Simulation of a branching model in continuous time

University of Technology Eindhoven

Bachelor Final Project

2WH40, 2015-2016

Author:
Stephan van Driel, ID 0858182

Supervisors:
Júlia Komjáthy
Remco van der Hofstad

August 18, 2016
# Contents

1 **Introduction**  \hspace{1cm} 1  
   1.1 Problem identification \hspace{1cm} 2

2 **Model description**  \hspace{1cm} 2  
   2.1 Assumptions \hspace{1cm} 2  
   2.2 Mathematical model \hspace{1cm} 3  
      2.2.1 General model \hspace{1cm} 3  
      2.2.2 Lifetime distributions \hspace{1cm} 3  
      2.2.3 Offspring distributions \hspace{1cm} 3  
      2.2.4 Cases of our branching process \hspace{1cm} 4  
   2.3 Research questions \hspace{1cm} 4

3 **Simulation**  \hspace{1cm} 5  
   3.1 Analysis of the distribution functions \hspace{1cm} 5  
      3.1.1 Lifetime distributions \hspace{1cm} 5  
      3.1.2 Offspring distribution \hspace{1cm} 7  
   3.2 Description of the simulation \hspace{1cm} 8

4 **Simulation results and analysis**  \hspace{1cm} 10  
   4.1 Theory for analysis \hspace{1cm} 10  
   4.2 Case 1: Poisson+1 offspring, exponential lifetime \hspace{1cm} 12  
   4.3 Case 2: Poisson+1 offspring, uniform lifetime \hspace{1cm} 17  
   4.4 Case 3: power law offspring, exponential lifetime \hspace{1cm} 23  
   4.5 Case 4: power law offspring, uniform lifetime \hspace{1cm} 28  
   4.6 Case 5: power law offspring, explosive lifetime \hspace{1cm} 32  
   4.7 Case 6: power law offspring, extreme explosive lifetime \hspace{1cm} 36

5 **Conclusion**  \hspace{1cm} 40

6 **Discussion**  \hspace{1cm} 41

**Bibliography**  \hspace{1cm} 41

**A Appendix**  \hspace{1cm} 42  
   A.1 List of variables \hspace{1cm} 42
1 Introduction

In this report we will look at the behaviour of branching processes. To introduce branching processes, we will describe applications of branching processes. After this short introduction we will explain our goal of this thesis in section 1.1 problem identification.

One of the first users of a branching process model are Galton and Watson, they developed their model in 1873 to explain the disappearance of certain family names in England [1]. This is an example of a branching process which models a certain population. This is now a days one of the applications of branching processes. In general terms, branching processes are used as models for the evaluation of a population. By populations you can think of all kind of animals, but also of biological populations. A specific example of a biological populations are amoebas. Amoebas are single-celled organisms. They are so tiny that you need a microscope to see them. They live in water, soil and as parasites in body parts of animals. Amoebas propagates by simple division. This means that one amoeba can split in half and make two identical new amoebas (this process is known as Mitosis). When an amoeba is cut in two, the part that contains the nucleus survives the other part will die. Amoebas reproduce themselves quickly and easily. This means that there can be born a lot of amoebas in little time [2].

Theodore E. Harris has made a book about the existing theory of branching processes in 1963 [3]. He studied many general branching processes for biological evaluations, but his focus was more on the total population size. The notation that Harris used for branching processes is in most cases used by other researchers. In the following years developed branching processes rapidly. One of the persons that has done research on branching processes is Wolfgang J. Bühler. He obtained related theorems concerning the generations and the relationship structure [4] [5]. Buhler refers to other studies about the distribution of generations present at a given time. Buhler refers in his paper The distribution of generations and other aspects of the family structure of branching processes [5] to Martin-Löf, Kharlamov and Samuels. Martin-Löf and Samuels considered general age dependent branching processes. Kharlamov and Bühler were mainly concerned with Markov branching processes, with exponential distribution of the lifetime. Beside that Kharlamov and Bühler also studied Galton-Watson branching processes, with constant lifetime.

Beside branching processes to model populations, they are also used by other comparable dynamics. An examples of the same dynamics are random graphs. Random graphs are introduced by Paul Erdős and Alfréd Rényi in 1959 and independently in the same year by Edgard Gilbert. They both have made models that we now call Erdős-Rényi random graph models. They are used to get a idea of networks, for example the internet or social relations. An example of an Erdős-Rényi random graph models is in figure 1. You can see it as networks, with a few branching processes. With random graphs you can determine how many people are connected with you or you can determine how many steps you are from an unfamiliar person [6].

Figure 1: Example of an Erdős-Rényi random graph [7].
1.1 Problem identification

In this report branching processes will be viewed as models for the evaluation of biological populations. But with minor adjustments it can be applied for all types of branching processes. So our branching process models a population with individuals. Each individual has his own identity. The characteristics of an individual is his generation and his lifetime. The branching processes starts with one individual. The lifetimes of the individuals is governed by a probability generating function. At the end of his lifetime he will replaced by a random number of new individuals, his children. These number of offspring is governed by an other probability generating function. The children of the first individual will live in the first generation. In general, children will get the \( k + 1 \)-st generation if their parent was a member of the \( k \)-th. We assume that all individuals have the same distribution functions for the lifetime and offspring.

During this report we will use different distribution function for the lifetime and for the offspring. We will get conservative and explosive branching processes. With a conservative branching process, we mean that we use distribution functions that gives small and predictable results. With explosive branching processes we mean the opposite, so we use distribution functions that gives small results and also really big results. If we use an distribution for the offspring with an infinity expectation, The total population will suddenly grow fast. And when we use an distribution for the lifetime with an infinity expectation. There will be very old individuals. The conservative branching are predictable. But at the explosive branching process the behaviour does not seem predictable. In this report, we want to find a particular behaviour for explosive branching processes.

We will look at the behaviour of the total population during time and the expected total population in time. We will also look at behaviour of the amount of individuals of a specific generations in time. And also the corresponding expectation of living individuals of a specific generation in time.

To make this more specific we will distinguish two properties of branching processes. We will look at the distribution of the generation of individuals that are alive at time. And we will look at the distribution of the amount of individuals that are alive at time, so the total population at a specific time.

2 Model description

2.1 Assumptions

To make a model we have to make assumptions. Preferably assumptions have no effect on the result, but this is not always the case. In this section we give the assumptions that we have to do and we will explain the influence of the assumptions on our model.

The assumptions that we need to do for our model:

- The lifetime and offspring between individuals are independent.
- The lifetime and offspring are time-independent.
- The power law offspring has a limit.

The first assumption that we have to explain is that all individuals have an independent offspring and lifetime. This assumption makes the model easier, because we can forget deceased individuals. So, the ancestor of new individuals do not influence the lifetime and offspring of the new individual. Is this assumption realistic? It depends on the situation. For example when you look at the propagation of amoebas this is a fact. But there are examples of branching processes where it has influence.

The second assumptions is similar to the first. The Lifetime and offspring is time-independent. That means that we can forget the time. So for example the number of offspring is the same in the winter as in the spring. This is not realistic in the most cases. We can also give examples of long periods which reject this assumption. For instance the baby boom after world war II. This is again a specific example, there are branching processes where the offspring and lifetime of individuals is time-independent. So, the independence of time can have influence on the model. We have to study the branching process and
will use averages as parameters.

The last assumption is that we have a limit on the power law offspring. That means that the offspring of an individual can not be big. We need this assumption because the computer has a limit of storing data. The limit on the offspring will be big and the chance that it occurs will be small. But it has influence on the reliability of our model. we want a explosive branching process, so we want big off springs. When we look at one branching process we get still a realistic model. But when we want to look at more branching processes en calculate averages, the limit will have influence.

2.2 Mathematical model

In this section we will give the mathematical model. We have already explain in words our model in section 1.1. We will now introduce mathematical terms that we will use in this report. Beside that we will explain how we will solve this model.

2.2.1 General model

We start our branching process at time \( t = 0 \). On that time we have 1 individual. This individual has generations \( G = 0 \). We calculate with the chosen random probability lifetime distribution his lifetime \( \delta_i \). We start the number of splits \( n \) at 0.

Our branching process can grow now. The individual with the lowest lifetime will go die. On the same time he will get his offspring. The number of new individuals \( D_i \) will be determined with the given random probability offspring distribution. For all new individuals we calculate there lifetimes. We increase the number of splits \( n \) to \( n + 1 \). This process will continue until the chosen number of total splits is reached or the individuals have become extinct.

During the branching process, we keep track of time and the number of living individuals \( Z(t) \). The notation that we use comes from Theodore E. Harris [3]. We will continue his notation in this report. \( Z_k(t) \) is the number of living individuals in the \( k \)-th generation. \( M(t) = E[Z(t)] \), so the expectation of the number of individuals present at time \( t \). \( M_k(t) = E[Z_k(t)] \), the expectation of the number of individuals in the \( k \)-th generation. We also define \( \rho_k(t) = M_k(t)/M(t) \). \( \rho_k(t) \) is the distribution of living individuals at each time \( t \). \( Var(G)(t) \) is the variance of generations of living individuals at time \( t \). \( T_n \). \( T_n \) is the time that the simulation has reached the \( n \)-th split. We will use the same notation, so for example \( Z(T_n) \) is the number of individuals present after \( n \) splits. To make this notation more clear, In Appendix A.1 stands a list of variables.

2.2.2 Lifetime distributions

Each individual has is own lifetime. We use the following lifetime distributions for our branching processes. We will explain why we have chose these distribution in section 2.2.4.

1. Lifetimes are generated by an exponential distribution with mean 1.
2. Lifetimes are generated by an uniform distribution on \([0,1]\).
3. Lifetimes are generated by the distribution function \( F_\delta(t) = \exp(-C/t^\beta) \) with \( C = 1 \) and \( \beta=0.99,0.90 \) and 0.85.
4. Lifetimes are generated by the distribution function \( F_\delta(t) = \exp(- \exp(C/t^\beta) + 1) \) with \( C = 1 \) and \( \beta=0.99,0.95 \) and 0.80.

2.2.3 Offspring distributions

The offspring is also a property of an individual. We will use the following offspring distributions for our branching processes. The reason for the chosen distribution functions are explained in section 2.2.4.

