

Summary of the work entitled *Inference in population-size-dependent branching processes*

Carmen Minuesa Abril

Joint work with Peter Braunsteins and Sophie Hautphenne

1 Context

The current work belongs to the field of Markov Stochastic Processes and more specifically, to the theory of Branching Processes. This theory finds its origin in the middle of the 19th century, motivated by the study of the extinction of certain family lines of the European aristocracy. However, the term “branching process” is more recent and due to Kolmogorov and Dmitriev (1947).

A common feature of branching processes is that they allow for the description of system dynamics where individuals (or elements, in general) reproduce independently of each others. Thus, the application areas of these processes are diverse: biology, epidemiology, genetics, medicine, nuclear physics, demography, actuarial mathematics, and algorithm and data structures.

2 State of Art

Since the appearance of the pioneer branching process, the Galton-Watson process, the theory of branching processes has experienced a consolidated growth and nowadays, it counts a rich range of discrete-time and continuous-time models. In view of its potential applications, this work focusses on a class of discrete-time branching processes: the *population-size-dependent branching processes* (PSDBPs).

These models were initially introduced in Labkovskii (1972). Similarly to the Galton-Watson process, the reproduction of each individual in the same generation is independent of the others. However, in the PSDBPs the probability distribution governing the reproduction

of individuals at each generation, called the offspring distribution, may depend on the current population size. This is a realistic assumption in several biological contexts. For instance, when the resources of an environment are limited, we cannot expect an indefinite growth of the population: the larger the number of individuals, the sparser the resources to support them. Thus, we expect a reduction in the reproductive capacity of the individuals when the population size is above a threshold value known as the *carrying capacity* of the ecosystem (see Jagers (2011)). Other situations can be found in other fields, such as chemistry. One example is the polymerase chain reaction (PCR), where DNA molecules replicate. In this context, individuals are identified with molecules, and their capacity to replicate is reduced when the reaction components become limiting as a consequence of a rise in the number of molecules present in the reaction (see Chigansky et al. (2018)). Other applications include blood cell or bone marrow cell formation (see Klebaner (1984b) for further examples).

Most recent research on PSDBPs tackles the description of branching models with a carrying capacity, both in continuous-time (see Jagers and Klebaner (2011); Hamza et al. (2016); Sagitov and Shaimerdenova (2013)) and in discrete-time (see Jagers (2011); Athreya and Schuh (2016)). The carrying capacity K is a threshold value that determines the critical behaviour of the process in the following way: when the current population size is below K , then the expected number of individuals in the following generation increases, while if the population size is above K , then we expect a drop in the number of individuals in the following generation. In the case where the population size equals K (if K is an integer), the expected number of individuals in the following generation is the same as in the current one.

The flexibility of the PSDBP versus other branching processes comes at a cost: the dependence on the current population size complicates the asymptotic analysis of PSDBPs (see Klebaner (1984a,b, 1985, 1993)), but also their statistical inference. As a consequence, only a few number of papers address the parameter estimation in PSDBPs (see Lalam and Jacob (2004); Lalam et al. (2004)). However, like in the majority of branching processes, research on inferential theory is of great importance given the strong relation between the value of the parameters of the processes and their behaviour.

A common feature of the above works on inference for PSDBPs is that they tackle the parametric estimation of the offspring distribution in cases where the processes have a positive chance of survival, and the asymptotic properties of the resulting estimators are stated in the non-extinction set. To our knowledge, no previous research has been performed regarding the estimation of the parameters in PSDBPs which become extinct with probability one, and this has motivated the study in the present paper. This question is of special interest in order to develop strategies for the management of species which are threatened by restricted habitat, where accurate approximations of the main parameters of the proposed models are essential.

3 Main contributions

In this work, we investigate for the first time the maximum likelihood estimation in PS-DBPs which become almost surely extinct. First, for a general PSDBP and assuming that we observe the number of individuals giving birth to k offspring, for all $k \geq 0$, until some generation, we determine the maximum likelihood estimators (MLEs) for the offspring distribution, the offspring mean, and the offspring variance at each population size. In addition, we provide the expectation of the MLEs for the offspring distribution and the offspring mean.

