Journal of Al-Nahrain University Vol.11(3), December, 2008, pp.117-127 Science

SIMULATION TO THE TWO BASIC GALTON PROBLEMS

Akram M. Al-Abood and Shayma A. Al-Dabbagh
Department of Mathematics and Computer Applications, College of Science,
Al-Nahrain University, Baghdad, Iraq

Abstract

In this paper a summary on simulating two basic problems in branching processes which was first posed by ***Galton***, that is the problem of finding the probability of extinction of a branching process and the problem of specifying the distribution of the offspring number in the nth generation. Three distributions were used in this simulation; namely binomial, Poisson and uniform distribution. The algorithms used in this simulation, the results obtained and the conclusions based on these results are all briefly given in this paper. It was seen practically that ,accoording to the simulation assumptions,the most suitable distribution for the first problem from the above three distributions is ***uniform*** and the next is ***Poisson***, for the second problem ***Poisson*** distribution is the most suitable.

1.Introdction:

Branching processes theory is that part of mathematics which deals with the growth and decay of populations of objects which multiply and replace one another, generation by generation, according to rules in which chance plays a prominent part.

A branching process is a ***Markov chain*** arising in the theory of population growth. In a population of like individuals, e.g. marks of a biological species neutrons in a physical substance, suppose that each individual, independently of all other individuals, is capable of giving rise to a number of offspring, this number being a random variable, say X_n , taking values 0, 1, 2…with probabilities P_0, P_1, P_2, … respectively. Starting with one individual we examine the development of its descendants generation by generation.

The initial individual is regarded as belonging to the zeroth generation, its offspring as belonging to the first generation, the total number of offspring of individuals in the first generation comprise the second generation and so on.

If all the individuals of a generation fail to reproduce, the population becomes extinct.

The probability of extinction is of our interest in this paper so that the assumption will be $P_0 > 0$, for if $P_0 = 0$ then each individual must have at least one offspring and extinction is impossible, also it must be that $P_0+P_1<1$; for if $P_0+P_1=1$ then the process is trivial because when that is happened then each individual can have at most one offspring and each generation can have number at most one, [11].

There are numerous examples of ***Markov*** branching processes that arise naturally in various scientific disciplines. Some of the most prominent ones are listed below:

(a) <u>Electron Multipliers:</u>

An electron multiplier is a device that amplifies a weak current of electrons. A series of paths are set up in the path of electrons emitted by a source.

Each electron, as it strikes the first plate, generates a random number of new electrons, which in turn strike the next plate and produce more electrons, etc.

Let X_0 be the number of electrons initially emitted, X_1 the number of electrons produced on the first plate by the impact due to the X_0 initial electrons; in general let X_n be the number of electrons emitted from the nth plate due to electrons emanating from the (n-1)st plate. The sequence of random variables $X_0, X_1, X_2, \ldots, X_n, \ldots$ constitutes a branching process.

(b) <u>Neutron Chain Reaction:</u>

A nucleus may splitied when it colide with a neutron. The resulting fission yields a random number of new neutrons. Each of these secondary neutrons may hit some other nucleus producing a random number of additional neutrons, etc. In this case the initial number of neutrons $X_0=1$.

The first generation of neutrons comprises all those produced from the fission caused by the initial neutron. The size of the first generation is a random variable X_1. In general the population X_n at the n^{th} generation is produced by the chance hits of the X_{n-1} individual neutrons of the $(n-1)^{st}$ generation.

(c) <u>Survival of Family Names:</u>

The family name is inherited by sons only. Suppose that each individual has probability P_k of having k male offspring. Then from one individual there result the 1^{st}, 2^{nd}, …, n^{th}, …generations of descendants.

The distribution of such random variables may be investegated as the number of descendants in the n^{th} generation, or the probability that the family will eventually become extinct.

(d) <u>Survival of Mutant Genes:</u>

Each individual gene has a chance to give birth to k offspring, k=1, 2, …, which are genes of the same kind. However, any individual has a chance to transform into a different type or mutant gene.

This gene may become the first in a sequence of generations of a particular mutant gene. Inquiry about the chances of survival of the mutant gene within the population of the original genesmay be done.

All the above examples posses the following structure:

Let X_0 denote the size of the initial population. Each individual gives birth, independently of the others, with probability P_k to k new individuals, where

$$P_k \geq 0, k=1, 2, \ldots, \text{and } \sum_{k=1}^{\infty} P_k = 1 \ldots\ldots\ldots (*)$$

The totality of all the direct descendants of the initial population constitutes the first generation whose size denoted by X_1. Each individual of the first generation independently bears a progeny whose size is governed by the probability distribution (*).