1. amount of new individuals are generated by \( \text{Poisson}(\lambda) + 1 \) with \( \lambda = 2, 3 \) and 4
2. amount of new individuals are generated by the distribution function: \( \mathbb{P}(D = k) = C/k^\alpha \) with \( C = 1 \) and \( \alpha = 0.99, 0.95 \) and \( 0.80 \).

### 2.2.4 Cases of our branching process

<table>
<thead>
<tr>
<th>Case</th>
<th>Lifetime distribution</th>
<th>Offspring distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 1: Cases that we simulate

Case 1 and 2 are basic branching processes (conservative branching processes). We have already explained conservative branching processes. The lifetime and offspring distribution function are predictable and thus also the branching process. For the lifetimes we use in case 1 the exponential and for case 2 the uniform distribution. The results of this distribution function will be relative small and will be around selected averages (the expectation is finite). Our model will give results that we can check. From case 2 and 3, we introduce the distribution functions for the offspring with an infinity expectation. We still use for the lifetime a distribution functions with finite expectation, for case 2 the exponential distribution and for case 3 the uniform distribution. But the offspring distribution will provide for a suddenly fast growth of the population. And there will be a generation that has a high proportion. We will chose the \( \alpha \) between 0 and 1. The chance on great results of the offspring distribution function will be higher with a \( \alpha \) closer to 0. We will be chose a \( \alpha \) closer to 1, because otherwise the offspring distribution is too large for the model. In case 3 and 4 we also use a distribution for the lifetimes with an infinite expectation. The population will grow fast and there will be individuals that are alive for a long time. We chose \( \beta \) between 0 and 1, and for \( C \) we chose 1 this has no really big influence. How closer the beta to 0, how more explosive is the distribution. The expectation of lifetime distribution in case 4 is greater then the lifetime distribution function in case 3. In the cases 1 and 2, we have a conservative branching process. In cases 3,4,5 and 6 we use offspring distribution with an infinite expectation. A conservative or an explosive branching process depends on the lifetime distribution. We expect that the branching processes in case 3 and 4 are conservative, in case 5 we have an explosive branching process and in case 6 we expect conservative ans explosive branching processes.

### 2.3 Research questions

We have now the different branching processes that we want to investigate. For each kind of branching process we want know his behaviour. We want to determine the density of \( R_k(t) \), the population at time \( t \). And we want to determine the density of \( \rho_k(t) \), generation of living individuals at time \( t \). For the conservative branching processes, the behaviour of this functions is know. We can determine \( R_k(t) \) and \( \rho_k(t) \) by hand. But we have also distribution functions that are explosive. The results of this function will suddenly be big. The behaviour of the branching processes with explosive distribution functions looks unpredictable. But we want to model also this branching processes, en maybe we can say something about his behaviour. As stated a above we can not determine with explosive branching processes \( R_k(t) \) and \( \rho_k(t) \) by hand. This is the reason that we need a simulation. We will also do a simulation for the conservative branching processes. We will compare the results of the simulation with the results that we have can determine by hand. For the explosive branching processes we will look at the density of \( R_k(t) \) and \( \rho_k(t) \). Our goal is to find comparable behaviours for explosive branching processes as for the conservative branching processes.
3 Simulation

3.1 Analysis of the distribution functions

We have made the simulation in Java. Java knows only the uniform distribution function on $[0,1]$. We want the distribution function that we have explained in section 2.2. In this section we transform the distribution functions in a workable form.

Inverse transform method

One of the methods that we use is the inverse transform method. It will work for continuous and discrete functions. We only need that the distribution function is cumulative. We define the cumulative function $F(x)$, $x \in \mathbb{R}$ as the function that we want transform in a workable form. We also define $F^{-1}(y)$, with $y \in [0,1]$ as $F^{-1}(y) = \min \left\{ x : F(x) \geq y \right\}$. Let $U$ uniform distributed on the interval $(0,1)$. We take $X = F^{-1}(U)$. Then is $X \sim F(x)$, in other words $\mathbb{P}(X \leq x) = F(x)$, with $x \in \mathbb{R}$ [8]. The proof of this method is in source [8].

3.1.1 Lifetime distributions

1

The exponential distribution function is, $F_\delta(x) = 1 - e^{-\lambda x}$.

We use the inverse transformation method to transform this function in a workable form. I have explained this method in 3.1. Determine a random number $U$ which is uniform distributed between 0 and 1. Set $U$ equal to $F_\delta(X)$, so $U = 1 - e^{-\lambda X}$. We determine now $X$ from this formula, we get:

\[
U = 1 - e^{-\lambda X}
\]
\[
\Rightarrow 1 - U = e^{-\lambda X}
\]
\[
\Rightarrow \log(1 - U) = -\lambda X
\]
\[
\Rightarrow X = \frac{-\log(1 - U)}{\lambda}
\]

We only need to calculate $U$ and we can find a exponential distributed $X$, with mean $\lambda$.

2

The uniform distribution function is, $F_\delta(x) = \frac{x-a}{b-a}$

We use the inverse transformation method to transform this function in a workable form. I have explained this method in 3.1. Determine a random number $U$ which is uniform distributed between 0 and 1. Set $U$ equal to $F_\delta(X)$, so $U = \frac{X-a}{b-a}$. We determine now $X$ from this formula, we get:

\[
U = \frac{X-a}{b-a}
\]
\[
\Rightarrow U \cdot (b - a) = X - a
\]
\[
\Rightarrow X = U \cdot (b - a) + a
\]

We only need to calculate $U$ and we can find a uniform distributed $X$ on interval $[a,b]$.

3

The distribution function that we use in case 5 is, $F_\delta(x) = \exp((-C/x^\beta))$

We use again the inverse transformation method. I have explained this method in 3.1. We determine a random number $U$ which is uniform distributed on the interval 0 and 1. Set $U$ equal to $F_\delta(x)$, so $U = \exp((-C/X^\beta))$. We determine now $X$ from this formula, we get:

\[
U = \exp((-C/X^\beta))
\]
\[
\Rightarrow \log U = -C/X^\beta
\]
\[
\Rightarrow \log(1/U) = C/X^\beta
\]
\[
\Rightarrow C/\log(1/U) = X^\beta
\]
\[
\Rightarrow X = (C/\log(1/U))^{\frac{1}{\beta}}
\]
We only need to calculate $U$ and we can find a $x$ which is distributed with $F_\delta(x)$.

Figure 2: inverse distribution function lifetimes $3$, with $C=1$, $\beta = 1$

4

The distribution function that we use in case 6 is, $F_\delta(x) = \exp(-\exp(C/x^\beta) + 1)$

We use the inverse transformation method. I have explained this method in 3.1. We determine a random number $U$ which is uniform distributed on the interval 0 and 1. Set $U$ equal to $F_\delta(x)$, so $U = \exp(-\exp(C/X^\beta) + 1)$. We determine now $X$ from this formula, we get:

$U = \exp(-\exp(C/X^\beta) + 1)$

$\Rightarrow \log U = -\exp(C/X^\beta) + 1$

$\Rightarrow 1 - \log U = \exp(C/X^\beta)$

$\Rightarrow \log (1 - \log U) = C/X^\beta$

$\Rightarrow C/(\log(1 - \log U)) = X^\beta$

$\Rightarrow X = (C/\log(\log(e/U)))^{1/\beta}$

We only need to calculate $U$ and we can find a $x$ which is distributed with $F_\delta(x)$. 
3.1.2 Offspring distribution

1

We use for the offspring distribution the poisson distribution with parameter $\lambda + 1$ in case 1 and 2. The corresponding distribution function is, $F_D(x) = 1 - e^{-\lambda x}$.

We use the exponential distribution (we have already done the transformation for the exponential distribution in 3.1.1), $X$ is the number of events in a time interval of length 1 if the inter-event times are independent and exponentially distributed with parameter $\lambda = 1$. $E_i$ is exponential distributed with mean 1. Let $i$ be the smallest index such that, $\lambda < E_1 + E_2 + ... + E_{i+1}$

We get now that $X = i$.

We only need to calculate different $E_i$ and we can find a $x$ which is distributed with Poisson($\lambda$) + 1.

2

The distribution function that we use in case 3,4,5 and 6 is, $F_D(x) = C/k^\alpha$. We use the inverse transformation method to transform this function in a workable form. I have explained this method in 3.1.

We determine a random number $U$ which is uniform distributed on the interval 0 and 1. Set $U$ equal to $F_D(x)$, so $U = C/X^\alpha$. We determine now $X$ from this formula, we get:

\[ U = C/X^\alpha \]
\[ \Rightarrow C/U = X^\alpha \]
\[ \Rightarrow X = (C/U)^{\frac{1}{\alpha}} \]

We only need to calculate $U$ and we can find a $x$ which is distributed with $F_D(x)$.
3.2 Description of the simulation

In this section we will describe the program which performs the simulation. On basis of the main code (the simulation) that we have translated in words. This code is a general code for our simulation, so we can use this code in all cases. But in the case of an exponential lifetime distribution, the code can more efficiently. In this section we will also explain this exception of our simulation.

Listing 1: Main code of program (simulation)

```
For All runs do:
    Make new list of living individuals.
    Put number of splits to 0.
    Calculate the lifetime of the first individual, with given lifetime distribution.
    Make first individual (with generation 0 and his lifetime).
    Add first individual to a sort list.
    Print time and number of living individuals to text file NumberOfIndividuals.

While splits are lower than total splits and number of living individuals is greater then 0 do:
    Take individual A with the lowest lifetime,
    Get time of death individual A.
    Get generation of individual A.
    Calculate the number of baby’s from individual A, with given offspring distribution.
    For all baby’s of individual A do:
        Calculate the lifetime of the new individual, with given lifetime distribution.
```
Make first individual (with generation and his lifetime plus local time).
Add individual to the list of amoebas.

Print local time and number of living individuals to text file
NumberOfIndividuals.
Increase number of splits with 1 split

Print of all living individuals the generation in text file Generations.

Individual B should be added to the list.
Get size of list, we call this number "inx".
While inx is greater then 0 do
    If time in list on place inx is less then or equal to the time of
death of individual B do:
        Stop while loop.
    End
    Update inx to inx − 1
End
Put individual B on place inx.

