The Effects of the Students' Daily Performance on Their Final Examination in Calculus

Yan Yang[1*]

[1]Department of Mathematics, Southern University of Science and Technology, Shenzhen, Guangdong, China, 518055
*Correspondence: yangy3@sustech.edu.cn

Abstract. Calculus is a fundamental course for the students, and it is also a challenging course for the students new to college. In this study, we first introduce how we collected the students' performance data of calculus II course in southern university of science and technology, where the in-class quizzes data are obtained by using the Rain Classroom teaching tool. Then we did an empirical study about the effect of the students' daily performance on their final examination. We will also investigate how a student's regular performance affects his probability of failing the course. The empirical results could serve as a basis for developing interventions or timely warning systems to alert students to devote more time to this course to avoid failing it. Finally, future improvements in teaching calculus are discussed.

Keywords: Calculus, Rain Classroom, Empirical study, Prediction

1 INTRODUCTION

Calculus is one of the fundamental courses for the undergraduate students majoring in science, technology, engineering and other relevant science disciplines. Calculus is concerned with problems of change and motion, and it deals with quantities that approach other quantities [1]. After taking calculus, the students should be able to understand the basic concepts of differential calculus and integral calculus and their applications. Calculus is not just a tool for other scientific and technical disciplines. The training gained in this course would also help students to improve their logic thinking and abstract thinking skills, which are essential for students' future studies and research in their own fields. Ultimately, the students are expected to build up abilities to model the real world problems and then use advanced math to solve those problems.

Normally, calculus is taught for the first-year undergraduate students. Typically, it is a two-semester course in many colleges and universities of China for non-math major students. The single variable calculus is covered in the first semester and the multivariable calculus would be the primary focus in the second semester. In China, almost all students have learned the basic knowledge of differentiation in high school and are good at finding derivatives for simple functions. Therefore, most students

P. Batista et al. (eds.), *Proceedings of the 2024 International Conference on Humanities, Arts, Education and Social Development (HAESD 2024)*, Advances in Social Science, Education and Humanities Research 892,
https://doi.org/10.2991/978-2-38476-344-3_43

don't have trouble in the first half of the first semester. Some of them might encounter some difficulties in the integration parts. However, the biggest challenge comes in the second semester, when the students get to the multivariate functions. Especially, the multiple integrals, line integrals and surface integrals are the most difficult parts for the students. Besides, calculus is a fast-paced and content-rich course, which adds more challenge to the students. The final grades are based on one midterm, homework assignments, attendance and the final exam. The final exam will account for the largest proportion of the final grades. Therefore, if we could predict students' final exam score based on their daily performance in advance, we may help more students to pass this course.

Several studies investigated the reform in advanced mathematics teaching. One measure is to build formative evaluation and feedback systems [2]. According to [3], students who progressed steadily throughout the semester tended to receive a higher overall grade. Students whose progress was less consistent, and thus exhibited higher variability in weekly activities completed, tended to receive a lower grade.

In our main analysis, based on the daily data we have collected, we will quantify the effects of the attendance, students' performance on homework, in-class quizzes and midterms on the students' final exam scores. We will also investigate how a student's regular performance affects his probability of failing the course.

The goal of this study is to understand how to use the modern teaching methods and tools to help students' learning and the instructors' teaching.

2 DATA

Our analyses rely on the data we collected from the calculus II classes in Southern University of Science and Technology. We adopted two lecture sections of the same semesters as our samples. Since the behavior pattern of the students retaking calculus II are significantly different from those who never take this course before, we only included the students who are taking this course for the first time. In each section, there are around 105 students in our sample. We will use one lecture section sample to build our model, we called it section 1, and another one to test the model, we called it section 2. During our lectures, we used Rain Classrooms, which is a smart teaching tool jointly developed by Xuetang Online and Tsinghua University. It is embedded in WeChat and PowerPoint. Instructors can distribute teaching materials to students' mobile phones via WeChat; the students, in turn, can preview learning material before class and review it after class [4]. What's more, during the lectures, we asked the students to sign in Rain Classrooms and occasionally we assign one or two in-class quizzes to test how well the students understood the teaching materials. Most of the quizzes are in form of multiple-choice questions so that the students could have instant feedback, which could motivate the students to maintain their learning interest.

Each week, we gave the students two sets of homework assignments. One is from the exercises after each section of the textbook-Thomas' Calculus. Those problems are relatively straightforward and uncomplicated. The solutions to those problems can be easily found online. So, most of the students seems doing well in that part and received high scores on that part. Another set of problems are chosen outside of the textbook, which are harder and more comprehensive. Therefore, the scores on this set of problems are lower for most students.