The descendant produced constitutes the second generation of size X_2. In general the n^{th} generation is composed of descendants of the $(n-1)^{st}$ generation each of whose members independently produces k progeny with probability P_k, k=1, 2,… . The population size of the n^{th} generation is denoted by X_n. Then X_n form a sequence of integer-valued random variables that generate a ***Markov*** chain.

For illustration, a simple numerical example with the following assumptions: [18]

(I) An appropriate sample space and probability measure must be first chosen.

(II) The 0^{th} generation is started with one genuine individual.

(III) 0, 1, 2, 3, … offspring will be seen in the first generation with probabilities P_0, P_1, P_2, P_3, … respectively.

(IV) If in the first generation there are k offspring, then in the second generation there will be $Z_1+Z_2+\ldots+Z_k$ offspring, where Z_1, Z_2, …,Z_k are independent random variables, each with the common distribution P_0, P_1, P_2, … .

The above assumptions enable us to construct a tree, and a tree measure, for any number of generations.

2. The History of the Process:

The subject of branching processes is now over a century old. Until 1972 it was believed that the theory of branching processes was originated with the “problem of the extinction of families ” posed by ***Francis Galton*** in the educational times magazine [7].

The first attempt at a solution to this problem was given by Reverend ***H. W. Watson*** [21]. Because of a mistake in algebra, he incorrectly concluded that a family name would always die out with probability 1. And from his joint paper in 1875 with ***Galton*** [8], the mathematical tool of branching emerged the ***Galton-Watson*** Process.

But in 1972 ***C.C. Heyde*** and ***E. Seneta*** had mentioned in their paper [12] a historical note on the discovery of the same problem, which was made by ***Bienayme´*** in 1845 [5] who anticipated ***Galton*** and ***Watson*** by 28 years. For this reason branching process is sometimes called the ***Bienayme´-Galton-Watson*** process. More details on the ***Bienayme´-Galton-Watson*** process are given in [1].

The ***Bienayme′ - Galton - Watson*** process had, since found, been used in many other areas, such as queuing theory (***Kendall,*** 1951) [14], genetics (***Ewens,*** 1969)[6], epidemiology (***Becker,*** 1977) [4], and demography (***Keyfitz,*** 1985) [17]. Statistical methods for these branching processes were first developed by ***Harris*** (1948) [10], and have been the target of extensive research for the last two decades in the twentieth century.[11], [18].

Also in the mid of 1980's ***Dynkin***, building on earlier work of ***Fisher*** and ***Feller*** on population genetics and of the Japanese school of Watanabe, Ikeda and Nagasawa on branching ***Markov*** processes, introduced the notion of super processes (with deep connections to the theory of partial differential equations) which arose as scaled limits of branching processes that allowed random movement of particles. This has become a major area of contemporary research in probability theory.

The subject of branching processes has had obvious implications for population dynamics, but with the development of computer science it has found new applications in area such as algorithms, data structures, combinatorics, and molecular biology especially in molecular DNA sequencing. This led to a conference titled ' Classical and Modern Branching Processes' at the IMA, Minneapolis where new developments were surveyed and open problems identified. The proceedings of the conference came out in 1997 [2].

Thus the area of branching processes is valid and well. New applications continue to be found and in turn inspire new questions for the subject.

The literature is vast of this exciting area of research and one has had to make a selection of topics. What is presented in this paper does reflect our interests and preferences. [3]

3. The Problem of Finding the Probability of Extinction:

This problem was first appeared in the last decades of the nineteenth century when ***Galton*** studied the problem of survival of family names in ***British*** peerage. The first attempt to solve this problem was made by ***Watson*** [21] while the correct solution appeared only in 1930's by ***Steffenson*** [19], [20]. This problem was first formulated by ***Galton*** as follows:

Let P_0, P_1, P_2, … be the respective probabilities that a man has 0, 1, 2, … sons, let each son have the same probability for sons of his own, and so on. What is the probability that the male line is extinct after n generations, and more generally, what is the probability for any given number of descendants in the male line in any given generation?, [11].

More details on the history and development of this problem and its solution can be found in [15], [16].

