# A random walk with catastrophes

Iddo Ben-Ari[*][†]     Alexander Roitershtein[‡]     Rinaldo B. Schinazi [§]

September 9, 2017; Revised: January 15, 2019

**Abstract**

Random population dynamics with catastrophes (events pertaining to possible elimination of a large portion of the population) has a long history in the mathematical literature. In this paper we study an ergodic model for random population dynamics with linear growth and binomial catastrophes: in a catastrophe, each individual survives with some fixed probability, independently of the rest. Through a coupling construction, we obtain sharp two-sided bounds for the rate of convergence to stationarity which are applied to show that the model exhibits a cutoff phenomenon.

*MSC2010:* Primary 60J10, 60J80, secondary 92D25, 60K37.
*Keywords*: population models, catastrophes, persistence, spectral gap, cutoff.

## 1 Introduction

### 1.1 Model

Consider a population with the following birth and death rules. Given two parameters $p \in (0,1)$ and $c \in (0,1]$, the population size is a discrete-time Markov chain $(X_t : t \in \mathbb{Z}_+)$ on the state-space $\mathbb{Z}_+$ of non-negative integers $(\mathbb{Z}_+ := \mathbb{N} \cup \{0\})$ with transition function

$$\mathfrak{p}^{p,c}(i,j) = \begin{cases} p & j = i+1 \\ (1-p)\binom{i}{j}(1-c)^j c^{i-j} & j \in \{0,\dots,i\} \end{cases}$$

When there is no risk of ambiguity, we will omit the superscripts $p, c$ and write $\mathfrak{p}$. In words, conditioned on the history of the process up to time $t$, the population size at time $t+1$ is determined by tossing an independent coin with probability $p$ of success. In the case of success, the population increases by 1, and in the case of a failure, also known as a

---

[*]Department of Mathematics, University of Connecticut, Storrs, CT 06269-1009, USA; iddo.ben-ari@uconn.edu

[‡]Deptartment of Statistics, Texas A&M University, College Station, TX 77843; alexander@stat.tamu.edu

[§]Department of Mathematics, University of Colorado, Colorado Springs, CO 80933-7150, USA; rschinaz@uccs.edu

*catastrophe*, the population is an independent binomial with parameters $X_t$ and $1 - c$. That is, in a catastrophe, each individual survives with probability $1 - c$ independently of the other, and is otherwise killed. Note that $\mathfrak{p}$ is aperiodic and irreducible. As will be shown below, $\mathfrak{p}$ is also geometrically ergodic in total variation.

The model is a version of subcritical branching process (the catastrophes) with linear migration (population increase), and belongs to a larger class of stochastic models with catastrophes extensively studied in the literature. The term "catastrophe" loosely refers to events where a large proportion or the entire population may be wiped out. There are many ways to model catastrophes and several have been studied, see Section 1.3 below. The particular model we study corresponds to binomial catastrophes of [25, Section 2].

## 1.2 Motivation

Our original interest in the model came from a curiously strong persistence feature we observed in simulations: repulsion from zero and long fluctuations in a narrow band before first hitting zero. Figure 1.2 shows a simulation of the model for $p = 0.4$ and $c = 0.01$, between times 0 and $10^5$. The initial population size is $X_0 = 10$. The population climbs quickly and fluctuates in a narrow band around an empirical mean close to 66.667 for a very long time. In Section 2 we show that the process is mean-reverting relative to the mean of its stationary distribution. Corollary 4.5 shows that already after 1500 steps the total variation distance between the process and its stationary distribution is bounded above by 0.001. These, along with the fact that the expectation of the first extinction time is of the order $10^{24}$, shown in Section 2, give at least a partial explanation to the simulations.

Additional motivation for our work on the model is in its amenability to coupling methods yielding sharp bounds on the rate of convergence to stationarity. These allow us to prove that the process exhibits the cutoff phenomenon. These results form the bulk of our work.

## 1.3 Literature

Stochastic models with catastrophes are studied in mathematical literature since mid-1970's, for a first systematic account and a review of the early literature see [3]. For a motivation and background in biological sciences see, for instance, [9, 10, 19, 23]. Most of the work in the literature concern with either continuous time (generalized) birth and death chains with catastrophes or ODE-based models with a random disturbance. For a recent review and an extensive bibliography see [16]. The persistence feature is discussed for models with catastrophes in, for instance, [6, 23]. We remark that despite the variety of mathematical approaches to modeling population catastrophes, some results seem to be of a universal nature and are exhibited by models of different types. As an example, we mention the logarithmic dependence of the first extinction time on the initial population size which we discuss in Section 5.2.3.

As mentioned above, the model we study is a particular version of binomial catastrophes case in the model introduced by Neuts in [25, Section 2]. A continuous-time analogue of our model was introduced in [3, Section 4]. For recent progress, see [1, 7, 16]. In our model, deaths occur in a branching fashion, and in Section 5.1 we reformulate and discuss the model as a special branching process with immigration in a random environment. The study of
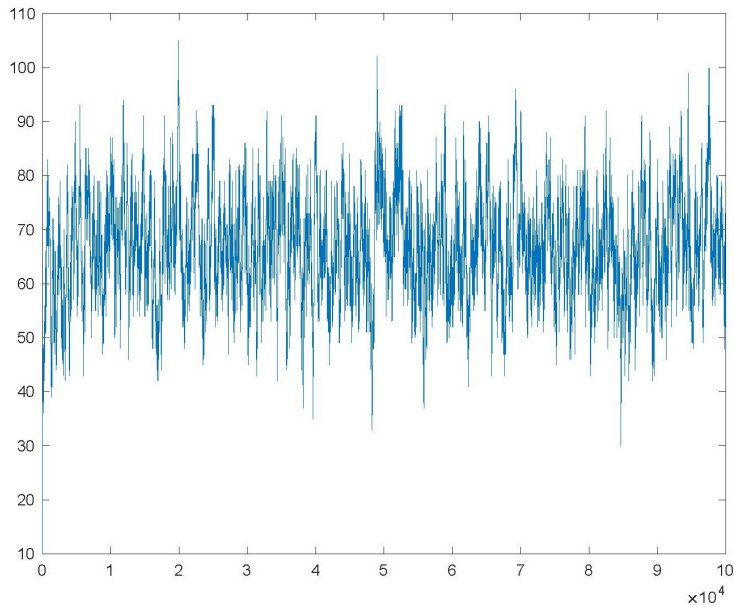
Figure 1: Long fluctuations of the random walk around its average before hitting zero

branching processes as models of population growth with catastrophes (or disasters) goes back to at least [15], where a branching process without immigration is considered. Due to their tractability, much attention in the literature has been received by models with a deterministic growth between catastrophes, so called semi-stochastic models [5, 10, 11, 20].

Many results in the literature focus on the phase transition between survival and non survival, see [2] and [14]. The results concerning first extinction times and stationary distributions are typically given in terms of Laplace transform or generating functions, see [2] and [16].

## 1.4 Organization

In Section 2 we give a probabilistic representation of the stationary distribution of the process. The bulk of our contribution is reported in Sections 3 and 4. In Section 3 we introduce a coupling and use it to compute sharp bounds on the total variation distance between the distributions of the process starting from two different initial states. In Section 4 we consider a sequence of models whose stationary distribution converges to a Poisson limit. We show that this sequence exhibits a cutoff phenomenon, namely on a certain time scale the total variation distance to the stationary distribution drops from one to zero in a narrow time window. Our study of both topics appear to be original in the context of stochastic models with catastrophes and we are not aware of similar results in the literature for any type of such models. Finally, in Section 5 we estimate the first extinction time and we use a branching representation for several purposes.

Throughout the paper, the notation $a_n \sim b_n$ stands for $\lim_{n\to\infty} \frac{a_n}{b_n} = 1$, and $X \overset{d}{=} Y$ indicates that the random variables $X$ and $Y$ have the same distribution.

# 2 Stationary distribution

## 2.1 Representation formula

Given a $\mathbb{Z}_+$-valued random variable $R$ and $\varepsilon \in [0, 1]$, write $\text{Bin}(R, \varepsilon)$ for the random variable which, conditioned on $R$, is binomial with parameters $R$ and $\varepsilon$. We begin with the following lemma whose proof is omitted.