We will first explain the general program. We start the simulation by determine the variables. We
are free to choose the lifetime distribution with associated parameters, the offspring distribution with
associated parameters, the number splits and the number of runs.
The program will now start the simulation, this is given in listing 1. The number of simulations is given
by number of runs. So one run means that we do one time the simulation and otherwise we do the same
simulation multiple times. We will now explain the simulation.
The simulation will start with zero splits and we make the first individual. This individual gets generation
0 and we calculate with the given lifetime distribution his lifetime. We will now store this individual in
a list. In this list are only living individuals. This list is sort by the times that an individual goes death.
This list is important because the individual on place 0 is the amoeba that goes first death.
We will now start a while loop (line 9), this loop will stop when the number of splits is reached or
when the individuals have become extinct. We take the individual on place 0, and obtain his lifetime
and generation. This individual will get baby’s. The numbers of baby’s will calculated by the offspring
distribution. Unfortunately the individual will die now, He will be thrown out of the list of living
individuals. All Individuals in this list will move one spot to the left. We can now make new individuals
(baby’s).
We will do the same for al new individuals. First we calculate the time that the individual goes death
and his generation. The time that he goes death is the lifetime (calculated by the lifetime distribution)
plus the local time and the generation of the new individual is calculated by the generation of is parent
plus 1. This individual is alive, so we want this individual in the list of living individuals. We have to
determine the place in the list where we put this individual. The list is sort by the time they die. This
is an important step in the simulation, the code of this process is in listing 2. We start at the end of the
list, we look the time of death of each individual until this time is lower then the time of death of the
new individual. The place where we stop will be the place of the new individual in the list.
The process of one individual that goes death and new individuals will born is called one split. So the
number of splits will now increase with one. When the total number of splits is reached, the simulation
will stop and their will start a new simulation until the number of runs is reached.
Exception: lifetimes exponential distributed

Listing 3: Code to determine individual from list

1. Get size of list, we call this number "inx".
2. Determine random number r between 0 and 1.
3. Around r*inx to nearest integer.
4. We call this number "int".
5. Take individual that is on place int.

A lifetime distribution with an exponential distribution has advantages. We can still use the main code that we have explain above and is shown in Listing 1. The difference is the way we determine individuals with the lowest lifetime. In this code an individual has no given lifetime. We use the independence property of the exponential distribution. All individuals that are alive will go die independence of each other. So we have still a list with individuals that are alive and the chance that an individual has the lowest lifetime is for all individuals 1/(number of living individuals). The code in which we determine an random individual is given in Listing 3. The advantages of this code is that he is faster and the number of data that he must keep is smaller. An disadvantages is that we not keep time, so we can not make a plot of the number of living individuals after time.

Results

After we have run the program we will get 2 results. We will get a text file Generations. This text file is a list of generations of living individuals. For each run we print the generations in the same text file. Printing the generations is displayed in listing 1 in line 24.

secondly, We print during the simulation the local time and the number of living individuals on that time in text file NumberOfIndividuals. This is only possible in the main program. in the program of the exponential lifetime distribution we do not have a local time. We print the time and number of living individuals in listing 1 line 7 and line 19. For each run we print the results in the same text file.

Limit on explosive distribution functions

We use in case 5 and 6 offspring distribution function with a infinite expectation. So we will get really big numbers. Unfortunately the computer had a limited buffer capacity. That is the reason that we have a limit $L$ on the offspring distributions. When a functions gives a results higher then $L$, the code will delete the simulation and the simulation will start again. The limit will have influence on the results. We will choose in this report limit $L=1000000$.

4 Simulation results and analysis

The simulation makes the branching processes. From this processes we get results. In this section we give the results of the simulation and we will analyse the results. We will give the results for each case. We split the results in two parts. The first part is the distribution of generations of living individuals and the second part is the distribution of living individuals. Before we start with the cases we give theory that we use for the analysis.

4.1 Theory for analysis

In this section we will explain general theory that we will use for the analysis. In the most cases we can only use the simulation to get the behaviour of branching processes. But there are also branching processes, where we can calculate the expected number of individuals after n splits and there are cases were we can calculate the behaviour of the generations of living individuals.
Distribution living individuals
We introduce $E[\delta]$, the expectation of the lifetime distribution. Each split is independent of another
split. We start with 1 individual and there will die 1 individual when there is a split. So in total there
will die n individuals after n splits. After n splits the expectation of new born individuals is $n \cdot E[\delta]$. So
when we combine this we get the expected individuals after n splits.

\[ M(T_n) = 1 - n + n \cdot E[\delta] \]  

An observation of equation 1 is that when $n - 1 > n \cdot E[\delta]$ the population will extinct.

We want also know the grow rate over time. Solving the Malthusian equation will give the exponential
growth rate.

\[ E[D] \int_0^\infty e^{-rt} f(t) \, dt = 1 \]  

In equation 2 is $f(t)$ the lifetime distribution, $E[D]$ the expected number of new individuals in 1 split and
$r$ the Malthusian parameter.

Distribution generations of living individuals
M. L. Samuels explains in one of his paper the distribution of the population among generations [9]. We
will use his method to predict the behaviour of generations of branching processes after n splits.

We introduce the new distribution function $F(t)$:

\[ F(t) = E[D] \int_0^t e^{-rx} f(t) \, dx \]  

in equation 12 is $f(t)$ the lifetime distribution, $E[D]$ the expected number of the offspring distribution
and $r$ the Malthusian parameter. We define $\mu$ as the mean and $\sigma^2$ as the variance of distribution function
$F$. We assume that the Malthusian parameter exist. We can calculate the Malthusian parameter
$r$ with equation 2. We also assume $\mu < \infty$, this is always true if $E[D] > 1$ and we assume that $\sigma^2 < \infty$. This
assumption is one of the reason that we can not use this theory for explosive branching processes.

Samuels’ theorem says: If $t \to \infty$ then applies:

\[ \frac{G(t) - t/\mu}{\sqrt{t \cdot \sigma^2 / \mu^3}} \to N(0,1) \]  

With $G(t)$ the generation of a uniform chosen individual at time t. So, we have that the generations at
time t are normal distributed with mean $t/\mu$ and variance $\sqrt{t \cdot \sigma^2 / \mu^3}$.

We only have to calculate $\mu$ and $\sigma$. We will determine $\mu$ and $\sigma$ not in general. We will do the calculation
of $\mu$ and $\sigma$ in each case separately.

We do not have the generations at time $t$, we have the generation at time $T_n$. We will transform $t$ to
$T_n$, we assume that we have the Malthusian parameter $r$. We claim that $T_n \approx -\log(w) + \log(n) / r$ because we
know for $Z(t)$ that it has the form of $Z(t) \approx e^{rt} \cdot w = n$. So, $T_n = t = \frac{\log(n) - \log(w)}{r}$.

When we fill in this result in the equation of Samuels we get for $T_n \to \infty$:

\[ \frac{G(t) - t/\mu}{\sqrt{t \cdot \sigma^2 / \mu^3}} = \frac{G(T_n)}{\sqrt{T_n \cdot \sigma^2 / \mu^3}} = \frac{G(T_n) - (\log(n)/(r \cdot \mu))}{\sqrt{\log(n) - \log(w) / (r \cdot \mu^3)}} + \frac{\log(w)/(r \cdot \mu)}{\sqrt{\log(n) - \log(w) / (r \cdot \mu^3)}} \]
\begin{align*}
&= \frac{G(T_n) - (\log(n)/(r \cdot \mu))}{\sqrt{(\log(n) - \log(w))\sigma^2}/(r \cdot \mu)}
&= \frac{G(T_n) - (\log(n)/(r \cdot \mu))}{\sqrt{\log(n)\sigma^2}/(r \cdot \mu)} \cdot \sqrt{\log(n)\sigma^2}/(r \cdot \mu)
&= \frac{G(T_n) - (\log(n)/(r \cdot \mu))}{\sqrt{\log(n)\sigma^2}/(r \cdot \mu)} \rightarrow N(0, 1)
\end{align*}

So, for \( T_n \to \infty \) we get that \( G(T_n) \) is normal distributed with mean \( \log(n)/(r \cdot \mu) \) and with variance \( \sqrt{\log(n)\sigma^2}/(r \cdot \mu) \).

### 4.2 Case 1: Poisson+1 offspring, exponential lifetime

Offspring \((D_i)\): Poisson(\(\lambda\))+ 1.

Lifetime \((\delta_i)\): Exp(1).

**Distribution living individuals**

![Figure 5: Case 1: Log number of individuals after time. 1 run, 300000 splits, \(\lambda = 2\)](Image)

We know for \( Z(t) \) that it has the form of \( w \cdot e^{r \cdot t} \), with \( r \) the Malthusian parameter and \( w \) a random number. In Figure (5) we have plotted the log(individuals) against the time. This means that we get \( \log(Z(t)) = \log(w \cdot \exp(r \cdot t)) = \log(w) + r \cdot t \) So in this specific example the slope is 1.998969 \((r)\) and the intercept is 0.6797076 \((\log(w))\). The red line in figure 5 represent the fit with this slope and intercept.

We want to determine \( r \) and we want explore the behaviour of \( w \). We plot the log\((Z(t))\) against time for more runs, we get figure 6.
The slopes of the different runs in figure 6 looks the same, but the intercepts are different. This is also what we expected. The slopes are the Malthusian parameter $r$ and the intercepts are the random number $\log(w)$.

It is possible to determine the slope by hand. We have to find the Malthusian parameter $r$ in equation 2. We will solve it for all numbers of $\lambda$. $E[D] = E[\text{Poisson}(\lambda) + 1] = \lambda + 1$ and $f(t)$ is the exponential distribution function $e^{-t}$.

\[
E[D] \int_0^\infty e^{-rt} f(t) \, dt = 1
\]

\[
\Rightarrow (\lambda + 1) \int_0^\infty e^{-rt} t \, dt = 1
\]

\[
\Rightarrow (\lambda + 1) \frac{1}{r + 1} = 1
\]

\[
\Rightarrow r = \lambda
\]

To get more information about the random number $w$, we have made a histogram of $\log(w)$ in figure 7. Beside that in table 2 stands the results of different simulations. We see again that the slope is a constant number, almost equal to $r$. And we have given the mean and variance of the random number $\log(w)$. The histogram of $\log(w)$ has a recognizable shape.
Bachelor final project

4 SIMULATION RESULTS AND ANALYSIS

Figure 7: Case 1: Log Number of individuals after time. 1000 run, 10000 splits, \( \lambda = 2 \).

