

A Stochastic Policy in Purifying Industrial Water using Branching Model

Maroussia Slavtchova-Bojkova

Department of Probability and Statistics, Institute of Mathematics and
Informatics, Bulgarian Academy of Sciences, 8, G. Bontchev str., 1113 Sofia,
(Bulgaria) e-mail: bojkova@math.bas.bg, fax: (00 359) 2 971 3649

1 Abstract

In general, the model originates from the problem of estimating the waiting time to a successful experiment in the purifying process of industrial wasted water, treated by bacterial culture systems. By "successful experiment" we mean the event when the certain type of bacterial population is adjusted to the media, so that is spreading out in this environment. We present an age-dependent branching model with immigration and theoretically analyze how one can extract exact information about some important characteristics of this model, as the mean reproduction and total progeny. On the other hand, using that model, we treat the problem of inference from expected waiting times and expected progeny on the fertility rates. This policy performs a comparison between different reproduction laws of bacteria and aims to show that the inference is not depending on the exact distribution law. Simulations were made for some reproduction laws and different reproduction means. These models depict accurately the actual problems in practice, because of the uncertainties associated with the conditions and measurements.

Key words: Bellman-Harris branching process, generating functions, cycle length, total progeny, reproduction mean, statistical inference, extinction bias, simulation

2 Introduction

The goal of the paper is to explain some phenomena arising in the biological treatment of wasted water by cells feeding on a substrate in a bio-reactor and give answer to questions which might be of interest for bio-technologists and environmentalists.

First, how long does take the final establishment of bacterial cultures in wasted water laboratory experiments? As there is always a positive probability of extinction, it is possible to have several unsuccessful trials before the bacterial cultures start to grow irreversibly.

Secondly, what conclusions one can draw from an early extinction of a bacterial culture in different types of wasted water? Does it imply that the offspring mean in these environments is low? Similarly, our study might help decision-makers to take a choice based on comparative laboratory results in similar wasted water cultivated with different bacterial strains. In general, such questions related to real world problems of industrial wasted water treatment are hard to answer. One of the major reasons for this difficulty is that the circumstances under which the experiments are made in natural and artificial basins, like lagoons, ponds and lakes, are not always the same.

In general, one of the best features of branching processes is that the exact theoretical results have a natural interpretation and can be directly used for numerical and simulation studies. We present an age-dependent branching model with immigration and the main idea is to point out that the duality between super-critical and sub-critical branching processes given extinction can make decision- makers take the wrong decision.

We would like to point out that in the discrete-time case, the problem concerning the total progeny was investigated by several authors (Jagers (1975), p. 39; Harris (1989), p. 32), however, their main interest is focused on questions of rates of growth, and this conditioned on survival. Karlin and Tavaré (1982) studied the asymptotic behavior of the probabilities of hitting the absorbing states, the times needed to hit these states, and the conditional distributions of the number of particles (for models allowing catastrophes).

The problem of inference from expected waiting times and expected progeny on fertility rates, was first treated by Bruss and Slavtchova-Bojkova (1999). They studied the simple case, in which all newly introduced populations are supposed to behave like independent identically distributed (i.i.d.) Bienaymé-Galton-Watson branching processes. As an example they provided population experiments with trout. Later on, Slavtchova-Bojkova (2000) theoretically generalized these results in continuous time, i.e. when the newly introduced populations are supposed to be Bellman-Harris branching processes, and proved that the results of inference remain valid. It is interesting to see how the duality is applied in a concrete application where it appears in a natural way.

Concerning the computation procedures, Jacobson (1985) was carried out

a numerical investigation of the rate of convergence of the extinction probability for a discrete-time age-dependent branching process. The most significant result illustrated by his simulations was the long tail of the distribution of life period for the case where the mean number of offspring was 1, as opposed to the quick convergence for the other runs.

The main goal of the present paper is to emphasize on the independence of inference on fertility rates from the reproduction law of the particles, provided extinction is observed. Simulation studies are also presented, corresponding to the binary, Poisson and geometric reproduction laws. Since Powell (1955) found that the life-period of bacteria follows a gamma distribution, and reproduction at death is characteristic of bacteria-like organisms, a discretized gamma density was used for all simulations. What is in common for the three cases is that conditioning on extinction makes sub-critical and super-critical processes undistinguishable.