In the middle of the semester, we will give a same midterm for all students. At the end of the semester, a final exam will be given. Both sections will have almost the same homework assignments, in-class quizzes and two exams.

2.1 Summary Statistics

The summary statistics of the variables we analyzed for section 1 are presented in Table 1.

Table 1. Summary statistics of the students' performance in section 1

	Attend%	**HW1**	**HW2**	**Quiz**	**Midterm**	**Final**
mean	89.1	94.3	85.8	3.3	78.7	79.8
Median	92.0	96.1	89.3	3.0	81.0	81.0
Std. Dev.	11.4	5.7	10.2	1.7	12.6	14.9
Skewness	-3.0	-2.7	-1.9	0.1	-1.0	-1.3
Kurtosis	12.8	11.0	7.3	2.6	4.2	6.0

Attend is the attendance rate we counted from the records in the Rain Classrooms. **HW1** is the students' average score of the textbook exercises. **HW2** is the average score of the harder assignments. The average and median of **HW1** are both higher than those of **HW2**, as we expected. The standard deviations are the reverse. It means that discrimination index of **HW2** is much higher than that of **HW1**.

2.2 Correlation matrix

Table 2 shows the correlations between the variables of the performance for the students in section 1.

Table 2. Correlation matrix

	Attend	**HW1**	**HW2**	**Quiz**	**Midterm**	**Final**	**Below_60**
Attend	1.00	0.46	0.48	0.35	0.15	0.31	-0.38
HW1		1.00	0.88	0.24	0.23	0.23	-0.19
HW2			1.00	0.37	0.39	0.39	-0.23
Quiz				1.00	0.18	0.37	-0.24
Midterm					1.00	0.58	-0.32
Final						1.00	-0.62
Below_60							1

Below_60 is a binary variable. If the final score is below 60, then it is equal to 1, otherwise it is 0. As we can see from the table, of course, midterm and final examination results are highly correlated. **HW1** is less correlated to the final score comparing to any other performance variables, but it is highly correlated to **HW2**.

3 RESULTS

From the data of the students in section 1, by using the ordinary least square method [6], we regress the variable **Final** on all other variables, we obtain the following regression line:

$$\textbf{Final} = 34.2+0.2\textbf{Attend}-0.57\textbf{HW1}+0.35\textbf{HW2}+1.65\textbf{Quiz}+0.57\textbf{Midterm} \qquad (1)$$

Equation (1) tells us that if other factors are the same, the students with 10 more points on midterm will get 5.7 points higher on final examination on average. If a student answer one more in-class quizzes problems correctly, he will get 1.65 points higher on final. It seems that the performance of students on **HW1** has a negative effect on their final examination. One reason for this happened is that **HW1** and **HW2** are highly correlated and therefore it might be the consequence of multicollinearity. Not surprisingly, **HW1** and **HW2** are not statistically significant due to this multicollinearity. But all the other independent variables are statistically significant. Since we are more interested in the prediction of the final examination score, we could ignore the multicollinearity problems.

We use equation (1) to predict the final examination score. If we plot the predicted score against the true score, the result is showed in figure 1.

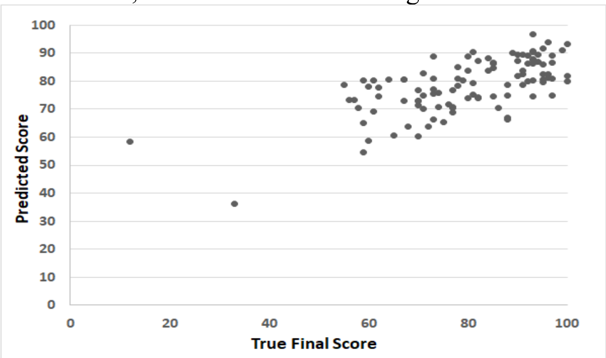

Fig. 1. The predicted scores for students in section 2

The average relative error, which is defined as the ratio of the absolute error to the true score, is around 14%. 50% of the difference between predicted score and true score is within 7 points.

We also use the data from section 2 to verify the goodness of the fitting. The prediction scores using equation (1) for the students in section 2 are illustrated in figure 2.

The average relative error is around 18%, and 50% the errors is within 8 points, which are not too worse and acceptable. This means the model we set up from the samples in section 1 also could be applied for the students in section 2.

If we only consider the students whose final are indeed above 60, which are showed in the right upper corner in the pictures, the results are even better for both sections.