**Lemma 2.1.** *Suppose that $R_0, R_1 \ldots$ are independent $\mathbb{Z}_+$-valued random variables and let $\varepsilon_0, \varepsilon_1, \varepsilon_2, \ldots$ be a sequence taking values in $[0, 1]$. Assume that $\sum \varepsilon_j E[R_j] < \infty$. For $j = 0, 1, \ldots$, let $Bin_j(R_j, \varepsilon_j)$ be $Bin(R_j, \varepsilon_j)$-distributed, with $(Bin_j(R_j, \varepsilon_j) : j \geq 1)$ independent, conditional on $(R_j : j \in \mathbb{Z}_+)$. Let $R = \sum_{j=0}^{\infty} Bin_j(R_j, \varepsilon_j)$ and $\varepsilon > 0$. Then the distribution of $\sum_{j=0}^{\infty} Bin_j(R_j, \varepsilon\varepsilon_j)$ coincides with the distribution of $Bin(R, \varepsilon)$.*

For $\alpha \in (0, 1]$, write $\text{Geom}^-(\alpha)$ for the shifted Geometric distribution with probability mass function equal to $(1 - \alpha)^k \alpha, k \in \mathbb{Z}_+$. Observe that if $R_0 \overset{d}{=} \text{Geom}^-(\alpha)$, then

$$E[s^{R_0}] = \alpha \sum_{k=0}^{\infty} s^k (1 - \alpha)^k = \frac{\alpha}{1 - (1 - \alpha)s}, \qquad s \in [0, 1]. \tag{1}$$

The following proposition gives the stationary distribution for $X$. Note that [25, formula (12)] gives the generating function of the stationary distribution for a class of Markov chains. Our model is in that class. The next proposition gives a probabilistic representation of the stationary distribution for our model. An interpretation through branching process representation is discussed in Section 5.1.

**Proposition 2.2.** *Let $R_0, R_1, R_2, \ldots$ be IID $Geom^-(1 - p)$, $\varepsilon_j = (1 - c)^j$ for $j \in \mathbb{Z}_+$. Let $R$ be as in Lemma 2.1 with $\epsilon = 1$, and let $\pi$ be its distribution. Then $\pi$ is stationary for $\mathfrak{p}$.*

In the degenerate case $c = 1$, $\pi$ is $\text{Geom}^-(1 - p)$-distributed. In Section 5.1 we discuss the case when $p$ and $c$ are both close to one.

*Proof of Proposition 2.2.* Suppose that $X_0 \overset{d}{=} R$. We verify that $X_1 \overset{d}{=} R$ through the generating function of $X_1$. For $s \in [0, 1]$, we have

$$E[s^{X_1}] = psE[s^{X_0}] + (1 - p)E\left[s^{\text{Bin}(X_0, 1-c)}\right]. \tag{2}$$

By Lemma 2.1, we have that

$$\text{Bin}(X_0, 1 - c) \overset{d}{=} \sum_{j=0}^{\infty} \text{Bin}_j\left(R_j, (1 - c)(1 - c)^j\right) \overset{d}{=} \sum_{j=1}^{\infty} \text{Bin}_j\left(R_j, (1 - c)^j\right).$$

Note that for $j = 0$, $Bin_j\left(R_j, (1 - c)^j\right) = R_0$. Hence, the sum of $\text{Bin}(X_0, 1 - c)$ and $R_0$ has the same distribution as $X_0$. In other words:

$$E\left[s^{\text{Bin}(X_0, 1-c)}\right] \cdot E[s^{R_0}] = E[s^{X_0}].$$

Thus (2) becomes

$$E[s^{X_1}] = \left(ps + \frac{1-p}{E[s^{R_0}]}\right)E[s^{X_0}] = E[s^{X_0}],$$

where the last identity is due to (1) with $\alpha = 1 - p$. $\qquad\qquad\square$

Before continuing to our next topic we briefly discuss several related observations. Using Proposition 2.2 and identity (1) we have

$$\pi(0) = \prod_{j=0}^{\infty} P\big(\text{Bin}_j\big(R_j, (1-c)^j\big) = 0\big) = \prod_{j=0}^{\infty} E\big[\big(1 - (1-c)^j\big)^{R_j}\big]$$

$$= \prod_{j=0}^{\infty} \frac{1-p}{1 - p(1 - (1-c)^j)}.$$

Let $\tau$ be the hitting time of 0, or the first extinction time,

$$\tau = \inf\{t \geq 1 : X_t = 0\}. \tag{3}$$

Thus,

$$E_0[\tau] = \frac{1}{\pi(0)} = \prod_{j=0}^{\infty}\left(1 + \frac{p}{1-p}(1-c)^j\right).$$

In the biological literature, this expected value is often referred to as the *persistence time* of the model [5, 22, 23]. Using that

$$x - \frac{x^2}{2} \leq \ln(1+x) \leq x, \qquad \forall\, |x| < 1, \tag{4}$$

we get for $p < 1/2$ that

$$\frac{p}{c(1-p)} - \frac{1}{2}\frac{p^2}{(1-p)^2(1 - (1-c)^2)} \leq \ln E_0[\tau] \leq \frac{p}{c(1-p)}.$$

For example, for $p = 0.4$ and $c = 0.1$, we get $E_0[\tau] \geq 244$. For $p = 0.4$ and $c = 0.01$, we get $E_0[\tau] \geq 10^{24}$.

## 2.2 Mean Reversal

It follows from Proposition 2.2 that

$$\mu := E_\pi[X_t] = \sum_{j=0}^{\infty} E[R_j](1-c)^j = \frac{p}{c(1-p)}. \tag{5}$$

Note that the local drift of $X$

$$\delta_t := E[X_{t+1}|X_t] - X_t = p - (1-p)cX_t = p\left(1 - \frac{X_t}{\mu}\right) \tag{6}$$

5

has the sign opposite to the deviation from $\mu$. Thus the random walk always drifts toward its expected value. We also comment that the probability to hit 0 in the next step decays geometrically with the state of the system, that is

$$P(X_{t+1} = 0 | X_t) = (1 - p)c^{X_t}.$$

These observations suggest that the process will tend to fluctuate about its mean before the first extinction, as can be seen in the simulation, see Figure 1.2.

# 3    Coupling and convergence to stationarity

## 3.1    Construction of the coupling

The key result of this section is a coupling of the probability laws $P_x(X_t \in \cdot)$ and $P_y(X_t \in \cdot)$, obtained from a simple representation of the process.

Let $x, y \in \mathbb{Z}_+$ with $x < y$. Set $X_0 = x$, $X_0' = y$ and $H_0 = y - x$. We continue inductively, assuming $((X_s, X_s', H_s), s \leq t)$ were defined and $X_s' = X_s + H_s$ for all $s \leq t$. Conditioned on $((X_s, X_s', H_s), s \leq t)$,

- With probability $p$, independently of the past, $H_{t+1} = H_t$, $X_{t+1} = X_t + 1$ and $X_{t+1}' = X_t' + 1$.

- Otherwise, that is with probability $1 - p$, set

  $$X_{t+1} = \text{Bin}(X_t, 1 - c) \text{ and } H_{t+1} = \text{Bin}(H_t, 1 - c),$$

  independent of each other and of the past. Moreover, set $X_{t+1}' = X_{t+1} + H_{t+1}$.

It immediately follows that $X$ and $X'$ are both copies of our Markov chain and that

$$X_t' = X_t + H_t$$

for all $t$. In addition, the process $(H_t : t \in \mathbb{Z}_+)$ is non-increasing. Write $P_{x,y}$ and $E_{x,y}$ for the joint distribution and expectation of $X$ and $X'$. Let $\xi$ be the coupling time of the two marginal processes, that is

$$\xi = \inf\{t \geq 0 : X_t = X_t'\} = \inf\{t \in \mathbb{Z}_+ : H_t = 0\} \tag{7}$$

If $H_t > 0$, then $H_{t+1} = H_t$ with probability equal to

$$p + (1 - p)(1 - c)^{H_t} < p + (1 - p)(1 - c) = 1 - c(1 - p).$$

Therefore, it immediately follows that under $P_{x,y}$, $\xi$ is stochastically dominated by a sum of $y - x$ independent copies of Geometric random variables with parameter $(1 - p)c$. Hence, $\xi < \infty$, $P_{x,y}$-a.s. and has a geometric tail. Furthermore, $X_t = X_t'$ for all $t \geq \xi$. Let $N_t$ denote the number of catastrophes up to time $t$. Then $N_t \overset{d}{=} \text{Bin}(t, 1 - p)$. It follows from the construction of the coupling that

$$P_{x,y}(H_t \in \cdot | N_t) \overset{d}{=} \text{Bin}(y - x, (1 - c)^{N_t}). \tag{8}$$

Therefore,

$$P_{x,y}(\xi > t) = P_{x,y}(H_t > 0) = 1 - E[(1 - (1 - c)^{N_t})^{y-x}] \leq (y - x)E[(1 - c)^{N_t}], \qquad (9)$$

where the last inequality is due to Bernoulli's inequality. Letting

$$\alpha = p + (1 - p)(1 - c) = 1 - c(1 - p), \qquad (10)$$

we have that

$$E[(1 - c)^{N_t}] = \alpha^t. \qquad (11)$$

Therefore

$$P_{x,y}(\xi > t) \leq (y - x)\alpha^t. \qquad (12)$$

We comment that this bound is asymptotically sharp as $t \to \infty$. That is

$$P_{x,y}(\xi > t) \sim (y - x)E[(1 - c)^{N_t}] = (y - x)\alpha^t, \qquad (13)$$

as can be seen by expanding the expression $(1 - (1 - c)^{N_t})^{y-x}$ through the binomial theorem and taking expectation.