We discuss an example of "extinction bias" which may mislead decision-makers in cultivation experiments.

3 Model Formulation

We will first outline an age-dependent branching process with immigration in the state zero. Consider a population process starting at time 0 with a single progenitor of age 0 whose life-length τ has distribution $G(t) = P(\tau \leq t)$, $G(0^+) = 0$. With probability p_k , $k \geq 0$ it produces at the end of its life k similar individuals (of age 0, with the same life-length τ and reproduction distribution, $(p_k, \sum_{k=0}^{\infty} p_k = 1$). The probability generating function (p.g.f.) of the number of ξ offspring is denoted by

$$f(s) = \sum_{k=0}^{\infty} p_k s^k, \quad |s| \leq 1, p_k = P(\xi = k).$$

Provided that there is at least one offspring, the death-and-reproduction process is repeated, and continues as long as individuals exist.

Let $\tilde{Z}(t)$ be the number of individuals existing in the population at time t or the state of the process $(\tilde{Z}(t))_{t \geq 0}$ at time t . Note that a line becomes extinct once $\tilde{Z}(t) = 0$ for some t (and for all t thereafter), and that the above process is "age-dependent" in the sense that the probability that an individual living at time t dies in the interval $(t, t + dt)$ is, in general, a non-constant function of t . The process $(\tilde{Z}(t))$ is the so-called Bellman-Harris

branching process (see, for example, Athreya and Ney (1972), pp. 137-144). Every time the process $(\tilde{Z}(t))$ hits the state zero-state we suppose to have an immigration of one particle from an outside source. With $(Z(t))_{t \geq 0}$ we shall denote the process with immigration in the state zero. Foster (1971) and Pakes (1971, 1975) first studied the discrete-time version of these processes.

For a branching process with immigration $(Z(t))_{t \geq 0}$ we call life periods (cycles) the intervals $(t_0, t_0 + T)$ of maximal length on which $\inf_{t_0 \leq t \leq t_0 + T} Z(t) > 0$. Thus $(Z(t))$ may have several life-periods, the last one always being infinite, provided the process is supercritical. If the process is sub-critical it will have almost surely (a.s.) infinitely many life periods.

We are interested in the last instant M of immigration, i. e. in the "birth time" of that process which will finally survive forever. Specifically, we shall derive the conditional distribution of the length T of the first life period and the conditional expectation of T , both conditioned on the event. $T < \infty$.

4 Preliminary Results

It is well known from the theory of branching processes (see e. g. Athreya and Ney (1972), pp. 139-144) that the probability q of eventual extinction of Bellman-Harris process $(\tilde{Z}(t))$ is the smallest non-negative root of the equation $f(s) = s$, and $q = 1 \iff m = f'(1) \leq 1$. The parameter m is called the reproduction mean, and the super-critical, critical and sub-critical cases correspond to the relations $m > 1$, $m = 1$ and $m < 1$, respectively. Thus before one attempts to find the smallest root of the above equation, the expected number should first be computed.

Although the probability of eventual extinction can be calculated easily using above equations, more interesting studies concern the probability of extinction at or by a certain time and how quickly the later probabilities approach the theoretical extinction probability. Let l be the maximum number of offspring an individual can have, r be the greatest age an individual can live to, and $g(\cdot)$ be the mortality density. There are two mutually exclusive ways a trajectory can become extinct by time t : the progenitor dies by time t with probability $1 - G(t)$ and leaves no offspring, or the progenitor dies at time $1 \leq s < t$ with probability $g(s)$, having had $1 \leq k \leq l$ offspring and each of the k offspring' lines becomes extinct by time t . If $F(t)$ denotes the probability of extinction by time t one obtains

$$F(t) = p_0 G(t) + \sum_{s=1}^{\infty} (t - s) \sum_{k=1}^l p_k F^k(t - s) g(s)$$

If $t > r$, (1) becomes $F(t) = p_0 + \sum_{s=1}^r \sum_{k=1}^l p_k F^k(t-s)g(s)$ since $G(t) = 1$ for $t \geq r$. Note that the scheme for computing $F(t)$ is recursive.

Let T_1, T_2, \dots be the lengths of the life periods of those consecutive processes dying out before the surviving process is initiated (i. i. d. copies of the r. v. T).

Let $v(t) = P(T \leq t)$.