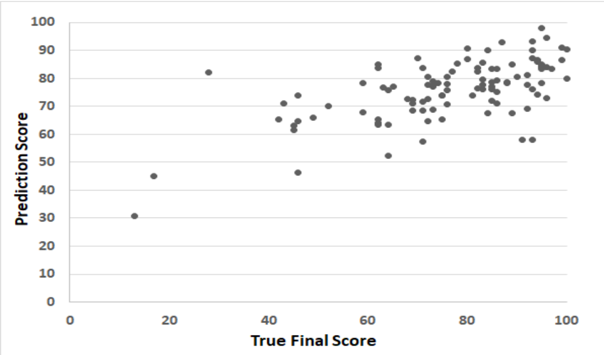

Fig. 2. The predicted scores for students in section 2

If we investigate the data more carefully, we found that whether a student's final score is below 60 depends mainly on his attendance, his performance on in-class quizzes, and midterm. For the in-class quizzes, the binary variable of whether he could answer at least two questions correctly has the strongest correlation with the variable **Below_60**, which is above 0.35. Next, we try to estimate the probability equation that the student might fail in the final exam, and the following regression equation was obtained:

$$\textbf{Below\_60} = \textbf{1.14-0.00679Attend+0.0885Quiz\_below\_2-0.006Midterm} \quad (2)$$

It turns out that attendance and the midterm are statistically significant at the 1% level, and the binary variable **Quiz_below_2** is statistically significant at the 15% level. According to equation (2), missing one class increases the probability of failing the final exam by 2.7%. Scoring 10 points lower on the midterm increases the failure probability by 6%. Students who cannot correctly answer at least two questions during the lectures will have an 8.85% higher probability of failing the final exam compared to those who perform relatively well in the quizzes. This equation quantifies the probability of failing the final exam, and it can also serve as a basis for developing interventions or timely warning systems to alert students to devote more time to this course to avoid failing it.

4 CONCLUSION

The analysis of student performance data indicates that attendance, quiz performance, and midterm scores have a positive impact on final exam results. Further analysis reveals that the probability of a student's final score being below 60 is highly correlated with attendance, quiz performance, and midterm scores.

The implications of these findings are that missing classes and poor quiz performance could be a very strong signal that the student is at a risk of failing this course. Interventions or warning systems could be developed based on these predictors to help students at risk of failing to improve their performance and avoid failure. The model's ability to predict failure probabilities suggests that it could be a useful tool for early identification of students who may need additional help.

5 FUTURE IMPROVEMENTS

From the above analysis, because the students have enough time to finish their homework assignments and sometimes the students could find the solutions online or get help from others, the homework does not play an important role when we predict the final exam score. We could put more weights on in-class quizzes since the students have to submit their solutions within a limited time frame, which more accurately reflects the student's true abilities. So, in the future, if we could have more data of that type, we could improve the accuracy of our prediction.

With the emergence of Generative AI, teaching and learning are changing rapidly. The use of artificial intelligence in higher education has risen quickly recently [5]. We could integrate AI tools and knowledge graph in our future teaching. In that way, we could let the students get customized homework assignments based on their own learning situation and provide customized prompt feedback to the students. The performance data will reflect their true learning status more accurately and thus we'll have a better prediction of their final examination scores.

ACKNOWLEDGEMENTS

This research is funded by SUSTech Teaching Reform Project ("Exploration of the Construction of a Linear Algebra Course Question Bank under the Background of Blended Learning", Grant No. XJZLGC202232).

REFERENCES

1. James Stewart. (2015) Calculus (8ed), Cengage Learning.
2. Jia, Hui-fang. (2024) A Preliminary Investigation and Practice of the Reform in Advanced Mathematics Teaching. Education and Teaching Forum, No.18
3. Ford, J., Erickson, R., Le, H., Vick, K., & Downey, J. (2024). Relating Consistent Improvement to Overall Performance in a Calculus I Course that Utilizes Standards-Based Grading. *PRIMUS*, *34*(8), 792–804. https://doi.org/10.1080/10511970.2024.2361374
4. Luo, M., Zhang, X.F. and Liu, W. (2024) Effect Evaluation of CBL Combined with Rain Classroom Teaching Method in Medical Statistics. Open Journal of Applied Sciences, 14, 1204-1213. https://doi.org/10.4236/ojapps.2024.145078
5. Crompton, H., Burke, D. Artificial intelligence in higher education: the state of the field. Int J Educ Technol High Educ 20, 22 (2023). https://doi.org/10.1186/s41239-023-00392-8
6. Jeffrey M. Wooldridge. (2019) Introductory Econometrics: A Modern Approach (7ed), Cengage Learning