<table>
<thead>
<tr>
<th>( \lambda )</th>
<th>( n ) (runs)</th>
<th>( r )</th>
<th>slope</th>
<th>95%-CI slope</th>
<th>( \log(w) )</th>
<th>95%-CI ( \log(w) )</th>
<th>( \text{Var}(\log(w)) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>10 (1000000)</td>
<td>2</td>
<td>1.967031</td>
<td>(1.964935, 1.969127)</td>
<td>-1.050117</td>
<td>(-1.055043, -1.045192)</td>
<td>6.316327</td>
</tr>
<tr>
<td>2</td>
<td>100 (100000)</td>
<td>2</td>
<td>2.004418</td>
<td>(2.002340, 2.006496)</td>
<td>-1.69117</td>
<td>(-1.708598, -1.673742)</td>
<td>7.906872</td>
</tr>
<tr>
<td>2</td>
<td>1000 (10000)</td>
<td>2</td>
<td>1.999933</td>
<td>(1.997857, 2.002009)</td>
<td>-1.753086</td>
<td>(-1.806716, -1.699455)</td>
<td>7.487370</td>
</tr>
<tr>
<td>2</td>
<td>10000 (1000)</td>
<td>2</td>
<td>2.002188</td>
<td>(2.000099, 2.004278)</td>
<td>-1.666017</td>
<td>(-1.822887, -1.509146)</td>
<td>6.405974</td>
</tr>
<tr>
<td>3</td>
<td>10 (1000000)</td>
<td>3</td>
<td>2.945849</td>
<td>(2.942852, 2.948847)</td>
<td>-1.722445</td>
<td>(-1.729618, -1.715271)</td>
<td>13.395</td>
</tr>
<tr>
<td>3</td>
<td>100 (100000)</td>
<td>3</td>
<td>3.006701</td>
<td>(3.003767, 3.009635)</td>
<td>-2.432874</td>
<td>(-2.455925, -2.409824)</td>
<td>13.83124</td>
</tr>
<tr>
<td>3</td>
<td>1000 (10000)</td>
<td>3</td>
<td>3.000984</td>
<td>(2.998062, 3.003906)</td>
<td>-2.43444</td>
<td>(-2.503239, -2.365642)</td>
<td>12.32143</td>
</tr>
<tr>
<td>3</td>
<td>10000 (1000)</td>
<td>3</td>
<td>2.99507</td>
<td>(2.996563, 3.002454)</td>
<td>-2.273036</td>
<td>(-2.48469, -2.061381)</td>
<td>11.66159</td>
</tr>
<tr>
<td>4</td>
<td>10 (1000000)</td>
<td>4</td>
<td>3.925898</td>
<td>(3.922006, 3.929789)</td>
<td>-2.463388</td>
<td>(-2.472808, -2.453767)</td>
<td>23.1033</td>
</tr>
<tr>
<td>4</td>
<td>100 (100000)</td>
<td>4</td>
<td>4.01044</td>
<td>(4.006657, 4.014223)</td>
<td>-3.142495</td>
<td>(-3.170623, -3.114368)</td>
<td>20.59454</td>
</tr>
<tr>
<td>4</td>
<td>1000 (10000)</td>
<td>4</td>
<td>4.001702</td>
<td>(3.997961, 4.005443)</td>
<td>-3.217744</td>
<td>(-3.305143, -3.130345)</td>
<td>19.88458</td>
</tr>
<tr>
<td>4</td>
<td>10000 (1000)</td>
<td>4</td>
<td>4.000469</td>
<td>(3.996709, 4.004229)</td>
<td>-3.109097</td>
<td>(-3.386368, -2.831825)</td>
<td>20.0131</td>
</tr>
</tbody>
</table>

Table 2: Case 1: Determine slope, mean \( \log(w) \) and \( \text{Var}(\log(w)) \)

We can also calculate \( M(T_n) \) by hand. This is the expected number of individuals after \( n \) splits. The number of new individuals is determined by \( \text{Poisson}(\lambda) + 1 \). We know for 1 split that the expected number of new individuals is given by:

\[
\mathbb{E}[\text{Poisson}(\lambda) + 1] = \mathbb{E}[\text{Poisson}(\lambda)] + 1 = \lambda + 1.
\]  

Each split is independent of another split. We have define \( M(T_n) \) as the expected number of individuals after \( T_n \). We have explain the formula of \( M(T_n) \) in equation 1.

\[
M(T_n) = 1 - n + n \cdot \mathbb{E}[\delta]
\]

\[
= 1 - n + n \cdot (\lambda + 1)
\]
\[ = 1 + n \cdot \lambda \] (9)

In table ?? are the \( Z(T_n) \) as result of the simulation, we see that this number is almost equal to \( M(T_n) \).

\[ \begin{array}{|c|c|c|c|} \hline \lambda & \text{Splits (runs)} & M(T_n) & Z(T_n) \\ \hline 2 & 10(10000000) & 21 & 20.99815 \\ 2 & 10(1000000) & 201 & 201.036 \\ 2 & 10(100000) & 2001 & 2001.038 \\ 2 & 10(1000) & 20001 & 20001.63 \\ 3 & 10(1000000) & 31 & 30.99582 \\ 3 & 10(100000) & 301 & 300.9112 \\ 3 & 10(10000) & 3001 & 3002.576 \\ 3 & 10(1000) & 30001 & 29998.89 \\ 4 & 10(1000000) & 41 & 40.99088 \\ 4 & 10(100000) & 401 & 401.0393 \\ 4 & 10(10000) & 4001 & 4000.933 \\ 4 & 10(1000) & 40001 & 40002.34 \\ \hline \end{array} \]

Table 3: Case 1: \( M(T_n) \) and \( Z(T_n) \).

**Distribution generations of living individuals**

We also want to look to the behaviour of the generations after \( n \) splits. We have done the simulations and we have calculate the expected mean and variance of the generations after \( n \) splits. The results are in table 4.


Table 4: Case 1: Results simulation.

Histrogram 8 shows \( \rho_k(T_n) \). In this histogram the total runs of the simulation is 1000 and \( n = 10000 \). The red line stands for the normal distribution with the mean and standard divination of \( G \).
4 SIMULATION RESULTS AND ANALYSIS

Figure 8: Case 1: Density of generations of living individuals ($\rho_k(T_n)$), 1000 runs, $n = 10000$, $\lambda = 2$

In Figure stands the plot of the mean of $G$ against log(n). We see that there is a straight line for all values of $\lambda$. The mean of the generations has a logarithmic growth, so $\overline{G} = \varphi \cdot \log(n)$ with $\varphi$ a constant number. The slope of the line is $\varphi$. To determine the slopes we use the following functions: $\varphi = \frac{\overline{G}(T_{10n}) - \overline{G}(T_n)}{\log(10)}$.

We can also determine the $\varphi$ exactly by hand, we already explain this calculation in section 4.1.

We have to determine the mean $\overline{G}$ and the variance $\overline{\sigma^2}$ of the distribution function, we will use the Malthusian parameter $r = \lambda$, calculation of this parameter stands in equation 7.
\[ \mu = \mathbb{E}[D] \cdot \int_{-\infty}^{t} e^{-r \cdot x} f(t) \, dx \]

\[ \Rightarrow (\lambda + 1) \cdot \int_{0}^{t} e^{-\lambda \cdot x} e^{-x} \, dx \]

\[ \Rightarrow \int_{0}^{t} (\lambda + 1) \cdot e^{-(\lambda + 1) \cdot x} \, dx \]

(10)

So, from equation 10 we get that \( F(t) \) is exponential distributed with mean \( \lambda + 1 \). We know for the exponential distribution the mean and variance, \( \mu = 1/(\lambda + 1) \) and \( \sigma^2 = 1/(\lambda + 1)^2 \). We can now fill in \( \mu \) and \( \sigma \) in \( G = \log(n)/(r \cdot \mu) \) and \( \text{Var}(G) = \sqrt{\log(n)\sigma^2}/(r \cdot \mu^3) \). We get that \( G = \frac{\log(n) \cdot (\lambda + 1)}{\lambda} \), and \( \text{Var}(G) = \sqrt{\log(n)/(\lambda + 1)} \).

We know that \( G = \varphi \cdot \log(n) \), so \( \varphi = \frac{\lambda + 1}{\lambda} \). The results of \( \varphi \) and \( \varphi^* \) are in Table 5.

| \( \lambda \) | \( n \) | \( \varphi \) | \( \varphi^* \) | | \( \varphi - \varphi^* \) |
|---|---|---|---|---|
| 2 | 10 | 3/2 | 1.463215 | 0.036785 |
| 2 | 100 | 3/2 | 1.495495 | 0.004505 |
| 2 | 1000 | 3/2 | 1.505723 | 0.005723 |
| 3 | 10 | 4/3 | 1.301816 | 0.039435 |
| 3 | 100 | 4/3 | 1.327692 | 0.005641 |
| 3 | 1000 | 4/3 | 1.335483 | 0.002150 |
| 4 | 10 | 5/4 | 1.220639 | 0.029371 |
| 4 | 100 | 5/4 | 1.245639 | 0.004361 |
| 4 | 1000 | 5/4 | 1.251684 | 0.001684 |

Table 5: Case 1: Results of calculation \( \varphi \) and \( \varphi^* \)

### 4.3 Case 2: Poisson+1 offspring, uniform lifetime

Offspring \( (D_i) \): Poisson(\( \lambda \))+ 1

Lifetime \( (\delta_i) \): Unif[0,1]
Distribution living individuals

We start by plotting the $Z(T)$ against time of one branching process. We expect again that $Z(t)$ has the form of $w \cdot e^{-rt}$. With $r$ the Malthusian parameter and $w$ a random number. We have plot $\log(Z(t))$ against time of one branching process in figure 15. We get that $\log(Z(t)) = \log(w) + r \cdot t$. The slope stands for the Malthusian parameter $r$ and the intercept is a random number $\log(w)$. For the specific example we get that the slope is 2.823818 ($r$) and the intercept is -0.1945777 ($\log(w)$). The red line in figure 15 represent the fit with this slope and intercept.

Figure 10: Case 2: Log Number of individuals after time. 1 run, 1000000 splits, $\lambda = 2$
We have now plot $\log(Z(T))$ for more branching processes. We get again, The slopes of the different runs in figure 11 are the constant Malthusian parameter $r$ and the intercepts are different a random number $\log(w)$.