## 3.2   Upper bounds on total variation

Recall that the total variation distance between two probability measures $Q_1$ and $Q_2$ on $\mathbb{Z}_+$ is defined as

$$\|Q_1 - Q_2\|_{TV} = \max_{A \subset \mathbb{Z}_+} |Q_1(A) - Q_2(A)| = \max_{A \subset \mathbb{Z}_+} (Q_1(A) - Q_2(A)).$$

For $x, y, t \in \mathbb{Z}_+$, let

$$d_t(x, y) := \|P_y(X_t \in \cdot) - P_x(X_t \in \cdot)\|_{TV}.$$

By Aldous' coupling inequality [29], $d_t(x, y) \leq P_{x,y}(\xi > t)$. By combining this inequality and (12) we have proved

**Proposition 3.1.** *Let $\alpha = 1 - c(1 - p)$. Then for $x, y, t \in \mathbb{Z}_+$,*

$$d_t(x, y) \leq |y - x|\alpha^t.$$

Recall from (5) that

$$\mu = \sum_y y\pi(y) = \frac{p}{c(1 - p)}.$$

We have

**Corollary 3.2.** *For all $x, t \in \mathbb{Z}_+$,*

$$d_t(x, \pi) = \|P_x(X_t \in \cdot) - \pi\|_{TV} \leq \left(x - \mu + 2\sum_{y>x}(y - x)\pi(y)\right)\alpha^t.$$

*In particular,*

$$d_t(0, \pi) \leq \mu\alpha^t.$$

7

*Proof.* For any $A \subset \mathbb{Z}_+$,

$$|P_x(X_t \in A) - \pi(A)| \leq \sum_{y=0}^{\infty} |P_x(X_t \in A) - P_y(X_t \in A)| \, \pi(y) \leq \sum_y |y - x| \pi(y) \alpha^t,$$

where the second inequality follows from Proposition 3.1. The result follows because of the identity $\sum_y |y - x| \pi(y) = x - \mu + 2 \sum_{y>x} (y - x) \pi(y)$. $\qquad\square$

## 3.3  Lower bounds on total variation

The goal of this section is to obtain a lower bound for $d_t(x, y)$ which is of the same order as the upper bound in Proposition 3.1. We comment that the difficulty in proving such a result stems from the fact that the state space is infinite, because couplings which preserve linear ordering on a finite state space always satisfy this property, see [4] for a proof in continuous-time setting.

We need to introduce some notation. Let

$$\widetilde{p} = \frac{p}{\alpha} = \frac{p}{1 - c(1 - p)}.$$

The notation $P_x^{(\widetilde{p})}$ is the law of the Markov chain $X$ with initial state $X_0 = x$, and transition function $\mathfrak{p}^{\widetilde{p},c}$. We will also refer to the corresponding stationary distribution as $\pi^{(\widetilde{p})}$.

The main result of this section is the following theorem.

**Theorem 3.3.** *Let $x, y, t \in \mathbb{Z}_+$ with $x < y$. Then*

$$d_t(x, y) \geq \alpha^t \max_{j \in \mathbb{Z}_+} \sum_{k=x}^{y-1} P_k^{(\widetilde{p})}(X_t = j).$$

Before turning to the proof, we note the following

**Corollary 3.4.** *Suppose that $j^* \in \mathbb{Z}_+$ maximizes $\pi^{(\widetilde{p})}(\cdot)$. Then*

$$\pi^{(\widetilde{p})}(j^*) \leq \liminf_{t \to \infty} \frac{d_t(x, y)}{|y - x|\alpha^t} \leq \limsup_{t \to \infty} \frac{d_t(x, y)}{|y - x|\alpha^t} \leq 1.$$

In particular, $\lim_{t \to \infty} \frac{1}{t} \log d_t(x, y) = \log \alpha$, thus the $L_\infty$ spectral gap of the Markov chain $X$ is $1 - \alpha = c(1 - p)$, see [18]. The upper bound is Proposition 3.1. As for the lower bound, the ergodicity of the chain $\mathfrak{p}^{\widetilde{p},c}$ shows that for every $j \in \mathbb{Z}_+$, each summand in Theorem 3.3 $P_k^{(\widetilde{p})}(X_t = j) \sim \pi^{(\widetilde{p})}(j)$ as $t \to \infty$.

We prove Theorem 3.3 through two lemmas.

**Lemma 3.5.**   *For all $x, t \in \mathbb{Z}_+$,*

1. $P_{x,x+1}(\xi > t) = \alpha^t.$

2. $P_{x,x+1}(X_t \in \cdot | \xi > t) = P_x^{(\widetilde{p})}(X_t \in \cdot).$

*Proof of Lemma 3.5.* The first claim follows immediately from (8) with $y = x + 1$. We turn to the second claim. Conditioned on $N_t$, $\xi$ and $X_t$ are independent. Therefore,

$$
\begin{aligned}
P_{x,x+1}(X_t = j, \xi > t | N_t = k) &= P_{x,x+1}(X_t = j | N_t = k) P_{x,x+1}(\xi > t | N_t = k) \\
&= P_x(X_t = j | N_t = k)(1 - c)^k.
\end{aligned}
$$

Since $N_t \overset{d}{=} \mathrm{Bin}(t, 1 - p)$,

$$
\begin{aligned}
(1 - c)^k P(N_t = k) &= \binom{t}{k}\big((1 - p)(1 - c)\big)^k p^{t-k} \\
&= \binom{t}{k}(\alpha - p)^k p^{t-k} \\
&= \alpha^t \binom{t}{k}(1 - \widetilde{p})^k \widetilde{p}^{t-k}.
\end{aligned}
$$

This gives

$$
P_{x,x+1}(X_t = j, \xi > t) = \alpha^t \sum_{k=0}^{\infty} P_x(X_t = j | N_t = k) P^{(\widetilde{p})}(N_t = k).
$$

The distribution of $X_t$ conditioned on $N_t$ does not depend on the parameter $p$, and from this we obtain

$$
P_{x,x+1}(X_t = j, \xi > t) = \alpha^t P_x^{(\widetilde{p})}(X_t = j),
$$

and the result follows. $\qquad \square$

**Lemma 3.6.** *For $j \in \mathbb{Z}_+$, let $A_j = \{0, \dots, j\}$. Then*

$$
P_x(X_t \in A_j) - P_{x+1}(X_t \in A_j) = \alpha^t P_x^{(\widetilde{p})}(X_t = j).
$$

*Proof of Lemma 3.6.* Clearly,

$$
\begin{aligned}
P_x(X_t \in A_j) - P_{x+1}(X_t \in A_j) &= E_{x,x+1}\big[\mathbf{1}_{A_j}(X_t) - \mathbf{1}_{A_j}(X_t'), \xi > t\big] \\
&= \sum_{k=0}^{j-1} E_{x,x+1}[\mathbf{1}_{\{k\}}(X_t) - \mathbf{1}_{\{k+1\}}(X_t'), \xi > t] \\
&\quad + P_{x,x+1}(X_t = j, \xi > t) - P_{x,x+1}(X_t' = 0, \xi > t).
\end{aligned}
$$

Since for $t < \xi$ we have $X_t' = X_t + 1$, it follows that the expectations under the summation sign are all zero, and that the last summand is also zero. Therefore,

$$
P_x(X_t \in A_j) - P_{x+1}(X_t \in A_j) = P_{x,x+1}(X_t = j, \xi > t) = \alpha^t P_x^{(\widetilde{p})}(X_t = j),
$$

where the equality follows from Lemma 3.5. The proof of the lemma is complete. $\qquad \square$

*Proof of Theorem 3.3.* Let $A_j$ be as in the proof of Lemma 3.6. Then

$$
d_t(x, y) \geq P_x(X_t \in A_j) - P_y(X_t \in A_j) = \sum_{k=x}^{y-1}\big(P_k(X_t \in A_j) - P_{k+1}(X_t \in A_j)\big),
$$

and the theorem follows by virtue of Lemma 3.6. $\qquad \square$

9

We conclude this section with the following generalization of Lemma 3.5.

**Theorem 3.7.** $P_{x,y}(X_t \in \cdot | \xi > t)$ *converges in distribution to* $\pi^{(\tilde{p})}$ *as* $t \to \infty$.