Now, the aim of our work in this section is to determine the probability that a branching process dies out by a particular generation, as well as the limit of these probabilities as the generation size increases. Let q_n be the probability of dying out by the nth generation. Then we know further that:

$$q_n = \varnothing(q_{n-1}) \qquad (3.1)$$

where $\varnothing(z)$ is the generating function for the number of offspring produced by a single parent. Equation (1.1) makes it easy to compute these probabilities, where it is known from the theory of branching processes that:

$$0 = q_0 \leq q_1 \leq q_2 \leq \ldots \leq q_n \leq 1 \qquad (3.2)$$

The necessary steps for evaluating the probability of extinction in each generation of a branching process living till a specified number of offspring **n** = 0, 1, ... are given in the following algorithm:

<u>PROB-OF-EXT Algorithm</u>:

```
1. Read n {the maximum number of generations}, m {the maximum number of offspring that could be produced by a single parent in any generation}, and the assumed initial probabilities $P_0$ , $P_1$ , $P_2$ , …, $P_m$ .
2. Set i = 1, $q_i = P_0 + P_1 * q_{i-1} + P_2 * (q_{i-1})^2 + \ldots + P_m * (q_{i-1})^m$.
3. Print "the probability of extinction of the i$^{th}$ generation is $q_i$".
4. If i = n then go to step 6.
5. i = i +1.
6. End.
```

where each $P_i = P_r(X=i)$, i = 0, 1, 2, … .

A computer program named ***probability of extinction*** was made according to the above-mentioned algorithm for evaluating the values of the probability of dying out (extinction) where 12 generations were considered for the case that a parent can produce at most two offspring with probabilities $P_0 = 0.2$, $P_1 = 0.5$, and $P_2 = 0.3$. The values of probability of extinction obtained from the above program are illustrated in the following Table:

Table (1)
The probability of extinction for 12 generations.

Generation	***Probability of dying out***
1	0.2
2	0.312
3	0.385203
4	0.437116
5	0.47589
6	0.505878
7	0.529713
8	0.549035
9	0.564949
10	0.578225
11	0.589416
12	0.598931

Table (1) shows that the probability of dying out (extinction) by 12 generations is about 0.6 and a satisfactory results will be obtained as the number of generations increases where the probability of extinction approaches 1 which satisfy inequality (1.2).

The next example shows, as a consequence of our assumption: $P_0 = 0.2$, $P_1 = 0.5$, and $P_2 = 0.3$, that the probability of eventually dying out is $\frac{2}{3}$, so that even 12 generations is not enough to give an accurate estimate for this probability.

Example (1), [18]:

This example is practically a special case of a branching process in which at most two offspring can be produced. Then:

$Ø(z) = P_0 + P_1 z + P_2 z^2$,

where $P_i = 0$ for $i = 3, 4, \dots$.

In this simple case the condition $z = Ø(z)$ yields the equation:

$q = P_0 + P_1 q + P_2 q^2$,

which is satisfied by $q = 1$ and $q = P_0/P_2$. Thus, in addition to the root $q = 1$, the second root is $q = P_0/P_2$. The ***mean*** number m1 of offspring produced by a single parent is

$$
\begin{aligned}
m1 = Ø'(1) \; &= P_1 + 2P_2 \\
&= 1 - P_0 - P_2 + 2P_2 \\
&= 1 - P_0 + P_2 \,.
\end{aligned}
$$

Thus, if $P_0 > P_2$, $m1 < 1$ and the second root is > 1. If $P_0 = P_2$, a double root $q = 1$will be gained. If $P_0 < P_2$, $m1 > 1$ and the second root q is less than 1 and represents the probability that the process will die out.

If $P_0=0.2$ and $P_2=0.3$ then the second root will have the value $\frac{2}{3}$ or about 0.67 which is nearer to 1 than the value obtained in the above Table.

In this work updating to the program "***probability of extinction***" is made three times to make three other different programs named ***probability of extinction 1, 2, and 3*** respectively. In each time a different probability distribution to the number of offspring produced by a single parent is considered. Therefore, in each time there will be different $P_0, P_1, \dots, P_m$, $m=0, 1, \dots$; where m represents the maximum number of offspring produced by a single parent (as mentioned in the above algorithm) and then a determination is made to the probability of extinction of the branching process up to the 10th generation to all the three distributions, and for comparison sake the value of the mean of the all three distributions will be fixed to 2.

The three different probability distributions that had considered in this work are the well-known distributions namely:

(1) Uniform distribution [the case U (0,4) is considered].
(2) Poisson distribution [the case $P(2)$ is considered].
(3) Binomial distribution [the case b (5,0.4) is considered].