We want also find the slope with the Malthusian equation 2 by hand. The Malthusian parameter is the slope. We will solve it for general cases. $E[D] = E[\text{Poisson}(\lambda) + 1] = \lambda + 1$ and $f(t)$ is the uniform distribution function.

$$
(\lambda + 1) \int_0^1 e^{-r \cdot t} \cdot 1 \, dt = 1 \\
\Rightarrow (\lambda + 1) \left[ \frac{\exp(-r \cdot t)}{-r} \right]_0^1 \\
\Rightarrow \frac{1 - \exp(-r)}{r} = \frac{1}{(\lambda + 1)} 
$$

So, If we determine the Malthusian parameter for $\lambda = 2$ we get $r \approx 2.82143894$.

To get more information about the random number $w$ have made a histogram of $\log(w)$ (Histogram 12). We get a comparable form as in case 1 (Histogram 7). The results of more kinds of branching processes are in table 6.
Table 6: Case 2: Determine slope, mean log($w$) and $\text{Var}(\log(w))$

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$n$ (runs)</th>
<th>$r$</th>
<th>slope</th>
<th>95%-CI slope</th>
<th>log($w$)</th>
<th>95%-CI log($w$)</th>
<th>$\text{Var}(w)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>10 (1000000)</td>
<td>2.8214</td>
<td>2.564236</td>
<td>(2.562166,2.566306)</td>
<td>-0.3677573</td>
<td>(-0.3699006,-0.365614)</td>
<td>1.19578</td>
</tr>
<tr>
<td>2</td>
<td>100 (100000)</td>
<td>2.8214</td>
<td>2.804233</td>
<td>(2.801909,2.806557)</td>
<td>-0.951015</td>
<td>(-0.9605002,-0.9415298)</td>
<td>2.342067</td>
</tr>
<tr>
<td>2</td>
<td>1000 (100000)</td>
<td>2.8214</td>
<td>2.817721</td>
<td>(2.815397,2.820045)</td>
<td>-1.014365</td>
<td>(-1.042904,-0.9858257)</td>
<td>2.120211</td>
</tr>
<tr>
<td>2</td>
<td>10000 (10000)</td>
<td>2.8214</td>
<td>2.820203</td>
<td>(2.817883,2.822524)</td>
<td>-1.08984</td>
<td>(-1.175284,-1.004396)</td>
<td>1.90051</td>
</tr>
<tr>
<td>3</td>
<td>10 (1000000)</td>
<td>3.9207</td>
<td>3.594179</td>
<td>(3.591263,3.597095)</td>
<td>-0.5843927</td>
<td>(-0.5871839,-0.5816015)</td>
<td>2.028077</td>
</tr>
<tr>
<td>3</td>
<td>100 (100000)</td>
<td>3.9207</td>
<td>3.905959</td>
<td>(3.902741,3.909177)</td>
<td>-1.238802</td>
<td>(-1.249824,-1.22778)</td>
<td>3.16251</td>
</tr>
<tr>
<td>3</td>
<td>1000 (100000)</td>
<td>3.9207</td>
<td>3.918054</td>
<td>(3.914793,3.921317)</td>
<td>-1.26119</td>
<td>(-1.293786,-1.228595)</td>
<td>2.765854</td>
</tr>
<tr>
<td>3</td>
<td>10000 (10000)</td>
<td>3.9207</td>
<td>3.921146</td>
<td>(3.917867,3.924424)</td>
<td>-1.330046</td>
<td>(-1.436089,-1.224002)</td>
<td>2.92735</td>
</tr>
<tr>
<td>4</td>
<td>10 (1000000)</td>
<td>4.9651</td>
<td>4.595586</td>
<td>(4.591788,4.599383)</td>
<td>-0.84692</td>
<td>(-0.8503963,-0.8434438)</td>
<td>3.145697</td>
</tr>
<tr>
<td>4</td>
<td>100 (100000)</td>
<td>4.9651</td>
<td>4.955101</td>
<td>(4.950994,4.959209)</td>
<td>-1.53983</td>
<td>(-1.552498,-1.527268)</td>
<td>4.142888</td>
</tr>
<tr>
<td>4</td>
<td>1000 (100000)</td>
<td>4.9651</td>
<td>4.968728</td>
<td>(4.964615,4.972841)</td>
<td>-1.620716</td>
<td>(-1.658021,-1.583412)</td>
<td>3.622659</td>
</tr>
<tr>
<td>4</td>
<td>10000 (10000)</td>
<td>4.9651</td>
<td>4.961575</td>
<td>(4.957587,4.965564)</td>
<td>-1.509627</td>
<td>(-1.621488,-1.397767)</td>
<td>3.257309</td>
</tr>
</tbody>
</table>

Table 6: Case 2: Determine slope, mean log($w$) and $\text{Var}(\log(w))$

We can again determine $M(T_n)$. This is the same as in case 1. We get equation 9: $M(T_n) = 1 + n \cdot \lambda$

$Z(T_n)$ and $(M_n)$ should be the same. So, we have for each branching process calculated these variables. The results are in table 7.
# 4 SIMULATION RESULTS AND ANALYSIS

### Table 7: Case 2: $M(T_n)$ and $Z(T_n)$.

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>Splits(runs)</th>
<th>$M(T_n)$</th>
<th>$Z(T_n)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 10(10000000)</td>
<td>21</td>
<td>20.99204</td>
<td></td>
</tr>
<tr>
<td>2 10(1000000)</td>
<td>201</td>
<td>201.0479</td>
<td></td>
</tr>
<tr>
<td>2 10(100000)</td>
<td>2001</td>
<td>2001.292</td>
<td></td>
</tr>
<tr>
<td>2 10(10000)</td>
<td>20001</td>
<td>20001.87</td>
<td></td>
</tr>
<tr>
<td>3 10(10000000)</td>
<td>31</td>
<td>31.00061</td>
<td></td>
</tr>
<tr>
<td>3 10(1000000)</td>
<td>301</td>
<td>300.9201</td>
<td></td>
</tr>
<tr>
<td>3 10(100000)</td>
<td>3001</td>
<td>3000.698</td>
<td></td>
</tr>
<tr>
<td>3 10(10000)</td>
<td>30001</td>
<td>30006.45</td>
<td></td>
</tr>
<tr>
<td>4 10(10000000)</td>
<td>41</td>
<td>40.99755</td>
<td></td>
</tr>
<tr>
<td>4 10(1000000)</td>
<td>401</td>
<td>400.9384</td>
<td></td>
</tr>
<tr>
<td>4 10(100000)</td>
<td>4001</td>
<td>4001.854</td>
<td></td>
</tr>
<tr>
<td>4 10(10000)</td>
<td>40001</td>
<td>39991.44</td>
<td></td>
</tr>
</tbody>
</table>

**Distribution of generations of living individuals**

The simulation gives a list of living individuals with his generation after $n$ splits. We have calculated of this generations the mean $\bar{G}$, Confident intervals and the variance. Beside that we have made from 1000 branching processes, an histogram of the generations (Histogram 13). The red line in this figure stands for the normal distribution with corresponding mean $\bar{G}$ and variance $Var(G)$.

### Table 8: Case 2: Results simulation.

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>Splits</th>
<th>Runs</th>
<th>$\bar{G}$</th>
<th>95%-CI of $\bar{G}$</th>
<th>$Var(G)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>10</td>
<td>100000</td>
<td>3.37434</td>
<td>(3.37239,3.37687)</td>
<td>1.310514</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>100000</td>
<td>6.155457</td>
<td>(6.144275,6.166639)</td>
<td>3.254942</td>
</tr>
<tr>
<td>2</td>
<td>1000</td>
<td>10000</td>
<td>8.95424</td>
<td>(8.909539,8.99894)</td>
<td>5.201592</td>
</tr>
<tr>
<td>2</td>
<td>10000</td>
<td>1000</td>
<td>11.73377</td>
<td>(11.57007,11.89748)</td>
<td>6.976627</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>100000</td>
<td>5.190485</td>
<td>(5.188265,5.192706)</td>
<td>1.28373</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>100000</td>
<td>5.670465</td>
<td>(5.659315,5.681615)</td>
<td>3.236223</td>
</tr>
<tr>
<td>3</td>
<td>1000</td>
<td>10000</td>
<td>8.17597</td>
<td>(8.131075,8.220865)</td>
<td>5.249266</td>
</tr>
<tr>
<td>3</td>
<td>10000</td>
<td>1000</td>
<td>10.66466</td>
<td>(10.49728,10.83149)</td>
<td>7.245991</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>100000</td>
<td>3.112513</td>
<td>(3.110283,3.114744)</td>
<td>1.295326</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>100000</td>
<td>5.483865</td>
<td>(5.472579,5.495151)</td>
<td>3.315707</td>
</tr>
<tr>
<td>4</td>
<td>1000</td>
<td>10000</td>
<td>7.867165</td>
<td>(7.821518,7.912811)</td>
<td>5.423959</td>
</tr>
<tr>
<td>4</td>
<td>10000</td>
<td>1000</td>
<td>10.2523</td>
<td>(10.08249,10.42212)</td>
<td>7.50663</td>
</tr>
</tbody>
</table>

Table 7: Case 2: $M(T_n)$ and $Z(T_n)$.

