Student Career

Computer-Aided Simulations Lab G4

I. Problem overview

Our goal with the present task is to simulate a student career and evaluate the graduation time and the final grade at a MSc course.

II. Proposed approach

A. Stochastic events

There are 3 different random elements implemented in the simulation: *(1)* number of exams taken by a student during each session; *(2)* the success or not of a student at a single exam and *(3)* the grade a student achieve at a passed exam.

1) Exams taken during a session: The chosen approach for the number of exams taken per session was to draw it from a binomial distribution. In this case, consider that the student takes, on average, a given number of exams $E[X]$ and the probability of taking each exam is $p \approx 50\%$. By doing that, we define X as a binomial random variable with parameters:

$$E[X] = np \Rightarrow \begin{cases} n = \text{round}\left(\frac{E[X]}{0.5}\right) \\ p_X = \frac{E[X]}{n} \end{cases}$$

$$X \sim \text{Bin}(n, p_X)$$

2) Approval or not at an exam: The "pass/not pass an exam" event is modeled as a Bernoulli random variable Y with parameter p_Y. For this implementation, we generate a random number u uniformly distributed between 0 and 1; if $u < p_Y$, the student passes the exam.

$$Y \sim \text{Bernoulli}(p_Y)$$

3) Grades distribution: The grades of the students are generated after they pass an exam and follow the distribution presented on Table I

Grade z	Count	$P(Z=z)$	$F(z) = P(Z \leq z)$
18	87	4.71%	4.71%
19	62	3.35%	8.06%
20	74	4.00%	12.06%
21	55	2.97%	15.04%
22	99	5.35%	20.39%
23	94	5.08%	25.47%
24	117	6.33%	31.80%
25	117	6.33%	38.13%
26	136	7.36%	45.48%
27	160	8.65%	54.14%
28	215	11.63%	65.77%
29	160	8.65%	74.42%
30	473	25.58%	100%
Total	1849	100%	

TABLE I: Grades distribution for passed exams.

Again, for drawing a value from the given distribution, we generate a random value u uniformly distributed between 0 and 1. The grade of the student for the exam will be z, such that $F(z-1) < u <= F(z)$. For example, $u = 0.04$ results in the grade 18, while $u = 0.60$ results in the grade 28.

For the presented distribution, we can compute its expected value, to which we obtain:

$$E[Z] = \sum_{z=18}^{30} zP(Z=z) = 26.045 \quad (1)$$

B. Input parameters

The considered input for the simulations are:

- Probability of passing an exam (p_Y);
- Average number of exams taken by session ($E[X]$);
- Total number of courses to graduate;
- Exam sessions per year (this allows converting the time to graduate from sessions to years);
- Minimum accuracy for the chosen measures: considering an interval with 98% confidence level;

C. Output measures

There are two output metrics to analyze: the *average grade* of students after graduation and the *average graduation time*, in years.

D. Data structures

During the simulation of a student's career, the main data structures used are:

- `grades`: array recording the grade in each passed exam;
- `exams_left`: integer initialized as the total number of courses;
- `n_sessions`: counter of the total number of sessions needed for the student to graduate.

The final grade will be taken as a simple average from `grades` and the graduation time is obtained by dividing the final value of `n_sessions` by the number of sessions per year.

III. Experiments and results

A. Validation

For the first experiment, we aim at validating our simulator by analyzing how it performs for different inputs. For example, for the following input parameters:

- Probability of passing an exam: 65%;
- Average number of exams taken by session: 3.5;
- Total number of courses to graduate: 16;
- Exam sessions per year: 3;
- Minimum accuracy: 98%;

we obtained the following results:

- Number of simulations for minimum accuracy:
  - Average grade: 19
  - Average graduation time: 646

- Average grade: 26.05 ± 0.09;
- Average graduation time: 2.73 ± 0.05;

We can change some of the values to see how the simulator behaves:

- Probability of passing an exam: 55%;
- Average number of exams taken by session: 3.0;
- Total number of courses to graduate: 18;
- Exam sessions per year: 3;
- Minimum accuracy: 98%;

For these input values, the output is:

- Number of simulations for minimum accuracy:
  - Average grade: 23
  - Average graduation time: 598
- Average grade: 26.08 ± 0.08;
- Average graduation time: 4.11 ± 0.08;

From these first experiments, we can already notice that it takes way less simulations for the average grade to achieve the desired accuracy than for the average graduation time. The reason may be in the fact that the distribution of the grades is well known, as we see that the obtained intervals for the average grade fall in the distribution average we computed in Equation 1.

