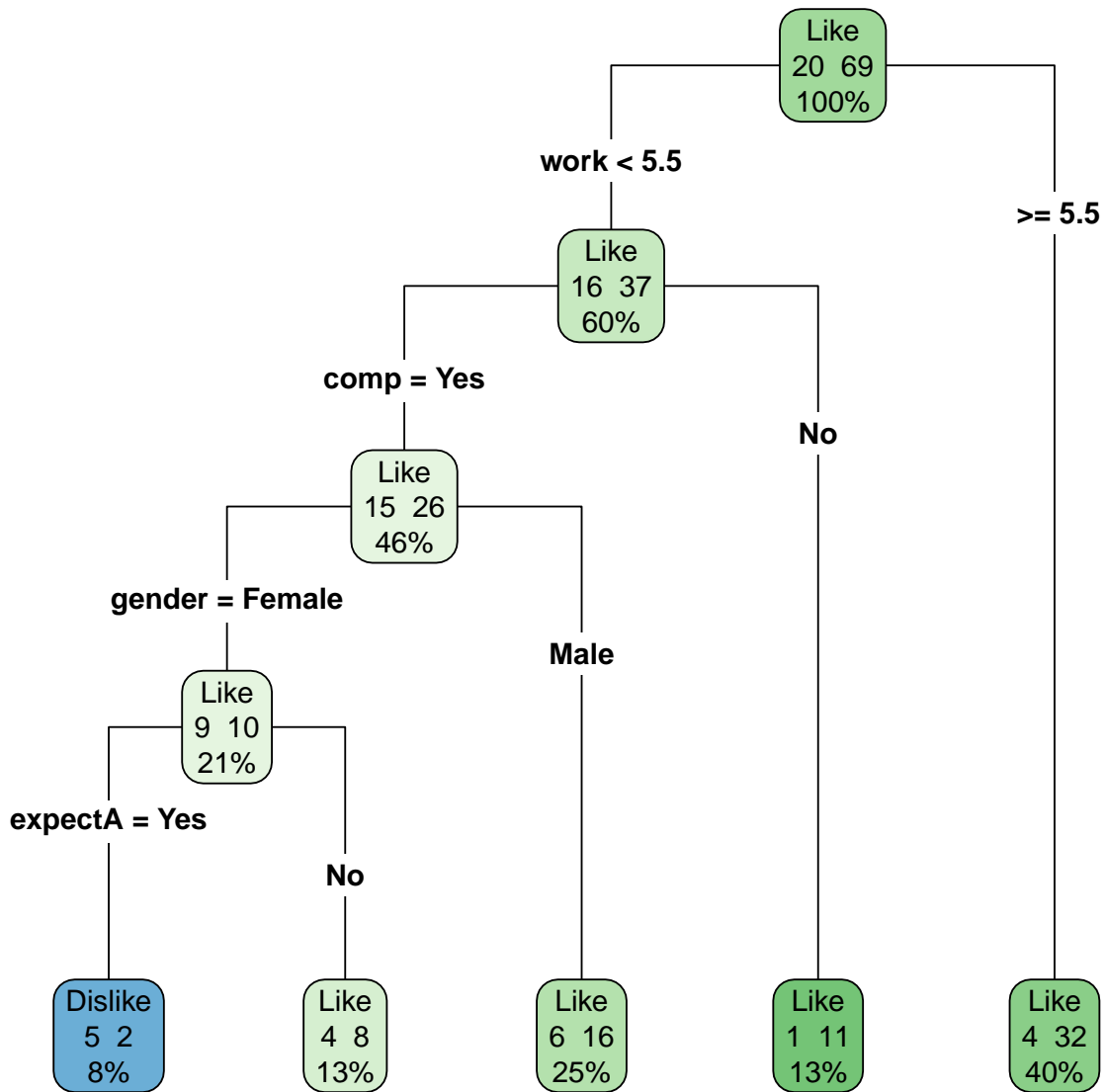
Figure 3: Decision tree for determining if a student likes or dislikes video games. The predictor variables used in the tree are gender, hours of paid work, own a computer, and expect an A in the course.