Table 8: Case 2: Results simulation.
The histogram changes in time. We want to know this behaviour. This is the reason that we have plotted $\bar{G}$. There is a straight line, for each $\lambda$. We calculate the slope of this line with equation 12.

$$\varphi = \frac{\bar{G}(T_{100}) - \bar{G}(T_n)}{\log(10)}$$

We can also determine $\varphi$ by hand. We need the theorem of Samuels that we have explain in section 4.1. We have to calculate $\mu$. We will use equation 11 to calculate the Malthusian parameter $r$. We use from the theorem that for $n \to \infty$ is $\bar{G} = \frac{\log(n)}{\mu}$ and $\bar{G} = \varphi \cdot \log(n)$, so $\varphi = \frac{1}{\mu \eta}$. We start by calculating $\bar{G}$. \( f(t) \) is the uniform distribution function and $E[D] = \lambda + 1 = m$.

$$\bar{G} = \mathbb{E}[D] \cdot \int_{-\infty}^{t} e^{-r \cdot x} f(t) \, dx$$

$$= m \cdot \int_{0}^{1} e^{-r \cdot x} \cdot 1 \, dx$$

$$= m \cdot (\frac{-x}{r} \cdot e^{-r \cdot x})_{0}^{1} - \int_{0}^{1} -(1/r)e^{-r \cdot x} \, dx$$

$$= m \cdot (\frac{-(1/r) \cdot e^{-r} - [1/r^2]e^{-r \cdot x})_{0}^{1})$$

$$= m \cdot (\frac{-(1/r) \cdot e^{-r} - (1/r^2) \cdot e^{-r} + (1/r^2)})$$

We have now to use the formula $\varphi = \frac{1}{\mu \eta}$. We can determine $r$ by equation 11 and we can now also determine $\mu$ by equation 13. We get:

$$\varphi = \frac{1}{\lambda + 1} \cdot \frac{\exp(r) \cdot r}{r - \exp(r) + 1}$$

Figure 13: Case 2: Density of generations of living individuals ($\rho_k(T_n)$). 1000 run, 10000 splits, $\lambda = 2$
4.4 Case 3: power law offspring, exponential lifetime

\( D_1: F_D = 1/U^\alpha \)

\( \delta_1: \text{Exp}(1) \)

Limit on the offspring distribution = 1000000

**Distribution living individuals**

We have plotted in figure 15 for 1 run and 10000 splits the log(individuals) against time. So the logarithmic growth of the individuals in time. Because the plot does not looks linear we have made also a plot in figure 16 the Log(Log(individuals)) against time. We get now a straighter line. So we get that \( \log \log(Z(t)) = \log(w) + y \cdot t \). where w and y are unknown variables, which we will not cover here. The plots shows us that there are leaps between the line segments. But in case 1 there was a straight line. It looks that the power law offspring provides the leaps between the lines.
Figure 15: Case 3: Log Number of individuals after time. 1 run, 10000 splits, $\alpha = 0.99$

Figure 16: Case 3: Log Number of individuals after time. 1 run, 10000 splits, $\alpha = 0.99$
Distribution generations of living individuals

For all cases we have simulate different branching processes. and for all these processes there is a table with results. In table 10 stands the results of case 3, concerning the generations after $n$ splits. We want to see if the results give a comparable behaviour in comparison with case 1 and case 2.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>Splits</th>
<th>Runs</th>
<th>$G$</th>
<th>95%-CI of $G$</th>
<th>Var($G$)</th>
<th>$Z(T_n)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.99</td>
<td>10</td>
<td>1000000</td>
<td>3.315627</td>
<td>(3.31269,3.318564)</td>
<td>2.246122</td>
<td>144.3266</td>
</tr>
<tr>
<td>0.99</td>
<td>100</td>
<td>100000</td>
<td>6.390087</td>
<td>(6.375124,6.405049)</td>
<td>5.827819</td>
<td>1386.565</td>
</tr>
<tr>
<td>0.99</td>
<td>1000</td>
<td>100000</td>
<td>8.937944</td>
<td>(8.880394,8.995494)</td>
<td>8.621087</td>
<td>13720.97</td>
</tr>
<tr>
<td>0.99</td>
<td>10000</td>
<td>100000</td>
<td>11.80317</td>
<td>(11.58931,12.01703)</td>
<td>11.90614</td>
<td>135536.3</td>
</tr>
<tr>
<td>0.95</td>
<td>10</td>
<td>1000000</td>
<td>3.351731</td>
<td>(3.348851,3.354612)</td>
<td>2.160334</td>
<td>186.3718</td>
</tr>
<tr>
<td>0.95</td>
<td>100</td>
<td>100000</td>
<td>6.277258</td>
<td>(6.26265,6.291866)</td>
<td>5.554887</td>
<td>1811.864</td>
</tr>
<tr>
<td>0.95</td>
<td>1000</td>
<td>100000</td>
<td>8.840869</td>
<td>(8.783085,8.898652)</td>
<td>8.691867</td>
<td>17700.76</td>
</tr>
<tr>
<td>0.95</td>
<td>10000</td>
<td>100000</td>
<td>11.5194</td>
<td>(11.31085,11.72794)</td>
<td>11.32188</td>
<td>173544.9</td>
</tr>
<tr>
<td>0.80</td>
<td>10</td>
<td>1000000</td>
<td>3.273372</td>
<td>(3.265016,3.281728)</td>
<td>1.817781</td>
<td>559.2707</td>
</tr>
<tr>
<td>0.80</td>
<td>100</td>
<td>100000</td>
<td>5.851969</td>
<td>(5.810781,5.893158)</td>
<td>4.416297</td>
<td>5624.333</td>
</tr>
<tr>
<td>0.80</td>
<td>1000</td>
<td>100000</td>
<td>8.2563</td>
<td>(8.084686,8.427732)</td>
<td>7.650485</td>
<td>55929.97</td>
</tr>
<tr>
<td>0.80</td>
<td>10000</td>
<td>100000</td>
<td>11.0729</td>
<td>(10.44622,11.69959)</td>
<td>10.22366</td>
<td>469385</td>
</tr>
</tbody>
</table>

Table 10: Case 3: Results simulation.

Figure 17 gives an example of 1 run, 500000 splits and $\alpha = 0.95$. It is clear to see that it has not a bell shape form. The redline is the normal distribution with mean and variance of the corresponding branching process. We want to know if the bell shape form comes when we make a histogram of more runs. We have done this two times, and the histograms should be the same. The histograms are in Figures 20 and 18. They both have more a bell shape form. But they are not the same, or comparable. We use in the simulation a limit on the offspring distribution of 1000000. In figure 19 stands a histogram where the limit was 10000000. And surprisingly there is no bell shape form.
4 SIMULATION RESULTS AND ANALYSIS

Figure 17: Case 3: Density of generations of living individuals ($\rho_k(T_n)$). 1 run, $n = 500000$, $\alpha = 0.95$.

Figure 18: Case 3: $\left(1^e\right)$ Density of generations of living individuals ($\rho_k(T_n)$). 1000 runs, $n = 10000$, $\alpha = 0.95$. 
Figure 19: Case 3: \( (2e) \) Density of generations of living individuals \( (\rho_k(T_n)) \). 1000 runs, \( n = 10000 \), \( \alpha = 0.95 \).

Figure 20: Case 3: \( (2e) \) Density of generations of living individuals \( (\rho_k(T_n)) \). 1000 runs, \( n = 10000 \), \( \alpha = 0.95 \) and Limit is 10000000.
Beside the histograms we have again determined $\varphi$. In case 1 and 2 we have determined $\varphi$ by hand, that is now not possible. In figure 21 stands the plot of $G$ against log(n) and in table 11 are the corresponding slopes of the lines between the dots ($\varphi$). The formula that we use to calculate the $\varphi$ is

$$\varphi = \frac{\bar{G}(T_{10n}) - \bar{G}(T_n)}{\log(10)}.$$ 

The lines look straight, but $\varphi$ is not a constant number. An observation is that the lines do not intersect.

![Figure 21: Case 3: Plot of $G$ at log(n), for $\alpha = 0.99, 0.95$ and $0.80$.](image)

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>n</th>
<th>$\varphi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.99</td>
<td>10</td>
<td>1.335221</td>
</tr>
<tr>
<td>0.99</td>
<td>100</td>
<td>1.10652</td>
</tr>
<tr>
<td>0.99</td>
<td>1000</td>
<td>1.244352</td>
</tr>
<tr>
<td>0.95</td>
<td>10</td>
<td>1.27054</td>
</tr>
<tr>
<td>0.95</td>
<td>100</td>
<td>1.113362</td>
</tr>
<tr>
<td>0.95</td>
<td>1000</td>
<td>1.163269</td>
</tr>
<tr>
<td>0.80</td>
<td>10</td>
<td>1.119871</td>
</tr>
<tr>
<td>0.80</td>
<td>100</td>
<td>1.044188</td>
</tr>
<tr>
<td>0.80</td>
<td>1000</td>
<td>1.223236</td>
</tr>
</tbody>
</table>

Table 11: Case 3: Results of calculation $\varphi$ and $\bar{G}$

4.5 Case 4: power law offspring, uniform lifetime

$D_i$: $F_D = 1/U^\alpha$

$\delta_i$: Unif(0,1)

Limit on the offspring distribution = 1000000

Distribution living individuals

When we look to the growth of the logarithmic individuals we do not see a straight line. We have plotted a example in figure 23. This was also in case 2. In case 2 there was a loglog growth of individuals. In this case we have also plotted in figure 23 the log log growth of individuals in time. There is a straight line, but there are still leaps between line segments.
Figure 22: Case 4: Log Number of individuals after time. 1 run, 10000 splits, $\alpha = 0.99$.

Figure 23: Case 4: Log log Number of individuals after time. 1 run, 10000 splits, $\alpha = 0.99$. 
Distribution generations of living individuals

In table 12 are results of the branching processes considering the generations. Beside that we have made histograms of the generations after \( n \) splits.

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>Splits</th>
<th>Runs</th>
<th>( \bar{G} )</th>
<th>95%-CI of ( \bar{G} )</th>
<th>Var(( G ))</th>
<th>( Z(T_n) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.99</td>
<td>10</td>
<td>100000</td>
<td>3.050678</td>
<td>(3.043186,3.058171)</td>
<td>1.461509</td>
<td>120.8319</td>
</tr>
<tr>
<td>0.99</td>
<td>100</td>
<td>1000</td>
<td>5.41905</td>
<td>(5.382859,5.45524)</td>
<td>3.409528</td>
<td>1485.62</td>
</tr>
<tr>
<td>0.99</td>
<td>1000</td>
<td>1000</td>
<td>7.614893</td>
<td>(7.461797,7.767988)</td>
<td>6.101397</td>
<td>15421.59</td>
</tr>
<tr>
<td>0.99</td>
<td>10000</td>
<td>100</td>
<td>9.879289</td>
<td>(9.367776,10.3908)</td>
<td>6.811098</td>
<td>141020.5</td>
</tr>
<tr>
<td>0.95</td>
<td>10</td>
<td>100000</td>
<td>3.150498</td>
<td>(3.143201,3.157795)</td>
<td>1.385949</td>
<td>183.1462</td>
</tr>
<tr>
<td>0.95</td>
<td>100</td>
<td>1000</td>
<td>5.280298</td>
<td>(5.244486,5.31611)</td>
<td>3.338539</td>
<td>1707.836</td>
</tr>
<tr>
<td>0.95</td>
<td>1000</td>
<td>1000</td>
<td>8.076957</td>
<td>(7.935086,8.218829)</td>
<td>5.239566</td>
<td>20727.54</td>
</tr>
<tr>
<td>0.95</td>
<td>10000</td>
<td>100</td>
<td>10.10831</td>
<td>(9.518323,10.69829)</td>
<td>9.061203</td>
<td>185435</td>
</tr>
<tr>
<td>0.80</td>
<td>10</td>
<td>100000</td>
<td>3.008334</td>
<td>(3.000721,3.015947)</td>
<td>1.508819</td>
<td>592.5338</td>
</tr>
<tr>
<td>0.80</td>
<td>100</td>
<td>1000</td>
<td>5.292721</td>
<td>(5.256715,5.328726)</td>
<td>3.374771</td>
<td>5661.482</td>
</tr>
<tr>
<td>0.80</td>
<td>1000</td>
<td>1000</td>
<td>7.721244</td>
<td>(7.582904,7.859583)</td>
<td>4.981907</td>
<td>62002.44</td>
</tr>
<tr>
<td>0.80</td>
<td>10000</td>
<td>40</td>
<td>10.08562</td>
<td>(9.240241,10.931)</td>
<td>7.441578</td>
<td>495601.3</td>
</tr>
</tbody>
</table>

Table 12: Case 4: Results simulation.

