Lab6 - Basic natural selection

Martina Martini s306163
Politecnico di Torino

Assignment Consider a simulator for natural selection with the following simplified simulation model: all the individuals belong to the same species; the initial population is equal to P; the reproduction rate for each individual is lambda; the lifetime LF(k) of individual k whose parent is d(k) is distributed according to the following distribution: LF(k)= uniform(LF(d(k), LF(d(k)*(1+alpha)) with probability prob_improve and uniform(0, LF(d(k)) with probability 1-prob_improve where prob_improve is the probability of improvement for a generation and alpha is the improvement factor (≥ 0).

1 Random elements, inputs and main assumptions

1.1 Introduction

The simulation of a basic natural selection mechanism brings us to pose some questions. Some of them may be the following. Which parameters are more effective in terms of time to the extinction of the species and the average lifetime of the population? In particular, is more impactful the varying of the reproduction rate or of the probability of improvement? But overall, more generally, what are the parameters more correlated between them and how do they influence each other?

To answer all of these questions, a simulation model is implemented as explained in the following sections.

1.2 Stochastic elements

To simulate such a natural selection model, some stochastic elements must be considered. Firstly, to evaluate how the model behaves when some parameters change, for each input a list of values uniformly distributed and randomly picked is set at the beginning of the simulation. Secondly, at the trigger of the event *born*, the choice of which individual of the population must reproduce itself is random.

As assigned, the improvement of the newborn depends on the probability *probability _to_improve* and, consequently, the lifetime of the child depends on this probability. In fact, if there is an improvement, its lifetime is uniformly distributed between the lifetime of the father and the lifetime of the father multiplied by (1 + alpha); instead, if there is not an improvement, its lifetime is uniformly distributed between 0 and the lifetime of the parent. Furthermore, as requested, the inter-birth time (i.e., the time between a *born* event and the following one) is distributed as a Poisson process with rate *reproduction_rate*. Concerning this point, 2 approaches are adopted and compared.

1° approach the reproduction rate is interpreted as a global parameter, equal for each individual;

2° approach the reproduction rate is intended as an individual parameter, characteristic of the single individual and different one from another, since it is chosen randomly from a list of possibilities at the moment of birth.

To conclude, a random seed is introduced for reproducibility purposes.

1.3 Input parameters

For what concerns the input parameters, I assume to set the following fixed variables: the initial population size is equal to 10, the reproduction rate to 0.3, the probability of improvement to 0.99, alpha to 0.2 and the initial lifetime of the first generation to 50. As mentioned above, to evaluate the system's behavior when changing the value of some parameters, I decided to vary one of them and keep fixed the others. When varying, the single input is taken from a list of values: the initial population size between 10 and 50, the reproduction rate between 0.4 and 1.2, the probability of improvement between 0.1 and 0.9.

1.4 Main assumptions

Regarding the assumption I made, some hypotheses are considered in the simulation. First of all, I considered the natural selection of one species of individuals. Secondly, I assumed that the parent is just one, so each child has only one father. Moreover, I worked in a homogeneous context, in terms of lifetime of the generation 0: all the individuals of the first generation have the same initial lifetime. Finally, as explained previously, I explored both cases in which the reproduction rate is equal for each individual and if it is characteristic.

2 Output metrics, data structures and further analyses

The output metrics used to evaluate the results of the simulation consist of what shown in Figure 2. Starting from the top, we can conclude the following.

plot 1 As long as the reproduction rate grows, the extinction time (i.e., no individual is reproducing anymore) slowly decreases;

plot 2 when varying the reproduction rate, the average lifetime does not seem to be affected since it remains mostly stable;

plot 3 both the approaches (global reproduction rate and individual reproduction rate) show that the extinction time is highly influenced by the probability of improvement such that these two variables are linearly dependent;

plot 4 in both the cases, the probability of improvement seems to not impact on the average lifetime of the species;

plot 5 the initial population size influences a little the time to extinction, in the individual rate approach more than in the other one;

plot 6 the growth of the average number of children depicts an increasing trend when evaluated with the extinction time for both situations;

plot 7 intuitively, the higher is the number of individuals of the population, the higher would be the extinction time.

Further analyses are added at the end of the simulation and consist in the general analysis of the correlation between all of the variables to evaluate which one influences the most another one. As we can interestingly evaluate from the Figure 1, in the case of **global reproduction rate**,

1) the reproduction rate is highly impactful for the average lifetime;

2) the probability of improvement, instead, is very effective for the computation of the time to extinction;

3) the average number of children is correlated with the average lifetime.

Instead, using the **individual reproduction rate** approach, we can notice that

1) the probability of improvement shows a very high correlation with the final number of individuals in the population, with the extinction time and with the average number of individuals per generation;

2) the time to extinction is correlated with the total number of individuals in the population.

In conclusion, I can say that the results respect our expectations and reflect the correct behavior of a natural selection model (e.g., it's also intuitive but perfectly shown that the initial population size influences the total number of individuals at the end and the average number of individuals per generation).

For clearer evaluations, please refer to the code provided.

