

Lab G4: Student career

Emanuele Pietropaolo

Politecnico di Torino

Student id: s319501

emanuele.pietropaolo@studenti.polito.it

Abstract—The aim of this laboratory is to simulate the career of a Politecnico student. The career path can be simulated by a variety of stochastic elements that will be described in this paper.

I. PROBLEM OVERVIEW

A student's career is filled with deadlines and exams to pass. Creating a system that satisfactorily describes a student's journey through university can be challenging. The purpose of this report is to describe the stochastic simulation system that I have developed to meet this challenge. Some assumptions have been made and they will be described in detail.

II. PROPOSED APPROACH

My system simulates each student's career path individually. For computational efficiency, I choose to compute the confidence interval on the output parametrics for every 10 student simulations.

The grade distribution has been taken from given data, and each grade has a probability calculated directly from that data:

$$P(x_i) = \frac{\tilde{x}_i}{\sum_{18}^{30} d_i}$$

where \tilde{x}_i is the distribution of the grade i shown in the data, and d are all the grades for all x_i grades from 18 to 30.

I have created different input parameters that can be simulated in different combinations. This is to be able to store different career tendencies based on parameters like *average number of exams per session* or *probability of passing the exam*.

After simulating all the grades for a single student, I calculate the final grade, which follows the Politecnico di Torino rule described by this formula:

$$x = \bar{g} \cdot \frac{110}{30} + \text{thesis} + \text{presentation} + \text{bonus}$$

where \bar{g} is the average grade, the points for the *thesis evaluation* can reach up to 4, the points for the *presentation* up to 2 and the *bonus points* also up to 2.

If $x > 112.5$, the student graduates with honours.

A. Assumptions

I assumed that a year would consist of mini-sessions rather than the usual main session. In this way I can simulate the fact that a student can take the same exam more than once in the same main session. Each mini-session represents half of a normal university session. Each of them will have a maximum number of exams that can be attempted.

I also assumed that different students, or the same student in different sessions, could try a different number of exams in each session. This is to capture the variety of behaviour that students can exhibit.

I have assumed that all exams have the same probability of success. While this is not entirely accurate, it does describe the system reasonably well and allows me to use the pass probability as an input parameter and try different combinations of career paths.

B. Random Elements

- 1) The first random element is represented by the **number of exams each student attempts in the same session**. In fact, knowing that each session has a maximum number of exams that can be tried and that each student will behave in a different way, I generated the number of exams to try from a Poisson distribution. This distribution is suitable for the task because it can effectively describe random events that occur at a fixed time interval from the mean, such as the number of exams attempted in a session.
- 2) The second random element is the **Bernoulli experiment**, which can simulate the attempt of an exam. In fact, an exam can have two outcomes: passed or failed. This outcome was simulated with a binomial.
- 3) The third random element is the **grade generation**. If the exam is passed, the grades are generated from the past data (as described above).
- 4) The last random elements are the **points awarded on graduation day**, the points for the thesis evaluation, the presentation, and the bonus points. Each of these random elements is generated with a uniform (from 0 to 4 for the thesis evaluation and from 0 to 2 for the other two).

C. Data Structures

My code does not use data structures. In fact, the career of a single student can be summarised in a single list of votes and two integers representing the number of total attempts and the number of sessions passed before graduation.

There is the main method that cures the simulation of all the input parameters, the methods that cure all the random elements, and the method for saving and plotting the graphs.

III. RESULTS

All the results has been calculated after reaching an acceptable level of accuracy (97%), and with a confidence level of 95% for all output metrics.

A. Output metrics

- 1) **Grades of exams:** this metric effectively represents a student's career success.
- 2) **Number of sessions passed:** this one can describe the years that a student has passed in the university.
- 3) **Total number of attempts:** describe the total number of attempts for taking all the exams.

B. Correlations

From Fig. 1 it can be seen how similar the grade distribution is to the historical data distribution, suggesting that the system is behaving as intended.

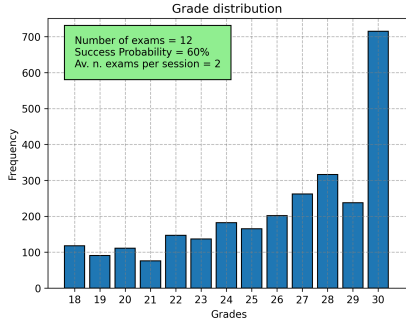


Fig. 1. Distribution of the grades.

From Fig. 2 and Fig. 3 show how the probability of success determines the average number of attempts and the time taken to complete. This makes sense as it is expected that harder exams will take longer and more attempts to pass.

Fig. 3 also shows that the length of time also depends on the average number of exams each student attempts per session. This makes sense, as the fewer exams a student attempts, the longer it takes to graduate.

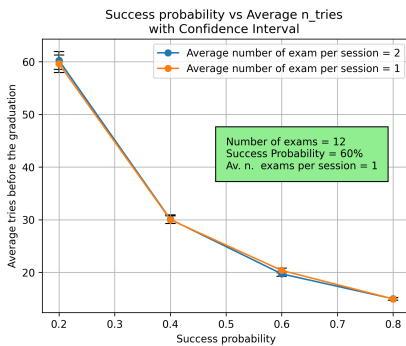


Fig. 2. Number of total attempts based on the success probability

Fig. 4 shows that the distribution of the time it takes students to graduate follows a Poisson distribution. This is

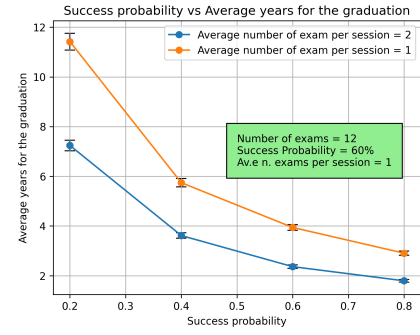


Fig. 3. Years for the graduation based on the success probability

reasonable as this distribution effectively describes the number of independent events that occur in a given period of time.

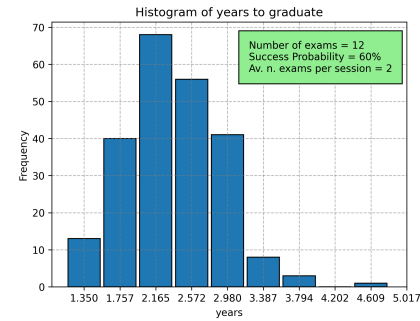


Fig. 4. Distribution of the number of years needed for the graduation

Fig. 5 describes the distribution of final grades for graduation and shows that the grades are distributed in a bell shape, although it deviates a little, and this can be explained by the presence of the random element that characterises this process.

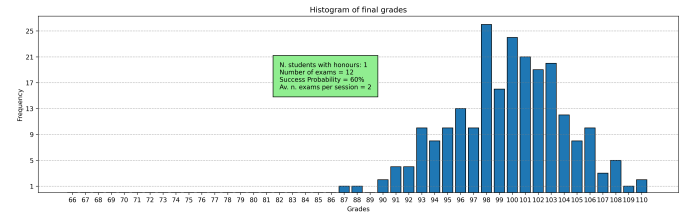


Fig. 5. Distribution of final grades.

IV. CONCLUSION

The system is behaving as expected and is leading to a satisfactory analysis of the results.