The following result is used for the computation of the conditional distribution of the length of a life cycle, its conditional expectation (given extinction) and the expectation of the total length of unsuccessful life cycles. The length T of the life-period and the last instant M of immigration of the process ($Z(t)$) have the following properties:

$$P(T \leq t | T < \infty) = \frac{v(t)}{q},$$

$t > 0$, $v(0) = P(Z(0) = 0) = 0$, where q is the probability of eventual extinction of the Bellman-Harris process ($\tilde{Z}(t)$);

$$E(T | T < \infty) = \frac{1}{q} \int_0^\infty 0(2)$$

$$E(M) = \frac{q}{(1-q)} E(T | T < \infty). (3)$$

The proofs of (2) and (3) could be seen in Slavtchova-Bojkova (2000).

5 Numerical results

To study the implications of the above equations for extinction probabilities we use Mapple 6. Nine computer simulations were then made with extinction probabilities up to 150 generations.

In Figure 1 one can see the results based on the binary reproductive rule (either zero or two offspring at death), with mean offspring of 0,5, 1, 1,5, both conditioned and unconditional. The results based on a Poisson and geometrical reproduction laws, again with mean offspring of 0,5, 1, 1,5 in both cases are presented in Figure 2. and Figure 3, respectively.

The most significant result shown in the graphs is that for super- and sub-critical branching models the conditional distributions of the life cycle given extinction coincide. The reason is explained in the next paragraph.

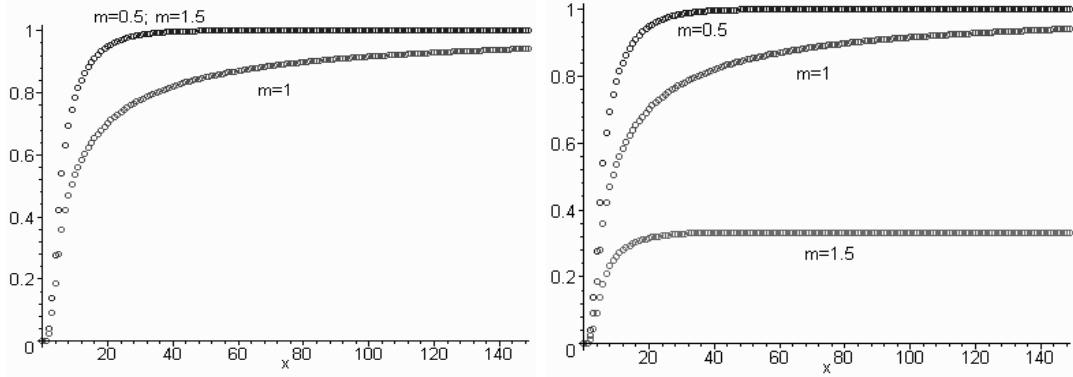


Figure 1. Conditional and unconditional extinction probabilities with binary reproduction law.

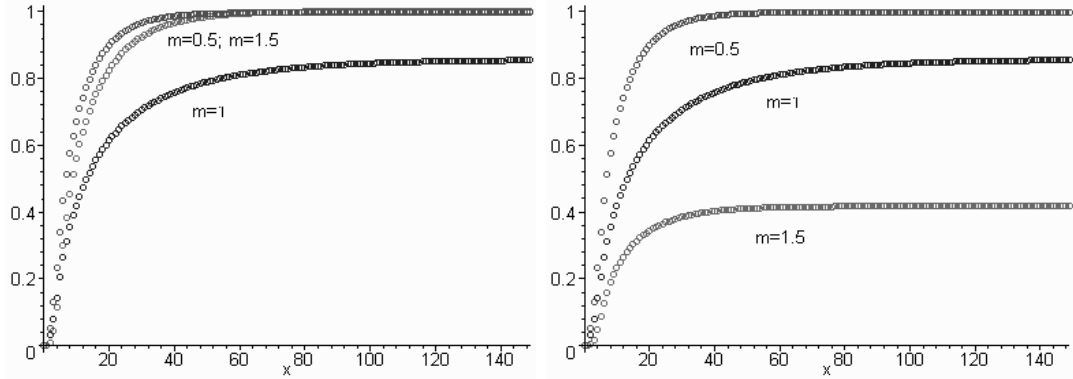


Figure 2. Conditional and unconditional extinction probabilities with Poisson reproduction law.