Since it is not easy somewhat to determine the value of the probability generating function Ø(s) and accordingly the solution of the equation $q = Ø(q)$ by using inequality (1.2) especially for $m \geq 3$, so the value of the probability of extinction for $m \geq 3$ won't be determined theoretically and it will be determined just for the case when m=2 for each assumed distribution but the rest will be determined practically in the mentioned programs (this is the reason that lead to use the computer programs for evaluating this value), as follows:

3.1 The probability of extinction for the b(5, 0.4) distribution [1]:

For a r.v. X~b(5,0.4) , the p.d.f. of X is given by:

$$f(x)=\begin{cases}\binom{5}{x}(0.4)^{x}(0.6)^{5-x} & ,x=0,1,2,3,4,5\\ 0 & ,\text{otherwise.}\end{cases}$$

Then $P_0 = f(0) = (0.6)^5 = 0.07776$, $P_1 = f(1) = 0.2592$, $P_2 = f(2) = 0.3456$,

$P_3 = f(3) = 0.2304, \dots$.

Then the results from the program ***probability of extinction 1***, that was run for 9 iterations in each of which we give a different value for **m**, are given in the following Table:

Table (2)
The values of the probability of extinction for the b(5,0.4) distribution.

		The maximum number of offspring produced by a single parent								
		2	***3***	***4***	***5***	***6***	***7***	***8***	***9***	***10***
The generations	***0***	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.0778	0.0778	0.0778	0.0778	0.0778	0.0778	0.0778	0.0778	0.0778
	2	0.1000	0.1001	0.1001	0.1001	0.1001	0.1001	0.1001	0.1001	0.1001
	3	0.1071	0.1074	0.1074	0.1074	0.1074	0.1073	0.1073	0.1073	0.1074
	4	0.1095	0.1099	0.1099	0.1099	0.1099	0.1097	0.1097	0.1097	0.1098
	5	0.1103	0.1107	0.1107	0.1107	0.1107	0.1104	0.1104	0.1104	0.1107
	6	0.1105	0.1110	0.1110	0.1110	0.1110	0.1107	0.1107	0.1107	0.1110
	7	0.1106	0.1111	0.1111	0.1111	0.1111	0.1108	0.1108	0.1108	0.1111
	8	0.1107	0.1111	0.1112	0.1112	0.1112	0.1108	0.1108	0.1108	0.1111
	9	0.1107	0.1112	0.1112	0.1112	0.1112	0.1108	0.1108	0.1108	0.1111
	10	0.1107	0.1112	0.1112	0.1112	0.1112	0.1108	0.1108	0.1108	0.1111

For the case when m=2, $\O(q) = P_0 + P_1 q + P_2 q^2$, where $P_i = 0$ for $i = 3, 4, \dots$; and according to example (1) where it was concluded that the solution of the equation $\O(q) = q$ is either q=1 or $q=\frac{P_0}{P_2}$, then in this case either q=1 or $q=\frac{0.07776}{0.3456}=0.225$ which is larger than all the corresponding practical values in the above table but they seem to be close. When m=3, the cubic equation $\O(q) = P_0 + P_1 q + P_2 q^2 + P_3 q^3$ is obtained, and the solution of the equation $\O(q) = q$ is readily obtained by using the mathcad software application which gives a suitable solution at q=0.111 where this value seems to be very close to the corresponding practical values in the above Table.

3.2 The probability of extinction for the Poisson P(2)distribution, [1]:

For a r.v. X~*P*(2) , the p.d.f. of X is given by

$$f(x)=\begin{cases}\dfrac{e^{-2}\,2^{x}}{x!} & ,\ x=0,1,2,\ldots\\ 0 & ,\text{otherwise.}\end{cases}$$

Then $P_0 = f(0) = e^{-2} = 0.135335283$, $P_1 = f(1) = 0.270670566$,

$P_2 = f(2) = 0.270670566$,

$P_3 = f(3) = 0.180447044$, ….

The results from the program ***probability of extinction 2***, also it was run for 9 iterations with different values for m, are given in Table (3):

Table (3)
The values of the probability of extinction for the P(2) distribution.

		The maximum number of offspring produced by a single parent								
		2	***3***	***4***	***5***	***6***	***7***	***8***	***9***	***10***
The generations	***0***	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.1353	0.1353	0.1353	0.1353	0.1353	0.1353	0.1353	0.1353	0.1353
	2	0.1769	0.1774	0.1774	0.1774	0.1774	0.1774	0.1774	0.1774	0.1774
	3	0.1917	0.1929	0.1930	0.1930	0.1930	0.1930	0.1930	0.1930	0.1930
	4	0.1972	0.1989	0.1991	0.1991	0.1991	0.1991	0.1991	0.1991	0.1991
	5	0.1992	0.2013	0.2015	0.2015	0.2015	0.2015	0.2015	0.2015	0.2015
	6	0.2000	0.2023	0.2025	0.2025	0.2025	0.2025	0.2025	0.2025	0.2025
	7	0.2003	0.2026	0.2029	0.2029	0.2029	0.2029	0.2029	0.2029	0.2029
	8	0.2004	0.2028	0.2031	0.2031	0.2031	0.2031	0.2031	0.2031	0.2031
	9	0.2005	0.2029	0.2031	0.2031	0.2031	0.2031	0.2031	0.2031	0.2031
	10	0.2005	0.2029	0.2031	0.2032	0.2032	0.2032	0.2032	0.2032	0.2032