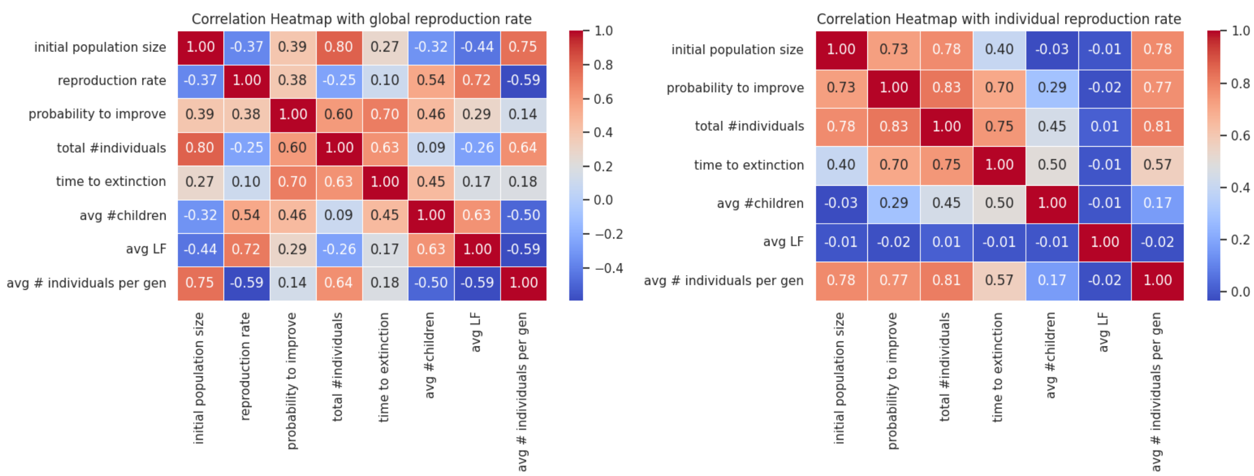

Figure 1: Correlation heatmaps

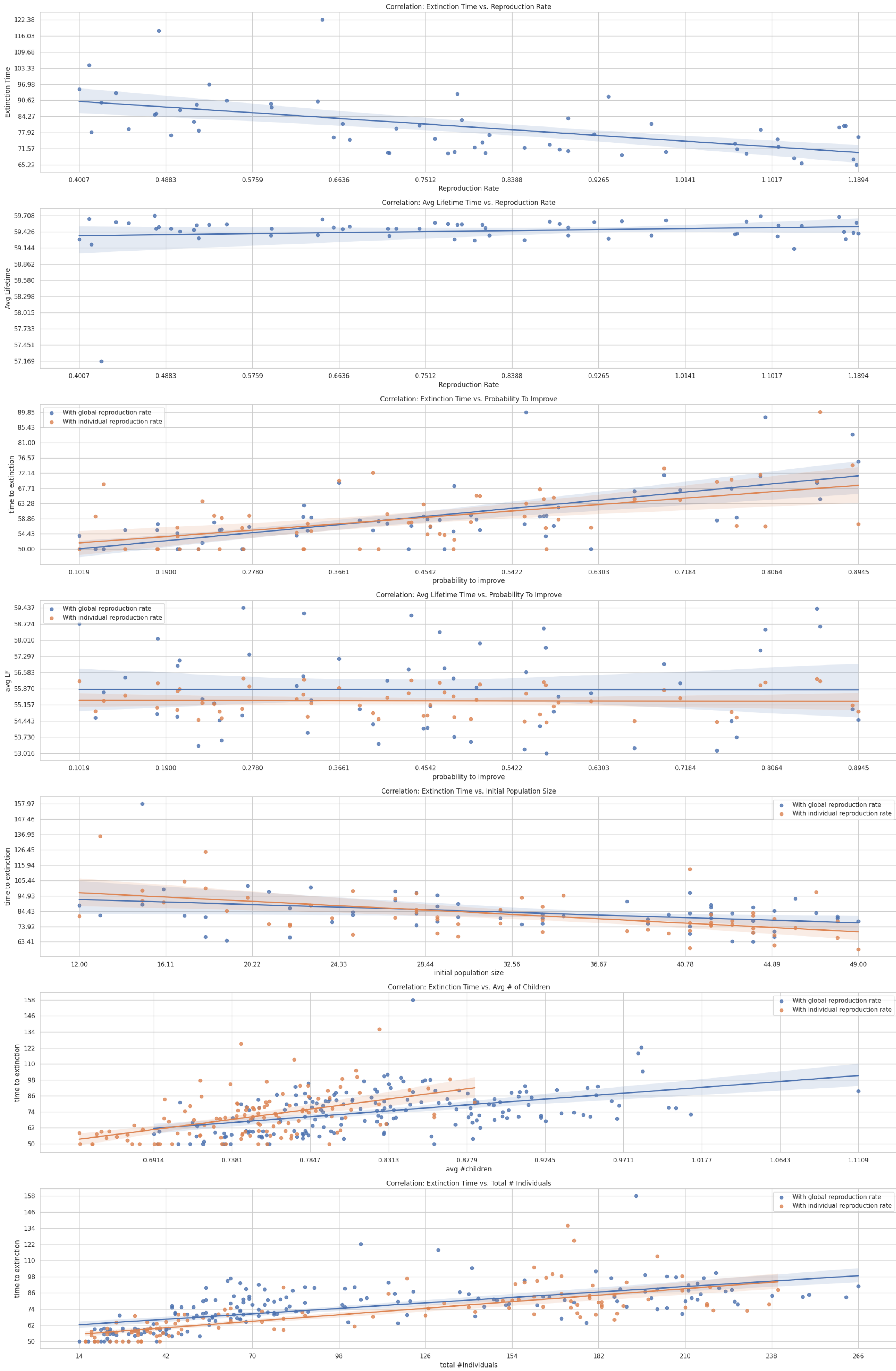

Figure 2: Main output metrics