*Proof.* First,

$$P_{x,y}(\xi > t) = \sum_{k=1}^{y-x} P_{x,y}(H_t = k).$$

Let $M = y - x$ and $\theta = 1 - c$. Then from (8), we have

$$P_{x,y}(H_t = k) = \binom{M}{k} E[(1 - \theta^{N_t})^{M-k} \theta^{N_t k}].$$

For $\rho < 1$, $E[\rho^{N_t}] = (\rho(1-p) + p)^t$, and it follows from the binomial formula that

$$P_{x,y}(H_t = k) = \binom{M}{k} (\theta^k(1-p) + p)^t (1 + o(1)) \text{ as } t \to \infty.$$

As a result,
$$P_{x,y}(\xi > t) = P_{x,y}(H_t = 1)(1 + o(1)) = (y - x)\alpha^t(1 + o(1)).$$

Next, repeating the argument in the proof of Lemma 3.5 we obtain

$$P_{x,y}(X_t = j, H_t = 1) = \sum_{k=0}^{\infty} P(X_t = j | N_t = k) P_{x,y}(H_t = 1 | N_t = k) P(N_t = k)$$

$$= \sum_{k=0}^{\infty} P_x^{(\tilde{p})}(X_t = j | N_t = k)(y - x)(1 - (1-c)^k)^{y-x-1}(1-c)^k P(N_t = k)$$

$$= \alpha^t(y - x) E_x^{(\tilde{p})}[\mathbf{1}_{\{j\}}(X_t)(1 - (1-c)^{N_t})^{y-x-1}]$$

$$= \alpha^t(y - x) P_x^{(\tilde{p})}(X_t = j)(1 + o(1))$$

where the last line follows from the binomial theorem and the bounded convergence theorem. Thus,

$$P_{x,y}(X_t = j | H_t = 1) = P_x^{(\tilde{p})}(X_t = j) + o(1).$$

Putting it all together,

$$P_{x,y}(X_t = j | \xi > t) = \frac{\sum_{k=1}^{y-x} P_{x,y}(X_t = j, H_t = k)}{P_{x,y}(\xi > t)}$$

$$= \frac{P_{x,y}(X_t = j | H_t = 1) P_{x,y}(H_t = 1) + P_{x,y}(H_t = 1)o(1)}{P_{x,y}(H_t = 1)(1 + o(1))}$$

$$= P_{x,y}(X_t = j | H_t = 1) + o(1)$$

$$= P_x^{(\tilde{p})}(X_t = j) + o(1)$$

$$= \pi^{(\tilde{p})}(j) + o(1),$$

where the second line follows from the fact that for $k \geq 2$, $P_{x,y}(H_t = k) = o(P_{x,y}(H_t = 1))$. The proof of the theorem is complete. $\quad\square$

# 4 Poisson limit and a cutoff phenomenon

In this section we let $p$ and $c$ tend to 0. We will work under the following assumption

**Assumption 4.1.** *For $n \in \mathbb{N}$ $p_n, c_n \in (0, 1)$ with $p_n \to 0$ and*

$$\lim_{n \to \infty} \frac{p_n}{c_n} = \beta \in (0, \infty).$$

We will use the superscript $(n)$ to denote the dependence of the total variation distance, probability, expectation, and stationary distribution of the parameters, e.g. the stationary distribution for the process with parameters $p_n$ and $c_n$ will be denoted by $\pi^{(n)}$.

**Theorem 4.2.** *Assume 4.1. Then $\pi^{(n)}$ converges in distribution to $Pois(\beta)$ as $n \to \infty$.*

The proof is a routine calculation of moment generating functions, and the proof appears at the end of the section. We note that the actual form of the limit distribution is irrelevant for our next and main result of this section, the cutoff phenomenon, although we do rely on the tightness of $(\pi^{(n)} : n \in \mathbb{N})$ to prove the second claim below.

**Theorem 4.3.** *Let Assumption 4.1 hold. Let $(y_n : n \in \mathbb{Z}_+)$ be a sequence of a real numbers satisfying $\lim_{n \to \infty} y_n = \infty$. Set*

$$t_n = \frac{\ln y_n}{c_n}.$$

*Then, for every $\epsilon > 0$*

*1.*

$$\lim_{\epsilon \to 0} \limsup_{n \to \infty} \sup_{t > t_n + \frac{1}{\epsilon c_n}} d_t^{(n)}(y_n, \pi^{(n)}) = 0.$$

*2.*

$$\lim_{n \to \infty} \inf_{t < t_n - b_n} d_t^{(n)}(y_n, \pi^{(n)}) = 1,$$

*where*

$$b_n = (1 + \epsilon)\Big(\frac{1}{2} \ln y_n + \frac{\ln \ln y_n}{c_n}\Big).$$

Therefore with a choice of parameters as in Theorem 4.3, the model exhibits a cutoff at $t_n$ with window size $O(\max(\ln y_n, \frac{\ln \ln y_n}{c_n}))$, see [21, p. 248].

To prove the theorem we will use the following lemma.

**Lemma 4.4.** *Assume the conditions for Theorem 4.3 hold. For $\theta > 0$, let*

$$\lambda_n(\theta) := \frac{\ln y_n + \theta}{c_n} \quad \text{and} \quad \nu_n(\theta) := \frac{\ln y_n - \ln \ln y_n - \ln \frac{p_n}{c_n} - \frac{\theta}{(\ln y_n)^{1/4}}}{- \ln(1 - c_n)}, \qquad n \in \mathbb{N}.$$

*Then*

*1. $\lim_{\theta \to \infty} \big\{ \limsup_{n \to \infty} \sup_{t > \lambda_n(\theta)} d_t^{(n)}(0, y_n) \big\} = 0$.*

2. $\lim_{\theta \to \infty} \left\{ \liminf_{n \to \infty} \inf_{t < \nu_n} d_t^{(n)}(0, y_n) \right\} = 1$.

*Proof.* Let $\alpha_n = 1 - c_n(1 - p_n)$. Recall that

1. By Proposition 3.1, for any $t > \lambda_n(\theta)$,

$$\ln d_t^{(n)}(0, y_n) \leq \ln y_n + t \ln \alpha_n \leq \ln y_n + \lambda_n(\theta) \ln \alpha_n$$

$$= \ln y_n + \left( \ln y_n + \theta \right) \frac{\ln(1 - c_n(1 - p_n))}{c_n(1 - p_n)} (1 - p_n)$$

$$\leq \ln y_n - (\ln y_n + \theta) = -\theta,$$

from which the first assertion of the lemma follows.

2. We will use the following Chernoff-Hoeffding bounds for a binomial distribution [12]. If $X \overset{d}{=} \text{Bin}(m, p)$ for some $m \in \mathbb{N}$ and $p \in (0, 1)$, then for any $\delta \in (0, 1)$,

$$P(X > (1 + \delta)pm) \leq e^{-\frac{\delta^2 pm}{2}} \quad \text{and} \quad P(X < (1 - \delta)pm) \leq e^{-\frac{\delta^2 pm}{3}}. \quad (14)$$

First, observe that under $P_0^{(n)}$, $X_t$ is stochastically dominated by the number of births up to time $t$ whose distribution is $\text{Bin}(t, p_n)$. Let

$$\gamma_n(\theta) := \left( 1 + \frac{\theta}{2(\ln y_n)^{1/4}} \right) p_n \nu_n(\theta). \quad (15)$$

In what follows, in order to simplify the notation, we will simply write $\nu_n$ and $\gamma_n$ instead of, respectively, $\nu_n(\theta)$ and $\gamma_n(\theta)$.

By the Chernoff-Hoeffding inequality, for any $t \leq \nu_n$,

$$P_0^{(n)}(X_t \geq \gamma_n) \leq P(\text{Bin}(t, p_n) \geq \gamma_n) \leq P(\text{Bin}(\nu_n, p_n) \geq \gamma_n)$$

$$\leq \exp\left( -\frac{\theta^2 p_n \nu_n}{8\sqrt{\ln y_n}} \right).$$

Therefore,

$$\lim_{n \to \infty} P_0^{(n)}(X_t \geq \gamma_n) = 0. \quad (16)$$

On the other hand, under $P_{y_n}^{(n)}$, $X_t$ stochastically dominates $\text{Bin}(y_n, (1 - c_n)^{N_t})$, which in turn, dominates $\text{Bin}(y_n, (1 - c_n)^t)$. Notice that

$$y_n(1 - c_n)^{\nu_n} = \frac{p_n}{c_n} \cdot \ln y_n \cdot e^{\frac{\theta}{(\ln y_n)^{1/4}}}. \quad (17)$$

Thus, for $n$ large enough, we have

$$\frac{\gamma_n}{y_n(1 - c_n)^t} \leq \frac{\left( 1 + \frac{\theta}{2(\ln y_n)^{1/4}} \right) p_n \nu_n}{y_n(1 - c_n)^{\nu_n}} = \frac{\left( 1 + \frac{\theta}{2(\ln y_n)^{1/4}} \right) c_n \nu_n}{(\ln y_n) e^{\frac{\theta}{(\ln y_n)^{1/4}}}} \leq \frac{1 + \frac{\theta}{2(\ln y_n)^{1/4}}}{e^{\frac{\theta}{(\ln y_n)^{1/4}}}}$$

$$\leq \left( 1 + \frac{\theta}{2(\ln y_n)^{1/4}} \right) \cdot \left( 1 - \frac{\theta}{2(\ln y_n)^{1/4}} \right) = 1 - \frac{\theta^2}{4\sqrt{\ln y_n}},$$

12

where at the last but one step we used the inequality $e^{-x} \leq 1 - \frac{x}{2}$, which is true for any sufficiently small $x > 0$, with $x = \frac{\theta}{(\ln y_n)^{1/4}}$.