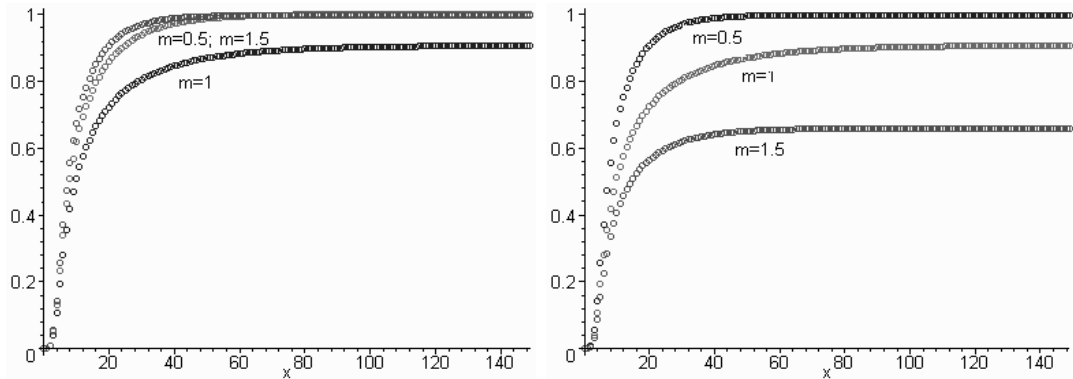


Figure 3. Conditional and unconditional extinction probabilities with geometric reproduction law.

6 Discussion

One must be careful not to draw hasty conclusions after failed experiments. To be specific, suppose that three different types of bacteria (α , β , γ , say) were introduced in similar wasted water and that each of these seems to have disappeared after some time, but that the α -type strains were reported in highest numbers or over the largest period of time. Is it then most promising to continue with α -type bacteria for a new experiment? The frequency of reports must be thought of as being positively correlated with the total progeny and the later with the reproduction mean of that bacterial culture. However such a conclusion would be erroneous. It is the number of times the process becomes extinct before it grows irreversibly, that will help to decide if the process is sub- or supercritical.

7 Conclusion

Extinction entails a very strong bias. If a decision maker decides to try again with that strains (α -type) which seems to have been best adapted so far he may exclude those strains with a much higher fertility rate m . The point is that he has to take that decision after extinction. It is simply very improbable that a process with a "comfortable" mean $m > 1$ would die out late. The higher the mean of a population the more probable it becomes that this population would, after extinction, be excluded from further experiments.

We conclude that the problem is of a greater significance that it might appear at the first sight. Independent control studies to assess prior probability of extinction are likely to be environment-bias.

8 Acknowledgements

This paper was partially supported by NSFI-Bulgaria, Grant No MM-1101/01.

9 References

- Athreya, K. and Ney, P.(1972): Branching Processes. Springer Verlag, Berlin, 287 pp.
- Bruss, F. T. and Slavtchova-Bojkova, M.(1999): On waiting times to populate an environment and a question of statistical inference. J. Appl. Probab., v. 36, 261-267.
- Foster, J. H.(1971): A limit theorems for a branching process with state-dependent immigration. Ann. Math. Stat.,v. 42, 1773-1776.
- Harris, T. (1989): The Theory of Branching Processes. Dover Publications Inc., New York, 230 pp.
- Jacobson, M. E.(1985): Computation of Extinction Probabilities for the Bellman-Harris Branching Processes. Math. Biosci., v. 77, 173 - 177.
- Jagers, P.(1975): Branching Processes with Biological Applications. Wiley, 268 pp.
- Karlin, S. and Tavaré, S.(1982): Detecting particular genotypes in populations under nonrandom mating. Math. Biosci., v. 59, 57-75.
- Pakes, A. G.(1971): A branching processes with a state-dependent immigration component. Adv. Appl. Prob., v. 3, 301-314.
- Pakes, A. G.(1975): Some results for non-supercritical Galton-Watson processes with immigration. Math. Biosci., v. 24, 71-92.
- Powell, E. D.(1955): Some features of the generation times of individual bacteria. Biometrika, v. 42, 16-44.
- Slavtchova-Bojkova, M.(2000): Computation of waiting time to successful experiment using age-dependent branching model. Ecological Modelling , v. 133, 125-130.