Second, we consider the less detailed sample containing the population sizes up to some generation n . We show that the MLE of the offspring mean $m(z)$ at a fixed population size z based on that sample, which we denoted by $\hat{m}_n(z)$, is exactly the MLE based on our initial sample up to generation n . Focussing our attention on the study of PSDBPs whose extinction is certain, we analyse the asymptotic behaviour of the MLE $\hat{m}_n(z)$ for a population size z fixed. In order to apply the law of large numbers to obtain consistency of the estimator, we need to condition on the survival of the process. Our results show that the behaviour of our estimator in this setting is strongly related to the Q -process associated with the original PSDBP, which can be interpreted as the Markov chain defined as the original process conditioned on not becoming extinct in the distant future. We prove that our estimator satisfies a new type of consistency, which we name Q -consistency. This states that, conditionally on $Z_n > 0$, $\hat{m}_n(z)$ converges “weakly” to $m^\uparrow(z)$ as $n \rightarrow \infty$, where the constant $m^\uparrow(z)$ is the counterpart in the Q -process of the offspring mean at the population size z in the original process.

Finally, we illustrate our results numerically through examples of branching processes with a carrying capacity and binary fission reproduction, using the statistical software and programming environment R. For these processes, we also analyse the difference between the functions $m^\uparrow(z)$ and $m(z)$, for $z \geq 1$. Our numerical results show that the discrepancy between these two functions is insignificant and decreases when the carrying capacity K increases. This allows us to conclude that, at least for the branching processes with a carrying capacity and binary fission reproduction, our estimators provide good approximations for the true value of the parameters.

References

- K. B. Athreya and H. B. Schuh. A Galton-Watson process with a threshold. *Journal of Applied Probability*, 53:614–621, 2016.
- P. Chigansky, P. Jagers, and F. C. Klebaner. What can be observed in real time PCR and when does it show? *Journal of Mathematical Biology*, 76(3):679–695, 2018. doi: 10.1007/s00285-017-1154-1.

- K. Hamza, P. Jagers, and F.C. Klebaner. On the establishment, persistence, and inevitable extinction of populations. *Journal of Mathematical Biology*, 72(4):797–820, 2016.
- P. Jagers. Extinction, persistence, and evolution. In F. A. C. C. Chalub and J. F. Rodrigues, editors, *The Mathematics of Darwin's Legacy*, pages 91–104. Springer Basel, 2011.
- P. Jagers and F.C. Klebaner. Population-size-dependent, age-structured branching processes linger around their carrying capacity. *Journal of Applied Probability*, 48(A):249–260, 2011.
- F. C. Klebaner. Geometric rate of growth in population-size-dependent branching processes. *Journal of Applied Probability*, 21(1):40–49, 1984a.
- F. C. Klebaner. On population-size-dependent branching processes. *Advances in Applied Probability*, 16(1):30–55, 1984b.
- F. C. Klebaner. A limit theorem for population-size-dependent branching processes. *Journal of Applied Probability*, 22(1):48–57, 1985.
- F. C. Klebaner. Population-dependent branching processes with a threshold. *Stochastic Processes and their Applications*, 46(1):115–127, 1993.
- A. N. Kolmogorov and N. A. Dmitriev. Branching stochastic processes. *Compte rendu de l'Académie des sciences de l'URSS*, 56:5–8, 1947.
- V. A. Labkovskii. A limit theorem for generalized random branching processes depending on the size of the population. *Theory of Probability and its Applications*, XVII(1):72–85, 1972.
- N. Lalam and C.; Jacob. Estimation of the offspring mean in a supercritical or near-critical size-dependent branching process. *Advances in Applied Probability*, 36(2):582–601, 2004.
- N. Lalam, C. Jacob, and P. Jagers. Modelling the PCR amplification process by a size-dependent branching process and estimation of the efficiency. *Advances in Applied Probability*, 36(2):602–615, 2004. doi: 10.1239/aap/1086957587.
- S. Sagitov and A. Shaimerdenova. Extinction times for a birth–death process with weak competition. *Lithuanian Mathematical Journal*, 53(2):220–234, 2013.