Therefore, by the Chernoff-Hoeffding inequality, for any $t \leq \nu_n$,

$$
\begin{aligned}
P_{y_n}^{(n)}\big(X_t \leq \gamma_n\big) &\leq P_{y_n}^{(n)}\big[\mathrm{Bin}\big(y_n, (1 - c_n)^t\big) \leq \gamma_n\big] \\
&\leq P_{y_n}^{(n)}\big[\mathrm{Bin}\big(y_n, (1 - c_n)^{\nu_n}\big) \leq \gamma_n\big] \\
&\leq \exp\Big(-\frac{y_n(1 - c_n)^{\nu_n}\theta^4}{48 \ln y_n}\Big).
\end{aligned}
$$

Hence,

$$
\sup_{t < \nu_n} P_{y_n}^{(n)}\big(X_t \leq \gamma_n\big) \leq \exp\Big(-\frac{y_n(1 - c_n)^{\nu_n}\theta^4}{48 \ln y_n}\Big).
$$

It follows from (17) that

$$
\limsup_{n \to \infty} \sup_{t \leq \nu_n} P_{y_n}^{(n)}\big(X_t \leq \gamma_n\big) \leq e^{-\frac{\beta\theta^4}{48}}.
$$

Taking in account (16) this implies

$$
\liminf_{n \to \infty} \inf_{t < \nu_n} d_t^{(n)}(0, y_n) \geq 1 - e^{-\frac{\beta\theta^4}{48}},
$$

from which the second claim of the lemma follows.

$\square$

In order to obtain easier expressions to work with, we observe that for $\theta$ large enough, independently of $n$, we have

$$
\nu_n(\theta) \geq \frac{\ln y_n - \ln \ln y_n - \theta}{-\ln(1 - c_n)} = \frac{\ln y_n - \ln \ln y_n - \theta}{c_n} \times \underbrace{\frac{1}{1 + c_n/2 + c_n^2/3 + \ldots}}_{(*)}.
$$

Since $(*) = 1 - \frac{c_n}{2} + O(c_n^2)$ and we have that

$$
\begin{aligned}
\nu_n(\theta) &= t_n - \frac{1}{2}\ln y_n + O(c_n)\ln y_n - \frac{\ln \ln y_n}{c_n} + \frac{1}{2}\ln \ln y_n - O(c_n)\ln \ln y_n - \frac{\theta}{c_n}(1 - o(1)) \\
&= t_n - \Big(\frac{1}{2}\ln y_n + \frac{\ln \ln y_n}{c_n}\Big) + O(c_n)\ln y_n + \frac{1}{2}\ln \ln y_n - \frac{\theta}{c_n}(1 - o(1)),
\end{aligned}
$$

and so for every $\theta > 0$ and $\epsilon > 0$,

$$
\nu_n(\theta) > t_n - (1 + \epsilon)\Big(\frac{1}{2}\ln y_n + \frac{\ln \ln y_n}{c_n}\Big),
$$

provided $n$ is large enough.

This leads to the following corollary. Recall that $b_n = (1 + \epsilon)\big(\frac{1}{2}\ln y_n + \frac{\ln \ln y_n}{c_n}\big)$.

13

**Corollary 4.5.** *Under the assumptions of Theorem 4.3,*

1. $\lim\limits_{\epsilon \to 0} \limsup\limits_{n \to \infty} \sup\limits_{t \geq t_n + \frac{1}{\epsilon c_n}} d_t^{(n)}(0, y_n) = 0.$

2. *For any $\epsilon > 0$, $\lim\limits_{n \to \infty} \inf\limits_{t \leq t_n - b_n} d_t^{(n)}(0, y_n) = 1.$*

We are ready to prove Theorem 4.3.

*Proof of Theorem 4.3.* We begin with the first claim. Recall that $\alpha_n = 1 - c_n(1 - p_n)$. Then from the triangle inequality and the Corollary 3.2 we obtain

$$d_t^{(n)}(y_n, \pi^{(n)}) \leq d_t^{(n)}(y_n, 0) + d_t^{(n)}(0, \pi^{(n)})$$
$$\leq d_t^{(n)}(y_n, 0) + \mu_n \alpha_n^t,$$

where $\mu_n = \sum y\pi^{(n)}(y) = \frac{p_n}{c_n(1-p_n)}$. Now $\mu_n \to \beta$, and $\ln(\alpha_n^t) = t \ln(1 - c_n(1 - p_n)) \leq -\frac{1}{2} c_n t$ provided $p_n \leq \frac{1}{2}$. Therefore

$$\lim_{n \to \infty} \sup_{t > t_n + \frac{1}{\epsilon c_n}} \mu_n \alpha_n^t = 0.$$

The result now follows from this, combined with the first claim in Corollary 4.5.

We turn to the second claim. Fix $\theta > 0$, and recall $\nu_n(\theta)$ from Lemma 4.4. From the proof of Lemma 4.4, it follows that for all $t < \nu_n(\theta)$, $\limsup_{n \to \infty} \sup_{t \leq \nu_n(\theta)} P_{y_n}(X_t < \gamma_n(\theta)) < e^{-\beta\theta^4/48}$, where $\gamma_n = \gamma_n(\theta)$ was defined in (15). Since $t_n - b_n < \nu_n(\theta)$ provided $n$ is large enough, it follows that

$$\lim_{n \to \infty} \sup_{t \leq t_n - b_n} P_{y_n}(X_t \leq \gamma_n) = 0. \tag{18}$$

By definition, $\gamma_n \geq p_n \nu_n(\theta) \to \infty$ as $n \to \infty$, and since

$$d_t(y_n, \pi^{(n)}) \geq P_{y_n}^{(n)}(X_t > \gamma_n) - \pi^{(n)}\left(\{\gamma_n, \gamma_n + 1, \dots\}\right),$$

the tightness of $(\pi^{(n)} : n \in \mathbb{N})$ along with (18) give

$$\lim_{n \to \infty} \sup_{t \leq t_n - b_n} d_t(y_n, \pi^{(n)}) = 1,$$

completing the proof. $\qquad\square$

We conclude this section with the proof of Theorem 4.2

*Proof of Theorem 4.2.* Let $Z_n$ be a random variable distributed according to $\pi^{(n)}$. By Proposition 2.2 we can write

$$Z_n = \sum_{j=0}^{\infty} B_j(G_j),$$

where $(G_j : j \in \mathbb{Z}_+)$ are IID Geom$^-(p_n)$, and $(B_j(k) : j, k \in \mathbb{Z}_+)$ are independent with $B_j(k) \overset{d}{=} \mathrm{Bin}(k, (1 - c_n)^j)$, all independent of the $G_j$'s.

