Exam Time Limits and Student Preparation: Look at the Variance

Walt A. Nelson
Missouri State University

Is there a practical and statistical difference in the amount of time a well-prepared student takes to complete an exam and the time it takes a poorly prepared student to complete an exam? Computer-based exams permit the measurement of the time students take to complete exams. This paper uses data from college-level online real estate courses to examine the relationship between test results and test duration.

MOTIVATION

The online sections I teach have grown from about 40 students per semester in 2013 to almost 200 students in 2016. Online students are given the option to take the online exam on Blackboard or to take a seated (printed) version with me as proctor in a classroom on campus. This option can save money for those students living close enough to campus to make use of the printed exam option. Students living too far away from campus must obtain (and pay for) a professional proctor, usually at a local junior college, to complete an exam. At the College of Business at Missouri State University at least one exam for each course must be proctored.

Over the past four years a few students who have taken the online course have raised the objection that not enough time is allotted to complete the exam in the real estate course. They are usually permitted about 47-49 minutes to complete 52 questions, all multiple choice. The time for final exams is mandated by administration to be two hours. About 8-11 questions require math solutions. Student are permitted to bring a financial calculator and one sheet of notes (8.5 x 11 inches) and notes may appear on both sides. Students taking the online version of the exam may do so any time Monday through Friday during the week the exam is scheduled.

A few students have invoked a "standard" that there must be one minute permitted to complete each question. Other students insist that they be permitted to take the exam without a proctor. Is there enough time to complete the exam? My approach is to cite the outcomes of the class: if time is too short, what explains the fact that there are many A's and B's in the course? Once embarked along this road, it is only natural to dig into the Blackboard data and see what can be learned.

Once a proctor loads the password into the link for the online exam, Blackboard keeps track of the starting time and finishing time for each student. Of course, Blackboard also keeps track of the student response to each question, the total value of all correct answers and attaches this data to a spreadsheet located online in the "Grade Center."

DATA

The real estate course I teach is a sophomore-level course open to all students from across campus. No previous knowledge of accounting or finance is required. The topics covered include land use

regulation, location theory, appraisal and mortgage finance, to name a few. Exams are multiple-choice. Each question is worth four points. The value of each exam is 208 points and there are three exams in the semester, so the total weight of the exams is a little over 60% of the course. Total points available during the semester are 1,000. All exams are built into Blackboard. All exams in this set have 52 questions.

There are just 68 observations in this set. Some observations come from exams with a time limit of about 49 minutes and some come from the final exam period where 120 minutes is mandated. Some observations are taken from the fall 2015 semester and others from the start of the spring 2016 semester. The dividing line between successful and less-than-successful students is somewhat arbitrary. Successful students are defined as those with scores of at least 85%. There are only 12 such observations. This collection of students is called Group A. By way of contrast, there are 56 students with scores below 85%. This collection of students is labelled Group L.

Not all students elected to take the online version of the exam. In fact many took the seated "printed" version. To allow for the fact that filling in the answer sheet takes significantly more time in the seated version, about 7 or 8 minutes extra is permitted. The students are made aware of this difference in advance, so those taking the online version do so deliberately rather than accidentally. No extra time is permitted for the final exam; all students can take up to two hours to complete it. Scores for students who completed the printed version of any exam are not included here, because it is nearly impossible to obtain an accurate measure of completion time for such individuals.

RESULTS

As shown in Table 1, below, the average test score as measured on a percentage basis was much higher for Group A than for Group L. Specifically, students in Group A scored 91% on average while students in Group L scored just 57% on average. The range of scores for Group A was 85% to 98%, while the range for Group L was 21% to 83%. Of course, breaking up the data in this way practically guarantees that the standard deviation will be lower for Group A than for Group L.

TABLE 1
TEST SCORES FOR TWO GROUPS

	Group A	Group L
N	12	56
Average Score	91%	58%
Standard Deviation	4.7%	14%
T-Test for Difference of Means	1.2418E-19*	
F-Test for Unequal Variance	.0003	
*One-tail t-test significant at the .01 level		

The more important issue is the measure of time each group takes to complete these exams. The same amount of time is available to each student electing to take the online version of the exam. That is, allocated time levels the playing field in a way which permits a reasonable view of student success. After all, if a student who does not do well complains, then how does he or she account for the fact that *others did so much better in less time?* As can be seen in Table 2, to complete an exam Group A takes just over 31 minutes on average, plus or minus 8 minutes and 53 seconds. Group L takes well over 39 minutes on average, plus or minus 17 minutes and 43 seconds.

TABLE 2
TEST DURATION (MINUTES AND SECONDS) TWO GROUPS

	Group A	Group L
N	12	56
Average Time	31:16	39:48
Standard Deviation	8:53	17:43
T-Test for Difference of Means	.0347*	
F-Test for Unequal Variance	.0272	
*One-tail t-test significant at the .05 level		

The results suggest that the students in Group A are more prepared as evidenced by their average test (percentage) score combined with the fact that they take less time to complete the exam and most important, they are more consistent in their use of time as shown by the related standard deviation of just under 9 minutes about their mean. The results shown in Table 2 table also show that the F-test for unequal variances for each variable is significant (the variances are not equal). The small data set limits the ability to generalize. However, the results indicate that it is worthwhile to accumulate more data so that the relationship between test score and test duration (time) can be examined.

APPLICATIONS AND FURTHER RESEARCH

The process of validating student achievement continues to be controversial. Leading authorities propose that traditional grading systems are flawed to the point where they should be thrown away. However, outcomes in the world of business and especially sports are treated with more respect. For example, on the golf course the placement of the hole on each putting green may appear to be arbitrary, but it is not. Hole location at major tournaments, such as *The Masters*, is based on observations recorded over several decades. That is to say, golfing professionals establish their own "peer-review" benchmarks.

Likewise, using the data provided by Blackboard allows students to compare themselves with their peers. Instead of blaming the professor, many (but not all) students who are disappointed by the results will adjust their efforts toward improvement. The Blackboard accuracy with respect to the duration of time each student takes to complete exams allows students to improve testing time management. The time allocated for each exam will be seen as less an arbitrary decision by the professor and more of a standard with which to be reckoned.

As the federal government appraises student outcomes in light of cost-benefit and loan default rates by degree and by institution, data available in Blackboard regarding testing procedures and results is truly a practical resource hiding in plain sight.