Figure 24: Case 4: Density of generations of living individuals. 1000 runs,1000 splits, \( \alpha = 0.95 \).
We have made a few histograms. This are specific examples. To predict the behaviour of the generations in time, we compare the means of the generations. We have plotted the means against \( \log(n) \) in figure 26. We have also made a table with the slopes of the line segments. We call these slopes \( \varphi \). The formula that we use is 

\[
\varphi = \frac{G(T_{\text{max}}) - G(T_{\alpha})}{\log(10)}
\]

The results are in table 13.

Figure 25: Case 4: Density of generations of living individuals. 1000 runs, 1000 splits, \( \alpha = 0.95 \).

Figure 26: Case 4: Plot of \( G \) at \( \log(n) \), for \( \alpha = 0.99, 0.95 \) and 0.80.
4.6 Case 5: power law offspring, explosive lifetime

\( D_i: F_D = \frac{1}{U^\alpha} \)

\( \delta_i: \delta_i(x) = \exp\left(\frac{-C}{x^\beta}\right) \)

Limit on the offspring distribution = 1000000

Distribution living individuals

In figure 27 is a plot of the logarithmic growth of individuals in time. The plot is an example of one branching process. We see that the branching process explodes. So, in this example around 600 seconds the individuals will grow to infinity. The moment that the explosion happens is random. To get a better view, in figure 28 the plot of figure 27 is zoomed in.

![Figure 27: Case 5: Log Number of individuals after time. 1 run, 10000 splits, \( \alpha = 0.99 \)](image-url)

Table 13: Case 4: Results of calculation \( \varphi \) and \( \overline{\varphi} \)

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>( n )</th>
<th>( \overline{\varphi} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.99</td>
<td>10</td>
<td>1.028571</td>
</tr>
<tr>
<td>0.99</td>
<td>100</td>
<td>0.9536425</td>
</tr>
<tr>
<td>0.99</td>
<td>1000</td>
<td>0.9834149</td>
</tr>
<tr>
<td>0.95</td>
<td>10</td>
<td>0.9249604</td>
</tr>
<tr>
<td>0.95</td>
<td>100</td>
<td>1.214574</td>
</tr>
<tr>
<td>0.95</td>
<td>1000</td>
<td>0.8822045</td>
</tr>
<tr>
<td>0.80</td>
<td>10</td>
<td>0.9920966</td>
</tr>
<tr>
<td>0.80</td>
<td>100</td>
<td>1.054694</td>
</tr>
<tr>
<td>0.80</td>
<td>1000</td>
<td>1.026835</td>
</tr>
</tbody>
</table>
Distribution generations of living individuals

In table 14 stands the results of the branching processes concerning the generations. We have done the simulations and we have as a result all generations of the individuals after \( n \) splits. To compare this results of different branching processes we determines the mean and variance of the generations.

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>Splits</th>
<th>Runs</th>
<th>( G )</th>
<th>95%-CI of ( G )</th>
<th>Var(( G ))</th>
<th>( Z(T_n) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.99</td>
<td>10</td>
<td>10000</td>
<td>3.658506</td>
<td>(3.647666,3.669345)</td>
<td>3.05842</td>
<td>137.3636</td>
</tr>
<tr>
<td>0.99</td>
<td>100</td>
<td>10000</td>
<td>9.945516</td>
<td>(9.852032,10.039)</td>
<td>22.74994</td>
<td>1362.592</td>
</tr>
<tr>
<td>0.99</td>
<td>1000</td>
<td>1000</td>
<td>19.03774</td>
<td>(18.55207,19.52341)</td>
<td>61.40184</td>
<td>14107.14</td>
</tr>
<tr>
<td>0.99</td>
<td>10000</td>
<td>100</td>
<td>27.26047</td>
<td>(25.16961,29.35133)</td>
<td>113.8029</td>
<td>138357.9</td>
</tr>
<tr>
<td>0.95</td>
<td>10</td>
<td>10000</td>
<td>3.82475</td>
<td>(3.813029,3.836471)</td>
<td>3.576482</td>
<td>187.1387</td>
</tr>
<tr>
<td>0.95</td>
<td>100</td>
<td>10000</td>
<td>9.884916</td>
<td>(9.79269,9.977141)</td>
<td>22.14141</td>
<td>2199.184</td>
</tr>
<tr>
<td>0.95</td>
<td>1000</td>
<td>500</td>
<td>16.89306</td>
<td>(16.19386,17.59227)</td>
<td>63.63269</td>
<td>16979.28</td>
</tr>
<tr>
<td>0.95</td>
<td>10000</td>
<td>100</td>
<td>25.52609</td>
<td>(25.13072,25.92145)</td>
<td>92.36292</td>
<td>184194.2</td>
</tr>
<tr>
<td>0.80</td>
<td>10</td>
<td>10000</td>
<td>3.499461</td>
<td>(3.488802,3.51012)</td>
<td>2.957344</td>
<td>579.1123</td>
</tr>
<tr>
<td>0.80</td>
<td>100</td>
<td>10000</td>
<td>8.923128</td>
<td>(8.842488,9.003767)</td>
<td>16.92791</td>
<td>5978.563</td>
</tr>
<tr>
<td>0.80</td>
<td>1000</td>
<td>500</td>
<td>13.44841</td>
<td>(12.91128,13.98554)</td>
<td>39.05378</td>
<td>65393.65</td>
</tr>
<tr>
<td>0.80</td>
<td>10000</td>
<td>50</td>
<td>18.9327</td>
<td>(16.84977,21.01562)</td>
<td>56.47044</td>
<td>517446.2</td>
</tr>
</tbody>
</table>

Table 14: Case 5: Results simulation.

To get a idea how the generations are distributed, we have made histograms of the generations. figure 29 shows a histograms of a branching process with 1 run and 10000 splits. The redline stands for a normal distribution with the mean and variance of the branching process. Figure 30 and 31 are histograms of branching processes with more runs. They do not have a remarkable form.
Figure 29: Case 5: Density of generations of living individuals. 1 run, 10000 splits, $\alpha = 0.99$.

Figure 30: Case 5: Density of generations of living individuals. 10000 run, 100 splits, $\alpha = 0.99$. 
To compare the different branching processes we have made a plot of the mean of generations. In case 1 and 2 there was a straight line. To get a result of this plot, we have calculated the slopes ($\varphi$) of all lines. The results are in table 15. The formula that we have used is $\varphi = \frac{G(T_{10n}) - G(T_n)}{\log(10)}$.

Figure 32: Case 5: Plot of $\overline{G}$ at $\log(n)$, for $\alpha = 0.99$, $0.95$ and $0.80$. 

Figure 31: Case 5: Density of generations of living individuals. 100 run,10000 splits, $\alpha = 0.99$. 

### 4.7 Case 6: power law offspring, extreme explosive lifetime

\( D_i: F_D = 1/U^\alpha \)

\( \delta_i: F_\delta(x) = \exp(-\exp(C/x^\beta) + 1) \)

Limit on the offspring distribution = 100000

**Distribution living individuals**

The plot in figure 33 looks the same as in case 5. On a random point in time there will be an explosion. In figure 33 this at 420 seconds. There also leaps in the plot. To get a better picture, in figure 34 is the same plot with other intervals. The form of the plot is not unique for this case, for example there are also branching processes were the population growth is like figure 23.

![Figure 33: Case 6: Log Number of individuals after time. 1 run, 10000 splits, \( \alpha = 0.99 \)](image-url)
Figure 34: Case 6: Log Number of individuals after time. 1 run, 10000 splits, $\alpha = 0.99$ (without the first 100 splits)