Other results from these experiments, we can use to validate our simulator, since by increasing the number of courses in the program and decreasing the probability of passing an exam and the average of exams taken by session, the graduation time increases as expected. The actual influence of two of these input parameter will be analyzed on the following experiments.

B. Approval probability

On this experiment, we fix the input parameters except the exam approval probability, and analyze the behaviour of the average graduation time. The considered input parameters are:

- Average number of exams taken by session: 3.5;
- Total number of courses to graduate: 16;
- Exam sessions per year: 3;
- Minimum accuracy: 98%;

We can consider the exam approval probability as a measure for the course difficulty. We expect the average approval probability for a whole MSc course to be between 40% and 70%, although we might find single courses with approval rate greater or less than those values. But as we can see from Figure 1, more difficult courses will take the students more time to graduate. As the difficulty decreases, so does the graduation time, as expected.

Furthermore, the graduation time will tend to a known value, due to the average number of exams taken by session, i.e. if the exam approval probability is $p_Y = 100\%$, the students will graduate, on average, in:

$$16 \text{ exams} \cdot \frac{1 \text{ session}}{3.5 \text{ exams}} \cdot \frac{1 \text{ year}}{3 \text{ sessions}} = 1.52 \text{ year}$$

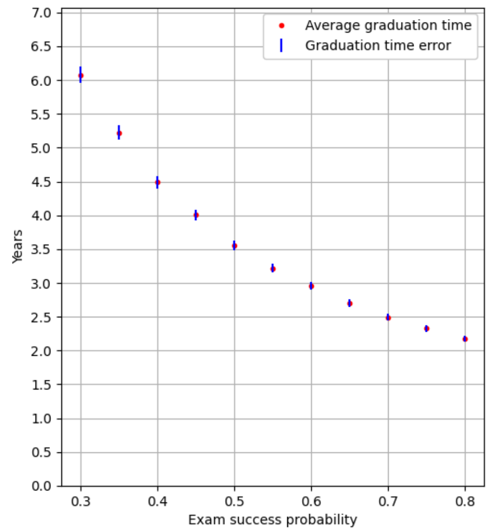

Fig. 1: Graduation time versus exam approval probability.

C. Average exams taken by session

As seen before, the average number of exams taken by the student during the exam session plays a crucial role on their expected time of graduation. In this experiment, we apply different values for this parameter while keeping the remaining fixed. This time, we use the following input parameters:

- Probability of passing an exam: 55%;
- Total number of courses to graduate: 14;
- Exam sessions per year: 3;
- Minimum accuracy: 98%;

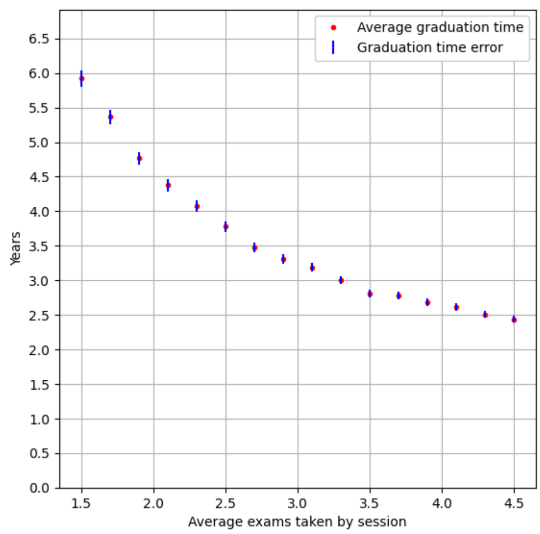

Fig. 2: Graduation time versus average exams taken by session.

In Figure 2, we observe a similar trend as the one from the previous experiment: students that take more exams at each exam session graduate sooner. This time, this might not be a completely realistic point of view, since students that take a great number of exams may have less time to prepare to each one of them and their probability of passing the exams may be lower than students that take less exams but are well prepared for all exams.