Let $\Lambda(t) = \ln E[e^{-tZ_n}]$. Then

$$\Lambda(t) = \sum_{j=0}^{\infty} \ln E[e^{-tB_j(G_j)}].$$

Now

$$\begin{aligned}
E[e^{-tB_j(G_j)}|G_j] &= \left(e^{-t}(1-c_n)^j + (1-(1-c_n)^j)\right)^{G_j} \\
&= \left(1 - q_n^j(1-e^{-t})\right)^{G_j} \\
&= e^{-\gamma_{n,j}(t)G_j},
\end{aligned}$$

where $q_n = 1 - c_n$ and $\gamma_{n,j}(t) = -\ln(1 - q_n^j(1-e^{-t}))$. Therefore

$$E[e^{-tB_j(G_j)}] = E[e^{-\gamma_{n,j}(t)G_j}] = \sum_{k=0}^{\infty} (1-p_n)p_n^k e^{-\gamma_{n,j}(t)k} = \frac{1-p_n}{1-p_n e^{-\gamma_{n,j}(t)}}.$$

Thus,

$$\begin{aligned}
\Lambda(t) &= -\sum_{j=0}^{\infty} \ln \frac{1 - p_n(1-(1-e^{-t})q_n^j)}{1-p_n} \\
&= -\sum_{j=0}^{\infty} \ln\left(1 + \frac{p_n q_n^j}{1-p_n}(1-e^{-t})\right).
\end{aligned}$$

For $x \in (0,1)$,

$$0 \le x - \ln(1+x) \le \frac{x^2}{2}$$

Therefore,

$$0 \le \underbrace{\sum_{j=0}^{\infty} \frac{p_n q_n^j}{1-p_n}(1-e^{-t})}_{(I)} + \Lambda(t) \le \underbrace{\frac{p_n^2(1-e^{-t})^2}{2(1-p_n)^2} \sum_{j=0}^{\infty} q_n^{2j}}_{(II)}. \tag{19}$$

Next,

$$(I) = \frac{p_n}{(1-p_n)c_n}(1-e^{-t}) \underset{n\to\infty}{\to} \beta(1-e^{-t}),$$

and since $\sum_{j=0}^{\infty} q_n^{2j} \le \sum_{j=0}^{\infty} q_n^j = \frac{1}{c_n}$,

$$(II) \le p_n \frac{p_n}{(1-p_n)^2 c_n} = p_n \beta O(1) \underset{n\to\infty}{\to} 0.$$

We have thus proved that $\lim_{n\to\infty} \Lambda(t) = -\beta(1-e^{-t})$. $\qquad\square$

15

# 5   Additional Topics

## 5.1   Branching process representation

We adopt a scheme of Key [17] for general branching processes with immigration in random environment to give a probabilistic interpretation of the particular instance of Neuts' formula [25]. Using the approach of [1] we compute the generating function of the extinction time in Section 5.2.3.

The process $X$ can be thought of as a branching process with immigration in random environment. Branching processes have been used to model growth of a population subject to random catastrophes by many authors (see, for instance, a comprehensive literature review in [16]), the idea goes back to at least [15] where a branching process in random environment (without immigration) was considered. In this section we use a branching representation of our process and Key's [17] representation of its stationary distribution for several purposes. First, it yields Lemma 5.1 below stating that the extinction time $\tau$ has exponential tails, next it provides an illuminating probabilistic representation of the invariant distribution $\pi$ for our process, including the extreme case of rare but nearly total catastrophes (see the discussion after Proposition 5.2 and Theorem 5.3 below).

Let

$$\omega_t = \begin{cases} 1 & \text{if} \quad \text{a birth event occurs at time } t \\ 0 & \text{if} \quad \text{a catastrophe occurs at time } t \end{cases} \tag{20}$$

We refer to the sequence $\omega := (\omega_t)_{t \in \mathbb{Z}_+}$ as a *random environment.* We denote the distribution of the environment by $\mathbb{P}$, the law of the process conditional on the environment by $P_\omega$, and the corresponding expectation by $E_\omega$.

The Markov process $X$ can be described using the following branching equation:

$$X_{t+1} = \sum_{k=1}^{X_t+I_t} U_{t,i} = \sum_{k=1}^{X_t} U_{t,i} + I_t, \tag{21}$$

where $I_t = \omega_t$ is interpreted as the number of immigrants joining the system at generation $t$ and $U_{t,i}$ as the number of progeny of the $i$-th particle living at generation $t$. Under the probability law conditional on the environment $\omega_t$, $U_{t,i}$ are independent Bernoulli variables with parameter $c_t := \omega_t + (1 - \omega_t)(1 - c)$ which are independent of $X_t$ :

$$P_\omega(U_{t,i} = 1) = c_t \qquad \text{and} \qquad P_w(U_{t,i} = 0) = 1 - c_t.$$

In statistical applications, this special type of branching processes with Bernoulli reproduction mechanism is often referred to as a RCINAR(1) *random coefficient integer-valued autoregressive process of order one* [32]. In this context, (21) is written as

$$X_{t+1} = (1 - c_t) * X_t + I_t, \qquad t \in \mathbb{Z}_+,$$

where $(1 - c_t)*$ describes the action of a *binomial thinning operator* [24, 28, 31].

Stationary distribution of branching processes with immigration in a random environment, in a general (and, in fact, multi-type) setting, was studied in [17]. In particular, it

follows from results in [17] that the random variable $\tau$ has an exponential distribution tail (in order to deduce this, one may replace $I_t$ by 1 in (21) to be able to formally use Theorem 4.2 in [17], and then apply a stochastic dominance argument). We state it formally as

**Lemma 5.1.** *There exists $a, b > 0$ such that $P_0(\tau > t) \leq ae^{-bt}$ for any $t \geq 0$.*

We next consider a branching process obtained from $X$ by sampling at the times when catastrophes occur. This auxiliary process has a slightly simpler structure than the underlying process $X$. We use it below to obtain an alternative probabilistic representation of the stationary distribution of $X$.

Let $T_0 = 0$ and

$$T_n = \inf\{k > T_{n-1} : \omega_k = 0\}. \tag{22}$$

Observe that the sequence $(T_n - T_{n-1} : n \geq 1)$ is an IID sequence of $\mathrm{Geom}(1-p)$ random variables. Let $Z_n = X_{T_n}$ and $Z := (Z_n)_{n \in \mathbb{Z}_+}$.

**Proposition 5.2.** *The Markov chain $Z$ has a unique stationary distribution $Z_\infty$, whose generating function is given by*

$$E[s^{Z_\infty}] = \prod_{k=1}^{\infty} \frac{1-p}{1 - p\big(s(1-c)^k + 1 - (1-c)^k\big)}, \qquad s \in [0,1].$$

*Thus, in the language of Proposition 2.2, $Z_\infty = R - R_0 = \sum_{j=1}^{\infty} Bin_j\big(R_n, (1-c)^j\big)$.*

*Proof.* Considering $R_t$ as an immigration process, $Z_t$ can be constructed as a branching process with immigration governed by the following branching identity:

$$Z_{t+1} = \sum_{k=1}^{Z_t + R_t} V_{t,k}, \qquad t \in \mathbb{Z}_+, \tag{23}$$

where $V_{t,k}$ are IID Bernoulli random variables, independent of the immigration process and $Z_0$, such that

$$P(V_{t,k} = 1) = 1 - c \qquad \text{and} \qquad P(V_{t,k} = 0) = c.$$

The result thus follows from Theorem 4.2 in [17]. $\qquad \square$

We remark that an auxiliary process similar to our $(Z_n)_{n \in \mathbb{Z}_+}$ has been used, for instance, in [7, 15] to derive the stationary distribution for different models with catastrophes.

Following the representation of the stationary distribution in [17], one can write

$$Z_\infty = \lim_{t \to \infty} \sum_{k=-t}^{-1} Z_{k,0} = \sum_{k=-\infty}^{-1} Z_{k,0}, \tag{24}$$

where $Z_{k,0} \overset{d}{=} Bin_{|k|}\big(R_{|k|}, (1-c)^{|k|}\big)$ is the number of descendant alive at time zero of a "demo" immigrant that arrived at time $k < 0$. Heuristically, in this representation $Z_\infty$ is the

17

population at time zero of a branching process that starts at minus infinity [17]. In between two regeneration times $T_n$, the process goes up $\mathrm{Geom}^-(1-p)$ number of times. When one observe the original chain in the stationary regime, time-wise the chain is in a random place between two random times $T_n$. This suggests (using the key renewal theorem) that the stationary distribution of the original Markov chain should be the convolution of $Z_\infty$ and an independent $\mathrm{Geom}^-(1-p)$ variable. The result is formally confirmed in Proposition 5.2.

We conclude this section with a brief discussion of the case of "severe but rare" catastrophes. For a biological motivation of this regime see, for instance, [13, 19, 26, 27, 30]. Specifically, a sequence of parameters $(p_n, c_n)$ such that $p_n \to 1$, $c_n \to 1$ as $n \to \infty$, and $\lim_{n \to \infty} \frac{1-c_n}{1-p_n} = \beta$ for some $\beta$. We will denote the stationary distribution for the $n$-th model, given by Proposition 2.2, by $R^{(n)}$. Observe that

$$E[s^R] = \prod_{k=0}^{\infty} \frac{1-p}{1 - p\big(s(1-c)^k + 1 - (1-c)^k\big)}. \tag{25}$$

With this, it is not hard to verify the following result:

**Theorem 5.3.** $R^{(n)} = R_0 + A_n$, where $A_n$ is independent of $R_0$ and converges in distribution, as $n \to \infty$, to $Poiss(\beta)$.