Distribution generations of living individuals

The simulations gives the generation of living individuals after $n$ splits. Of each branching process we have determined the mean and variance of the generations and we have calculated the $Z(T_n)$. $Z(T_n)$ is the mean of all present individuals after $n$ splits. The results are in table 16.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>Splits</th>
<th>Runs</th>
<th>$G$</th>
<th>95%-CI of G</th>
<th>$\text{Var}(G)$</th>
<th>$Z(T_n)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.99</td>
<td>10</td>
<td>100000</td>
<td>4.006031</td>
<td>(3.994773,4.017289)</td>
<td>3.29911</td>
<td>143.942</td>
</tr>
<tr>
<td>0.99</td>
<td>100</td>
<td>10000</td>
<td>9.5174962</td>
<td>(9.48216,9.671709)</td>
<td>24.36555</td>
<td>1551.081</td>
</tr>
<tr>
<td>0.99</td>
<td>1000</td>
<td>1000</td>
<td>16.24897</td>
<td>(15.83155,16.64641)</td>
<td>41.11821</td>
<td>15724.67</td>
</tr>
<tr>
<td>0.99</td>
<td>10000</td>
<td>100</td>
<td>21.56288</td>
<td>(20.11174,23.01403)</td>
<td>54.81816</td>
<td>131527.5</td>
</tr>
<tr>
<td>0.95</td>
<td>10</td>
<td>100000</td>
<td>3.843293</td>
<td>(3.832141,3.854444)</td>
<td>3.237326</td>
<td>170.7294</td>
</tr>
<tr>
<td>0.95</td>
<td>100</td>
<td>5000</td>
<td>8.993739</td>
<td>(8.880243,9.107234)</td>
<td>16.76608</td>
<td>1969.455</td>
</tr>
<tr>
<td>0.95</td>
<td>1000</td>
<td>500</td>
<td>15.21448</td>
<td>(14.65604,15.77293)</td>
<td>40.5917</td>
<td>17851.02</td>
</tr>
<tr>
<td>0.95</td>
<td>10000</td>
<td>50</td>
<td>20.61524</td>
<td>(18.54211,22.68836)</td>
<td>55.94048</td>
<td>237902.5</td>
</tr>
<tr>
<td>0.80</td>
<td>10</td>
<td>50000</td>
<td>3.853663</td>
<td>(3.839132,3.868194)</td>
<td>2.748334</td>
<td>542.1569</td>
</tr>
<tr>
<td>0.80</td>
<td>100</td>
<td>5000</td>
<td>8.007683</td>
<td>(7.907552,8.107813)</td>
<td>13.0499</td>
<td>5416.8</td>
</tr>
<tr>
<td>0.80</td>
<td>1000</td>
<td>500</td>
<td>12.01485</td>
<td>(11.58427,12.44544)</td>
<td>24.13151</td>
<td>57094.94</td>
</tr>
<tr>
<td>0.80</td>
<td>10000</td>
<td>50</td>
<td>16.61459</td>
<td>(14.95663,18.27255)</td>
<td>35.77838</td>
<td>504262.2</td>
</tr>
</tbody>
</table>

Table 16: Case 6: Results simulation.

In figure 35 is a histogram of the generations of one branching process. The red line represent the normal distribution function with mean and variance of the corresponding branching process. Figure 36 shows a histogram of the generations for more runs.
Figure 35: Case 6: Density of generations of living individuals. 1 run, 10000 splits, $\alpha = 0.99$.

Figure 36: Case 6: Density of generations of living individuals. 100 run, 10000 splits, $\alpha = 0.99$.

To compare the different branching processes, we have made a plot of the mean of generations. The
We call $\varphi$ the slope of the line segments between dots. The formula that we use is $\varphi = \frac{G(T_{\text{nth}}) - G(T_n)}{\log(10)}$.

Figure 37: Case 6: Plot of $G$ at $\log(n)$, for $\alpha = 0.99, 0.95$ and 0.80.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$n$</th>
<th>$\varphi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.99</td>
<td>10</td>
<td>2.418556</td>
</tr>
<tr>
<td>0.99</td>
<td>100</td>
<td>2.898485</td>
</tr>
<tr>
<td>0.99</td>
<td>1000</td>
<td>2.307802</td>
</tr>
<tr>
<td>0.95</td>
<td>10</td>
<td>2.23681</td>
</tr>
<tr>
<td>0.95</td>
<td>100</td>
<td>2.701635</td>
</tr>
<tr>
<td>0.95</td>
<td>1000</td>
<td>2.345516</td>
</tr>
<tr>
<td>0.80</td>
<td>10</td>
<td>1.804068</td>
</tr>
<tr>
<td>0.80</td>
<td>100</td>
<td>1.740292</td>
</tr>
<tr>
<td>0.80</td>
<td>1000</td>
<td>1.99764</td>
</tr>
</tbody>
</table>

Table 17: Case 6: Results of calculation $\varphi$ and $\overline{\varphi}$
5 Conclusion

On basis of six cases we have examined different branching processes. We have simulate the branching processes and generate data. With the collected data we have determined the behaviour of the branching processes. Case 1 and 2 were the basis of our research. The distribution functions that we have used in this cases for the offspring and lifetimes were predictable. So, beside the simulation we have done analysis by hand. We have compare the simulation with the calculated numbers and they agree. This is a sign that the simulation gives good results. For case 1 and 2, we get from the theory and the simulation that the population has a exponential growth in time. We get for the number of individuals $Z(t) = w \cdot e^{r \cdot t}$, with $r$ the Malthusian parameter and $w$ a random number. We have calculate the Malthusian parameter for both cases. To get more information about the random number $w$, we have made a histogram of $\log(w)$. The histogram has a recognizable shape, but we do not have the specific distribution. Beside time we have calculated the growth of the simulation after $n$ splits. For case 1 and 2 we get a linear growth of the population in correlation with the number of splits. In addition to the growth of the population we have data of the generations of individuals after $n$ splits. With this data we have calculate the mean and variance of the generations. We have also made histograms of the generations after $n$ splits of one branching process and of more branching processes. In case 1 and 2 we get that the generations are normal distributed. With theorem we were able to calculate the mean and variance of the generations. We have also made a plot of the generations against $\log(n)$. There is a straight line, the slope that we have calculated corresponds to the theory. The slopes look to convert to a specific number. This means that the distribution of the generation will convert in a solid form.

So, for case 1 and 2 we have results which describes the behaviour of the branching process. And we can check the results with theorem. For the other cases we could only use the simulation. The theorem uses that the distribution function have a finite expectation. But in case 3 and 4 the power law offspring distribution does not have a finite expectation and in case 5 and 6 the lifetime and offspring distribution does not have a finite expectation. But we have done the same simulation and calculation as in case 1 and 2. In case 3 the population will grow in time with a log log growth. In case 4 we have comparable results, and it look that the population also growth with log log. In case 5 and 6 we see other results. The population will explode on a random time. That means that de population will grow to infinity on a random time. But when we look to the growth of the population for each split, we see still a linear growth. But we must not forget that we have used a limit on the power law offspring distribution. We shall return to this in the discussion. We have also analysis the data concerning the generations after $n$ splits. In case 1 and 2 we get that the generations are normal distributed. In all other cases with some fantasy there is also a bell shape form. But when we look to the histogram of the generation of one branching process there is not a bell shape form. And still when we combine more branching processes there is not a bell shape form. The power law offspring provides occasionally really big numbers. Despite that we generate a lot of numbers, they do not eliminate the big numbers. The variance of the generations after $n$ splits becomes larger for each cases when $n$ grows. This means that the density of each generation becomes smaller. For each case we have make a plot of the mean of $G$ in addition to $\log(n)$. In case 1 and 2 there was a straight line. In the other cases the line was not straight, but the lines make a positive grow. So, the slopes of the line segments were not constant but they are all positive. So in general when we introduce the power law offspring the distribution of the generations after $n$ splits is not normal distributed. And the population growth explodes in case 5 and 5 when we introduce the distribution functions with infinite expectation.
6 Discussion

To get the behaviour of branching processes we have simulate the branching processes with a simulation. From the generated data we have come to a conclusion. In this section we will discuss if the conclusion are reliable and we will discuss if there are any improvements. The main goal of the simulation is to get data which represent a branching process. In case 1 and 2 the results of the simulation agree with the theorem, but in the other cases we could not use the theorem. From case 1 and 2 we can conclude that the program which executes simulation is working well. The only difference with the case 1 and 2 and the other cases is the limit in the power law offspring distribution. The reason why we have made a limit is because the limit of the buffer of the computer. The limit of the offspring distribution was really big. But the reason that we have chosen for that distribution function were the big numbers. So, the question is now gives the simulation still the results that we want and was there a better way to do the simulation. The limit will have effect on the results, because the distribution does not have an infinite expectation. We have made the branching processes more predictable. But when we look for example to the histograms, there is still not a normal distribution. The limit on the offspring distribution was big enough to get a unpredictable behaviour. Raising the limit is a good follow-up study. One of the reason that we do not have done this is the way that we have made the simulation. The simulation makes a branching process and will during this process generate data. With this data we have done calculations and analysis. An advantage of generate data is that we can always do new calculations. The data that we have generate was most of the time to big, so programs like R could not read this big files. Making an simulation and calculation in one program will solve this problem, but you have to know in advance what you will do with the data. For example I have print all generations after n splits in a text file and calculate the mean in an other statistics program. It was also possible to calculate the mean immediately in the simulation. An other improvement of the simulation is to increase the numbers of runs. In this thesis the numbers of runs dependent on the buffer size. Increasing the number of runs is also possible with a program which does a simulation and calculations. The simulation will give better results with higher number of runs and it is still a question of the generations after n splits will be normal distributed when the number of runs goes to infinity.

References


A Appendix

A.1 List of variables

<table>
<thead>
<tr>
<th>Variables</th>
<th>Definition</th>
<th>Further explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z(t)$</td>
<td>Number of individuals present at time t.</td>
<td></td>
</tr>
<tr>
<td>$Z_k(t)$</td>
<td>Number of individuals present at time t in the k-th generation.</td>
<td></td>
</tr>
<tr>
<td>$M(t)$</td>
<td>expectation of the number of individuals present at time t.</td>
<td>$M(t) = \mathbb{E}[Z(t)]$</td>
</tr>
<tr>
<td>$M_k(t)$</td>
<td>expectation of the number of individuals in the k-th generation at time t.</td>
<td>$M_k(t) = \mathbb{E}[Z_k(t)]$</td>
</tr>
<tr>
<td>$\rho_k(t)$</td>
<td>Density of generation of living individuals at time t.</td>
<td>$\rho_k(t) = M_k(t)/M(t)$</td>
</tr>
<tr>
<td>$R_k(t)$</td>
<td>Density of number of living amoebas at time t.</td>
<td>$R_k(t) = Z_k(t)/Z(t)$</td>
</tr>
<tr>
<td>$G(t)$</td>
<td>Mean of generations of living individuals at time t.</td>
<td></td>
</tr>
<tr>
<td>$Var(G(t))$</td>
<td>Variance of generations of living individuals at time t.</td>
<td></td>
</tr>
<tr>
<td>$T_n$</td>
<td>Time after n splits.</td>
<td></td>
</tr>
<tr>
<td>$\delta_i$</td>
<td>lifetime of individual i.</td>
<td></td>
</tr>
<tr>
<td>$D_i$</td>
<td>offspring of individual i.</td>
<td></td>
</tr>
<tr>
<td>$\mathbb{E}[\delta]$</td>
<td>Expectation of the lifetimes.</td>
<td></td>
</tr>
<tr>
<td>$\mathbb{E}[D]$</td>
<td>Expectation of the offspring.</td>
<td></td>
</tr>
</tbody>
</table>

Table 18: List of variables