As in section 3.1, the theoretical value of the probability of extinction is either q =1 or q = $\frac{P_0}{P_2}$ when m=2, and since in this case $P_0 = 0.135335283$ and $P_2 = 0.270670566$, then either q =1 or q = $\frac{0.135335283}{0.270670566} = 0.5$ which is larger than all the corresponding practical values in the above table.

Also by the same way as in the previous section, by the mathcad software application, the theoretical value of the probability of extinction with ***Poisson*** distribution assumption is 0.203 which also seems very close to the corresponding values in the above Table.

3.3 The probability of extinction for the U(0,4) distribution[1]

For a r.v. X~U(0,4) , the p.d.f. of X is:

$$f(x)=\begin{cases}\dfrac{1}{4} & ,\ 0<x<4\\ 0 & ,\text{otherwise.}\end{cases}$$

Then $P_0 = f(0) = \frac{1}{4} = 0.25$,

$P_1 = f(1) = \frac{1}{4} = 0.25$, $P_2 = f(2) = \frac{1}{4} = 0.25$,

$P_3 = f(3) = \frac{1}{4} = 0.25$, ….

The results from the program ***probability of extinction 3,*** where it was run by the same way of the above two programs, are shown in the table below:

Table (4)

The values of the probability of extinction for the U(0,4) distribution.

		The maximum number of offspring produced by a single parent								
		2	***3***	***4***	***5***	***6***	***7***	***8***	***9***	***10***
The generations	***0***	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.2500	0.2500	0.2500	0.2500	0.2500	0.2500	0.2500	0.2500	0.2500
	2	0.3281	0.3320	0.3330	0.3333	0.3333	0.3333	0.3333	0.3333	0.3333
	3	0.3589	0.3697	0.3733	0.3744	0.3748	0.3749	0.3750	0.3750	0.3750
	4	0.3719	0.3892	0.3960	0.3985	0.3995	0.3998	0.3999	0.4000	0.4000
	5	0.3776	0.3999	0.4099	0.4140	0.4156	0.4163	0.4165	0.4166	0.4166
	6	0.3800	0.4060	0.4187	0.4245	0.4269	0.4279	0.4283	0.4285	0.4285
	7	0.3811	0.4094	0.4246	0.4318	0.4351	0.4365	0.4371	0.4373	0.4374
	8	0.3816	0.4114	0.4285	0.4372	0.4412	0.4431	0.4439	0.4442	0.4443
	9	0.3818	0.4126	0.4311	0.4411	0.4460	0.4482	0.4492	0.4497	0.4499
	10	0.3819	0.4133	0.4329	0.4440	0.4497	0.4523	0.4536	0.4541	0.4544

As in section 3.2, the theoretical value of the probability of extinction is either q =1 or q = $\frac{P_0}{P_2}$, and since in this case P_0 = 0.25, P_2 = 0.25, then it will be only one solution q =1 which is larger than the entire corresponding practical values in the above table. And by using the mathcad software application, the obtained theoretical value of the probability of extinction with the uniform distribution assumption is 0.414 which seems to be very close to the corresponding values in the above Table.

4. The problem of specifying the Distribution of the offspring, [1]:

At the best of our knowledge, the exact distribution of the number of offspring in the n^{th} generation is not specified yet except for some special cases, and that was the second problem raised by ***Galton***.

So simulating this problem by putting three assumptions to the distribution of the number of offspring produced by a single parent was done and then testing the hypothesis in each assumption that whether or not this distribution is suitable for this branching process by ***Chi***-square goodness of fit test.

In each case X_1 is assumed to be a random variable generated from a certain distribution then the branching process structure for generating randomly the number of offspring in the 2^{nd} ,3^{rd} , …, n^{th} generation is applied, where **n** is a given fixed number. This procedure of the same assumption and the same way of generation is repeated for ***k*** prescribed independent times to get a vector of ***k*** independent random variables (r. v.'s) each of which represents the number of offspring in the n^{th} generation, that is the sequence $\{X_{n_r}\}_{r=1}^{k}$ will be obtained.