*Proof.* Recall (25), and set $x_n(k) := \frac{p_n(1-c_n)^k}{1-p_n}$, $k \in \mathbb{Z}_+$, $n \in \mathbb{N}$, so that

$$\ln E\big[s^{R^{(n)}}\big] = -\sum_{k=0}^{\infty} \ln\big(1 + x_n(k)(1-s)\big), \qquad s \in [0,1].$$

To estimate the right-hand side, one can apply to $x_n(k)$ the inequality $x - \frac{x^2}{2} \le \ln(1+x) \le x$ which is true for all $x > 0$ sufficiently small (and hence, uniformly on $k$, for all $x_n(k)$ with $n$ large enough). The result follows from the fact

$$\sum_{k=1}^{\infty} x_n(k) = \frac{p_n(1-c_n)}{(1-p_n)c_n} \to \beta, \qquad \text{as } n \to \infty,$$

and

$$\sum_{k=1}^{\infty} \big(x_n(k)\big)^2 \le x_n(1) \cdot \sum_{k=1}^{\infty} x_n(k) \to 0 \cdot \beta = 0, \qquad \text{as } n \to \infty.$$

where we took in account that $x_n(k)$ is monotone decreasing in $k$. Thus $\ln E\big[s^{R^{(n)}}\big]$ converges, as $n \to \infty$, to $-\beta(1-s)$, and the proof of the theorem is complete. $\qquad \square$

Note that in view of Proposition 5.2, $Poiss(\beta)$ is the limit in distribution of $Z_\infty$. Furthermore, using (24) and a similar representation for the underlying branching process $X$, one can by virtue of the renewal theorem interpret $-R_0$ as the time of the last catastrophe before time zero and $A_n$ as the distribution of the population right after the last catastrophe in the stationary branching process $(X_t)_{t \in \mathbb{Z}}$.

## 5.2 First Extinction Time

### 5.2.1 Overview

In this section we discuss the following two aspects related to the first extinction time $\tau$ :

- Asymptotic behavior of $\tau$ under large initial population.

- Generating function for $\tau$.

### 5.2.2 Asymptotic for large population

In this section we discuss the asymptotic behavior of the first extinction time $\tau$ when the process starts from a large population. To do that we will use the coupling construction of Section 3.1. Consider the processes $X_t^{(0)}$ and $X_t^{(n)}$ with initial populations 0 and $n$, respectively. From our coupling we know that for every $t \geq 0$ we have

$$X_t^{(n)} = X_t^{(0)} + H_t^{(n)}.$$

Let $\tau^{(n)}$ and $\xi^{(n)}$ be the hitting time of 0 by $X^{(n)}$ and $H^{(n)}$, respectively:

$$\tau^{(n)} = \inf\{t \geq 0 : X_t^{(n)} = 0\}, \ \xi^{(n)} = \inf\{t \geq 0 : H_t^{(n)} = 0\}.$$

Then $\tau^{(n)}, \xi^{(n)}$ are both nondecreasing.

Let $T_0 = 0$ and let $T_1, T_2, \dots$ be the increasing sequence of times $X^{(0)}$ visits 0. Then clearly,

$$\tau^{(n)} = \inf\{T_k : T_k \geq \xi^{(n)}\}.$$

This is because $X_t^{(n)} = 0$ if and only if $H_t^{(n)} = 0$ and $X_t^{(0)} = 0$. Now let $\rho^{(n)} = \tau^{(n)} - \xi^{(n)}$. Then $\rho^{(n)}$ depends on the past of the coupled system only through the size of the population $X_{\xi^{(n)}}^{(n)}$. Thus its distribution coincides with the distribution of $\tau^{(X_{\xi^{(n)}}^0)}$. By ergodicity of $X^{(0)}$, and the fact that $\xi^{(n)} \nearrow \infty$ a.s. as $n \to \infty$, it follows that the distribution of $\rho^{(n)}$ converges to the distribution of $\tau$, the hitting time of 0 under $\pi$. We have proved the following:

**Proposition 5.4.** $\tau^{(n)} - \xi^{(n)}$ *converges in distribution to* $P_\pi(\tau \in \cdot)$ *as* $n \to \infty$.

It follows from (9) that

$$P(\xi^{(n)} \leq t) = E[(1 - (1 - c)^{N_t})^n].$$

Let $\epsilon \in (0, 1/2)$, and let $A_t = \{|N_t/t - (1-p)| < \epsilon\}$. Then by the Law of Large Numbers $P(A_t) \to 1$. We have the following two-sided bounds:

$$E[(1 - (1 - c)^{(1-\epsilon)(1-p)t})^n, A_t] \leq E[(1 - (1 - c)^{N_t})^n] \tag{26}$$
$$\leq E[(1 - (1 - c)^{t(1+\epsilon)(1-p)})^n] + P(A_t^c).$$

Let

$$d_n = -\frac{\ln n}{(1 - p) \ln(1 - c)}.$$

19

If $t \leq (1 - \epsilon)d_n$, then it follows from the second inequality in (26) that

$$P(\xi^{(n)} \leq t) \leq \left(1 - n^{-(1-\epsilon^2)}\right)^n \left(1 + o(1)\right) + o(1) \to 0,$$

while if $t \geq (1 + 2\epsilon)d_n$, it follows from the first inequality in (26) that

$$P(\xi^{(n)} \leq t) \geq \left(1 - n^{-(1+2\epsilon)(1-\epsilon)}\right)^n \left(1 + o(1)\right) \to 1.$$

Thus $\xi^{(n)}/d_n \to 1$ in probability. This, and Proposition 5.4 give

**Proposition 5.5.** $\tau^{(n)}/d_n \to 1$ *in probability as $n \to \infty$.*

### 5.2.3   Generating function

For $s \in [0, 1]$, let $a_n(s) = E_n[s^\tau]$ and $\psi(s, z) = \sum_{n=1}^\infty a_n z^n$. Note that $a_0 = 1$. The process has the following first-step decomposition:

$$X_{t+1} = \mathbf{I}_{\{\omega_t = 1\}} \cdot (X_t + 1) + \mathbf{I}_{\{\omega_t = 0\}} \cdot \text{Bin}(X_t, 1 - c), \tag{27}$$

where $\mathbf{I}_A$ stands for the indicator of the event $A$, namely $\mathbf{I}_A(\omega) = 1$ if $\omega \in A$ and $\mathbf{I}_A(\omega) = 0$ if $\omega \notin A$., and $\omega_t$ is defined in (20). The generating function $\psi(s, z)$ can be evaluated using (27) and an analytical method of [1]. In particular, we have

**Theorem 5.6.** *For $s \in [0, 1]$, let $\eta_0(s) = 1$ and*

$$\eta_n(s) = (-1)^n (1 - c)^{\frac{n(n-1)}{2}} \left(\frac{(1 - p)s}{1 - ps}\right)^n \prod_{k=1}^n \frac{1}{1 - (1 - c)^k}, \qquad n \in \mathbb{N}.$$

*Then*

$$E_1[s^\tau] = 1 + \frac{1 - s}{ps} - \frac{\sum_{n=1}^\infty \eta_n(s)}{\sum_{n=1}^\infty \eta_n(s)\frac{ps(1-c)^n}{1-ps+ps(1-c)^n}}. \tag{28}$$

The proof of the theorem is similar to the proof of Theorem 3.1, part (ii), in [1]. Namely, an application of (27) leads to a recursive equation for the generating function $\psi(s, z)$ of a type that has been analyzed in [1]. We comment that through the recurrence relation (29), we can obtain an explicit formula for $E_n[s^\tau]$ for each $n \in \mathbb{N}$. The proof below is provided for the sake of completeness.

*Proof.* We assume throughout the argument that $s, z \in (0, 1)$. For simplicity of notation, we will occasionally suppress the dependence of underlying functions on the parameter $s$. Using (27), we obtain

$$a_n = ps a_{n+1} + (1 - p)s \sum_{k=0}^n \binom{n}{k} c^{n-k}(1 - c)^k a_k, \qquad n \in \mathbb{N}.$$

Multiplying by $z^n$ and summing over $n$ from 1 to $\infty$ yields

$$
\psi(z) - 1 = \frac{ps}{z}\{\psi(z) - 1 - a_1 z\} - (1-p)s
$$

$$
+ (1-p)s \sum_{k=0}^{\infty} z^k (1-c)^k a_k \sum_{n=k}^{\infty} \binom{n}{k} (cz)^{n-k}
$$

$$
= \frac{ps}{z}\{\psi(z) - 1 - a_1 z\} - (1-p)s + \frac{(1-p)s}{1-cz} \sum_{k=0}^{\infty} \left(\frac{z-cz}{1-cz}\right)^k a_k
$$

$$
= \frac{ps}{z}\{\psi(z) - 1 - a_1 z\} - (1-p)s + \frac{(1-p)s}{1-cz} \psi\left(\frac{z-cz}{1-cz}\right),
$$

where we used the negative binomial formula $\sum_{n=k}^{\infty} \binom{n}{k} x^{n-k} = (1-x)^{-k-1}$ with $x = cz$. Thus

$$
\psi(s,z) = \frac{ps}{ps-z} - \frac{z}{ps-z}\big(1 - psa_1(s) - (1-p)s\big)
$$

$$
- \frac{(1-p)sz}{(1-cz)(ps-z)} \psi\left(s, \frac{z-cz}{1-cz}\right)
$$

$$
= 1 + \frac{z}{ps-z}\big(psa_1(s) + (1-p)s\big)
$$

$$
- \frac{(1-p)sz}{(1-cz)(ps-z)} \psi\left(s, \frac{z-cz}{1-cz}\right). \tag{29}
$$

Let

$$
g(s,z) := ps - z + z\big(psa_1(s) + (1-p)s\big) \tag{30}
$$

and

$$
\varphi(z) = \psi(s,z)(ps-z).
$$

For $k \geq 1$, let $h(z) = \frac{z-cz}{1-cz}$, $h_0(z) = z$, and $h_k(z) = h\big(h_{k-1}(z)\big)$ for $k \in \mathbb{N}$. It is easy to verify that

$$
h_k(z) = \frac{z(1-c)^k}{1 - \big(1 - (1-c)^k\big)z}. \tag{31}
$$

In this notation, (29) can be rewritten as

$$
\varphi(s,z) = g(s,z) + \frac{(1-p)sh(z)}{(1-c)(h(z)-ps)} \varphi\big(s, h(z)\big). \tag{32}
$$

Note that $h_k(z) \in (0,z)$ for $z \in (0,1)$. Consequently, taking in account (31) and that $a_n(s) \in (0,1)$ for all $s \in (0,1)$,

(i) For any $z \in (0,1)$, $h_k(z)$ decreases, as $k \to \infty$, to zero, which is the smallest of two fixed points of $h$.

(ii) $\psi\big(s, h_k(z)\big) \leq \psi(s,z) \leq \sum_{n=0}^{\infty} z^n < \infty$ for all $k \in \mathbb{Z}_+$.

21

(iii) We have:

$$-1 < -z(1-s) < ps - z + z(1-p)s \leq g(s,z) \leq ps - z + z\big(ps + (1-p)s\big) < 1,$$

and hence $g(s,z)$ is uniformly bounded for $s, z \in (0,1)$.

(iv) For $z \leq ps$ and $k \in \mathbb{Z}_+$,

$$0 \leq \frac{(1-p)sh_k(z)}{(1-c)\big(ps - h_k(z)\big)} \to 0, \qquad \text{as } k \to \infty. \tag{33}$$

Thus, one can iterate (32) to obtain

$$\varphi(s,z) = g(s,z) + \sum_{n=1}^{\infty} g\big(s, h_n(z)\big) \prod_{k=1}^{n} \frac{(1-p)sh_k(z)}{(1-c)\big(h_k(z) - ps\big)}.$$

Plugging in into this formula $z = ps$ yields, taking into account that $\varphi(s,z) = 0$,

$$0 = g(s,z) + \sum_{n=1}^{\infty} g\big(s, h_n(z)\big) \frac{(1-p)^n s^n (1-c)^{\frac{n(n-1)}{2}}}{(ps-1)^n} \prod_{k=1}^{n} \frac{1}{1 - (1-c)^k}. \tag{34}$$

This yields (28) by virtue of (30) and (31). In fact, after a suitable renaming of variables, equation (34) for $a_1(s)$ is analogous to (3.12) in [1], while our (28) is its solution (3.4) in [1]. $\qquad\square$

# References

[1] J. R. Artalejo, A. Economou and M. J. Lopez-Herrero (2007). Evaluating growth measures in populations subject to binomial and geometric catastrophes. Math. Biosci. Eng. 4, 573-594.

[2] P. J. Brockwell (1986). The extinction time of a general birth and death process with catastrophes. J. Appl. Probab. 23, 851-858.

[3] P. J. Brockwell, J. Gani and S. I. Resnick (1982). Birth, immigration and catastrophe processes. Adv. in Appl. Probab. 14, 709-731.

[4] K. Burdzy and W. S. Kendall (2000). Efficient Markovian couplings: examples and counterexamples. Ann. Appl. Probab. 10, 362-409.

[5] B. J. Cairns (2009). Evaluating the expected time to population extinction with semi-stochastic models. Math. Popul. Stud. 16, 199-220.

[6] B. J. Cairns and P. K. Pollett (2005). Approximating persistence in a general class of population processes. Theoret. Population Biol. 68, 77-90.

[7] A. Economou (2004). The compound Poisson immigration process subject to binomial catastrophes. J. Appl. Probab. 41, 508-523.

[8] A. Economou and D. Fakinos (2008). Alternative approaches for the transient analysis of Markov chains with catastrophes. J. Stat. Theory Pract. 2, pp.183-197.

[9] W. J. Ewens, P. J. Brockwell, J. M. Gani, and S. I. Resnick (1987). Minimum viable population size in the presence of catastrophes. Pp. 59-68 in M. E. Soule, editor. Viable Populations for Conservation. Cambridge University Press, Cambridge, UK.

[10] F. B. Hanson and H. C. Tuckwell (1978). Persistence times of populations with large random fluctuations. Theoret. Population Biol. 14, 46-61.

[11] F. B. Hanson and H. C. Tuckwell (1981). Logistic growth with random density independent disasters. Theoret. Population Biol. 19, 1-18.

[12] W. Hoeffding (1963). Probability inequalities for sums of bounded random variables. J. Amer. Stat. Assoc. 58, 1330.

[13] T. E. Huillet (2011). On a Markov chain model for population growth subject to rare catastrophic events. Physica A: Statistical Mechanics and its Applications 390, 4073-4086.

[14] V. Junior, F. Machado and A. Roldan-Correa (2016). Dispersion as a survival strategy. J. Stat. Phys. 164, 937-951.

[15] N. Kaplan, A. Sudbury and T. S. Nilsen (1975). A branching process with disasters. J. Appl. Probab. 12, 47-59

[16] S. Kapodistria, T. Phung-Duc and J. Resing (2016). Linear birth/immigration-death process with binomial catastrophes. Probab. Engrg. Inform. Sci. 30, 79-111.

[17] E. S. Key (1987). Limiting distributions and regeneration times for multitype branching processes with immigration in a random environment. Ann. Probab. 15, 344-353.

[18] I. Kontoyiannis and S. P. Meyn (2012). Geometric ergodicity and the spectral gap of non-reversible Markov chains. Probab. Theory Related Fields 154, 327-339.

[19] R. Lande (1993). Risks of population extinction from demographic and environmental stochasticity and random catastrophes. American Naturalist 142, 911-927.

[20] M. C. A. Leite, N. P. Petrov and E. Weng (2012). Stationary distributions of semistochastic processes with disturbances at random times and with random severity. Nonlinear Anal. Real World Appl. 13, 497-512.

[21] D. A. Levin, Y. Peres and E. L. Wilmer (2009). Markov chains and mixing times. American Mathematical Society, Providence, RI.

[22] S. N. Majumdar and D. Dhar (2001). Persistence in a stationary time series. Phys. Rev. E 64, 046123.

[23] M. Mangel and C. Tier (1993). Dynamics of metapopulations with demographic stochasticity and environmental catastrophes. Theoret. Population Biol. 44, 1-31.

[24] E. McKenzie (2003). Discrete variate time series. In: D. N. Shanbhag and C. R. Rao (eds), Handbook of Statistics, Elsevier Science, pp. 573-606.

[25] M. F. Neuts (1994). An interesting random walk on the non-negative integers. J. Appl. Probab. 31, 48-58.

[26] R. R. Paine (2000). If a population crashes in prehistory, and there is no paleodemographer there to hear it, does it make a sound? Am. J. Phys. Anthropol. 112, 181-190.

[27] J. F. Silva, J. Raventos, H. Caswell, and M. C. Trevisan (1991). Population responses to fire in a tropical savanna grass, andropogon semiberbis: a matrix model approach. J. Ecol. 79, 345-355.

[28] F. W. Steutel, K. van Harn (1979). Discrete Analogues of Self-Decomposability and Stability. Ann. Prob. 7, 893–899.

[29] H. Thorisson (2000). Coupling, Stationarity, and Regeneration, Springer, New-York.

[30] S. Tuljapurkar (1990). Population Dynamics in Variable Environments. Lecture Notes in Biomathematics, Vol. 85, Springer-Verlag.

[31] C. H. Weiß (2008). Thinning operations for modeling time series of counts - a survey. AStA Adv. Stat. Anal. 92, 319-341.

[32] H. Zheng, I. V. Basawa and S. Datta (2007). First-order random coefficient integer-valued autoregressive processes, J. Statist. Plann. Inference 173 (2007), 212-229.