The necessary steps for generating random variates from discrete distributions are described in the following general algorithm (and the algorithm for generating from uniform distribution, which is a continuous distribution, is a known pakage in the ***Pascal programming language*** and other programming languages):

<u>DISTN-Algorithm</u>:

```
1. Set X = 0, P(X) = q(X) = $P_0$ .
2. Generate U from U(0,1)-distribution.
```

```
3. If $U \le q(X)$, deliver X as a r. v. generated from $P_x = P_r(X=x)$, go to step 7.
4. Set A(X), where $A(X) = \frac{P_{x+1}}{P_x}$.
5. $X = X + 1$, $P(X) = A(X - 1)\,P(X - 1)$, $q(X) = q(X - 1)+P(X)$.
6. Go to step 2.
7. End.
```

So for ***binomial*** distribution, the following algorithm will be used:

BINOMIAL-DISTN-Algorithm:

```
1. Read n, p.
2. Set $X = 0$, $P(X) = q(X) = P_0 = (1 - p)^n$.
3. Generate U from U(0,1)-distribution.
4. If $U \le q(X)$, deliver X as a r. v. generated from b(n,p), go to step 9.
5. Set $A(X) = \frac{p}{(1-p)(x+1)(n-x)}$.
6. $X = X + 1$, $P(X) = A(X - 1)\,P(X - 1)$, $q(X) = q(X - 1) + P(X)$.
7. If $X = n$, deliver $X = n$ as a r. v. generated from b(n,p), go to step 9.
8. Go to step 3.
9. End.
```

And for ***Poisson*** distribution, the following algorithm is applied:

POISSON-DISTN-Algorithm:

```
1. Read $\lambda$, n.
2. Set $X = 0$, $P(X) = q(X) = e^{-\lambda}$.
3. Generate U from U(0,1)-distribution.
4. If $U \le q(X)$, deliver X as a r. v. generated from $P(\lambda)$, go to step 9.
5. Set $A(X) = \frac{\lambda}{x+1}$.
6. $X = X + 1$, $P(X) = A(X - 1)\,P(X - 1)$, $q(X) = q(X - 1)+P(X)$.
7. If $X = n$, deliver $X = n$ as a r. v. generated from $P(\lambda)$, go to step 9.
8. Go to step 3.
9. End.
```

After that the ***Chi***-square goodness of fit test is applied to test the hypothesis that the above sequence represents a random sample (r.s.) of size ***k*** drawn from the same above distribution but with mean equals to the nth multiple of the mean of the assumed distribution for the 1st generation.

In this simulation, as in the previous section, three assumptions to the distribution of X_1 were made, that is, the well-known distributions ***uniform***, ***Poisson***, and ***binomial*** were considred, separately, as the distribution of X_1.

More specifically, a special case of each of the above distributions was considered in a suitable manner that fits the branching process structure of generation in the ***turbo Pascal programming language*** and for the sake of comparison between the above three assumptions to X_1 by fixing the mean of the three distributions to 2. That is why each of the U(0,4) , P (2), and b(5,0.4) distributions were considered as a distribution of X_1 in a separate program.

So three separated programs were made, each of which generates random vectors and then tests whether each of these random vectors drawn from the same original distribution (i.e., the assumed distribution for the r.v. X_1) but with mean equals to $\{E(X_1)\}^n$ by ***Chi***-square goodness of fit test (notice that $E(X_1)=m1$).

In simulation investigation for the offspring distribution the generation of the random numbers in all generations was according to the above three mentioned (proposed) distributions for the first generation.The observed data representing the values of the nth generation with repetition of 100 times is distributed into 10 cells.

A ***Chi***-square goodness-of-fit test was made with the following hypotheses separately:

1. H_0: data from ***binomial*** distribution with mean $(m1)^{\mathbf{n}}$ ***Versus*** H_1: any other alternative.
2. H_0: data from ***Poisson*** distribution with mean $(m1)^{\mathbf{n}}$ ***Versus*** H_1: any other alternative.
3. H_0: data from ***uniform*** distribution with mean $(m1)^{\mathbf{n}}$ ***Versus*** H_1: any other alternative.

According to the above three hypotheses, the data obtained were as follows:

- For the ***binomial*** hypothesis, the following acceptance percentage table was constructed with three different significant levels:

Table (5)
The acceptance percentages for the b(5,0.4) distribution.

The Iterations	Significant levels (α)		
	0.025	0.05	0.1
1	8 %	9 %	30 %
2	8 %	14 %	14 %
3	8 %	10 %	16 %
4	12 %	8 %	16 %
5	7 %	10 %	20 %
6	14 %	12 %	16 %
7	9 %	4 %	24 %
8	11 %	3 %	12 %
9	12 %	6 %	14 %
10	13%	8%	10%

- For the ***Poisson*** hypothesis, the following acceptance percentage table was obtained with three different significant levels:

Table (6)
The acceptance percentages for the P(2) distribution.

The Iterations	Significant levels (α)		
	0.025	0.05	0.1
1	88 %	86 %	88 %
2	83 %	79 %	84 %
3	90 %	71 %	93 %
4	88 %	80 %	85 %
5	83 %	80 %	85 %
6	90 %	76 %	87 %
7	88 %	82 %	89 %
8	81 %	82 %	97 %
9	89 %	85 %	85 %
10	89 %	76 %	93 %

- Finally for the ***uniform*** hypothesis, the obtained acceptance percentage table with three different significant levels was:

Table (7)
The acceptance percentages for the U(0,4) distribution.

The Iterations	Significant levels (α)		
	0.025	0.05	0.1
1	78 %	79 %	79 %
2	75 %	83 %	76 %
3	76 %	74 %	76 %
4	76 %	76 %	75 %
5	69 %	81 %	82 %
6	84 %	78 %	79 %
7	71 %	83 %	75 %
8	62 %	80 %	75 %
9	75 %	73 %	85 %
10	83 %	81 %	80 %

5. Conclusions and Suggestions:

Throughout this work, a large scale of ***Monte Carlo*** investigation is made for the two basic problems, namely the probability of extinction and the distribution of the offspring number in the n^{th} generation, where **n** is a prescribed fixed number representing the maximum number of generations.

- The board conclusions reached in finding the probability of extinction were:

1. In general, the probability of extinction of any generation is larger than its value in the proceeding generation for all assumed distributions.
2. In the investigation for the probability of extinction with the ***binomial*** distribution assumption, the conclusions observed were:
    I. The largest value of this probability in the n^{th} generation for all the iterations is ranging over the interval [0.1107, 0.1112].
    II. This probability is stable in much successive iterations although there

appear some differences due to randomness.

III. We expect that if the number of iterations increases then this probability will increase and have the value 0.1112 as an upper bound.

IV. When **r**=2 the theoretical value of this probability seems to be close to its corresponding practical values but this closeness appears more clearly when **r**=3, where **r** is the number of offspring produced by a single parent.

3. For the probability of extinction with the ***Poisson*** distribution assumption, the investigation lead to the following:
   I. The largest value of this probability in the n^{th} generation for all the iterations is ranging over the interval [0.2005, 0.2032].
   II. Tables of the probability of extinction show stability in the fourth to the ninth iteration and so it could be expected that this probability would be the same even in the tenth iteration and more.
   III. It was seen that the theoretical value of this probability when **r** = 2 seems to be close to its corresponding practical values but this closeness appears more clearly when r = 3.
4. And for the probability of extinction with the ***uniform*** distribution assumption, it is seemed that:
   I. The largest value of this probability in the n^{th} generation for all the iterations is ranging over the interval [0.3819, 0.4544].
   II. The values of the probability of extinction with this assumption increase step by step from iteration to another.
   III. Tables of the probability of extinction show different increasing values for this probability, i.e. no stability is found for this probability with the uniform distribution assumption but the upper bound to these values seems to be 0.5 as the no. of iterations increases.
   IV. It was seen that the theoretical value of this probability when r = 2 does not seem to be close to its corresponding practical values but it is close when r = 3.
5. Overall aggregate the good observer will notice that the value of the probability of extinction with the ***uniform*** distribution assumption reaches one faster than its value with the two other assumptions and the next is with the ***Poisson*** distribution assumption.

- For the problem of finding the offspring distribution, the following were observed :

I. For the ***binomial*** distribution assumption it was found that its tables of the acceptance percentages consists frequently of small percentages for all the significant levels, where these values (percentages) differ from 3 % to 30 % , which will be rejected.

II. For the ***Poisson*** distribution assumption it was found that its tables of the acceptance percentages consists of good results for all the significant levels and these results differ from 71 % to 97 % , which will be accepted.

III. For the ***uniform*** distribution assumption it was found that its tables of the acceptance percentages consists also of quite good results for all the significant levels and these results differ from 62 % to 85 % which person could be a little hesitate to accept it.

iv. ***Poisson*** distribution is the most suitable distribution in comparison with the two other considered distributions (namely ***binomial***, and ***uniform***) to the offspring distribution according to the results from the ***Chi***-square goodness of fit test and the ***uniform*** distribution may also be considered as a next suitable one, but the ***binomial*** distribution seems to be, according to the C***hi***-square goodness of fit test, not suitable as a distribution to the offspring number in the n^{th} generation.

- In branching processes, a relation or more between the probability of extinction and the offspring distribution is expected and that requires more research to be made which is kept open to the other researchers in future.
- Another suggestion is that other researchers could make further assumptions depending on other distributions for the two problems and test their hypotheses by suitable tests.

6.References and Bibliographies:

[1] Al-Dabbagh, S. A., "On Branching processes", MSc. thesis, Al-Nahrain University, 2002.

[2] Athreya, K. B., and P. Jagers, "Classical and Modern Branching Processes, Papers from IMA workshop held at the University of Minnesota, Minneapolis, MN, June 13-17, 1994, IMA volumes in Mathematics and Applications, 84, Spriger-Verlag, New York, 1997. Cited by Athreya and Vidiashankar in [3].

[3] Athreya, K. B., and A. N. Vidyashankar, "Branching Processes". Internet online research: http://citeseer.nj.nec.com/rd/0,211428,1,0.25,Download/http://citeseer.nj.nec.com/cache/papers/cs/89/athreya99branching.pdf.

[4] Becker, N., "Estimation for discrete-time branching processes with applications to epidemics". Biometrics 33 (1977), 515-522. Cited by Guttorp and Lockhart in [9].

[5] Bienaymé, I. J., "De la loi de multiplication et de la durée des familles". Soc. Philomath. Paris Extraits, Sér. 5, 37-39, 1845. Cited by Jagers in [13].

[6] Ewens, W. J., "Population Genetics". London:Methuen, 1969. Cited by Guttorp and Lockart in [9].

[7] Galton, Francis, "Problem 4001", Educational Times,1 April, p. 17, 1873. Cited by Jagers in [13].

[8] Galton, F., and H. W. Watson, "On the probability of the extinction of families". Journal of Anthropol. Soc. London (Royal Anthropol. Inst. G. B. Ireland) 4, 138-144, 1875. Cited by Jagers in [13].

[9] Guttorp, P., and Richard A. Lockhart, "Maximum likelihood estimation of the offspring variance in a Bienaymé-Galton-Watson branching process". Internet online research: http://www.stat. sfu.ca/~lockhart/richard/paper/paper.html.

[10] Harris, T. E., "branching processes". Journal of The American Mathematical Society 19(1948), 474-494.

[11] Harris, T. E., "The theory of branching processes". Springer-Verlag, Berlin, 1963.

[12] Heyde, C. C., and E. Seneta, "The simple branching process, etc.; a historical note on I. J. Bienaymé". Biometrica, 59 (1972), 680-683. Cited by Jagers in [13].

[13] Jagers, P., "branching processes with Biological Applications". Wiley Interscience, New York, 1975.

[14] Kendall, D. G., "Some problems in the theory of queues". Journal of the Royal statistical society, ser. B 13 (1951), 151-185. Cited by Guttorp and Lockart in [9].

[15] Kendall, D. G., "Branching processes since 1873". Journal of London Mathematical Society 41 (1966), 385-406.

[16] Kendall, David G., "The genealogy of genealogy, Branching processes before (and after) 1873". Bulletin of London mathematical Society 7 (1975), 225-253.

[17] Keyfitz, N., "Applied mathematical demography". 2nd edition, New York:Springer, 1985. Cited by Guttorp and Lockart in [9].

[18] Snell, J. Laurie, "Introduction to probability". Mc Graw-Hill, 1989.

[19] Steffensen, J. F., "Om sandsynligheden for at afkommet uddØr. Matematisk Tidsskrift B 1 (1930), 19-23. Cited by Harris in [11].

[20] Steffensen, J. F., "Deux Problémes du calcul des Probabilitiés". Ann. Inst. H. Poincaré 3 (1932), 319-344. Cited by Harris in [11].

[21] Watson, H. W., "Solution to problem 4001". Educational Times, 1 August, pp. 115-116, 1873. Cited by Kendall in [15].

الخلاصة

في هذا البحث موجز محاكاة مسألتين أساسيتين في العمليات التصادفية الفرعية وكانت أول من طرحها العالم كاليتون وهي مسألة ايجاد أحتمال الانقطاع في العملية التصادفية الفرعية ومسألة تحديد توزيع عدد الافراد للجيل النوني. فرضت ثلاث توزيعات في المحاكاة وهي توزيع بواسون، التوزيع الثنائي والتوزيع المتجانس واستخدمت خوارزميات ونوقشت النائج والاستنتاجات على ضوء مخرجات المحاكاة وكان الاقرب للمسألة الاولى هو التوزيع المتجانس ثم توزيع بواسون، في حين كان توزيع بواسون هو الانسب للمسألة الثانية.