Applied Probability Trust (23 January 2024)

BRANCHING PROCESSES IN NEARLY DEGENERATE VARYING ENVIRONMENT

PÉTER KEVEI,* *University of Szeged*
KATA KUBATOVICS,** *University of Szeged*

Abstract

We investigate branching processes in varying environment, for which $\overline{f}_n \to 1$ and $\sum_{n=1}^\infty (1-\overline{f}_n)_+ = \infty$, $\sum_{n=1}^\infty (\overline{f}_n - 1)_+ < \infty$, where $\overline{f}_n$ stands for the offspring mean in generation n. Since subcritical regimes dominate, such processes die out almost surely, therefore to obtain a nontrivial limit we consider two scenarios: conditioning on non-extinction, and adding immigration. In both cases we show that the process converges in distribution without normalization to a nondegenerate compound-Poisson limit law. The proofs rely on the shape function technique, worked out by Kersting [15].

Keywords: branching process in varying environment, Yaglom-type limit theorem, immigration, shape function

2020 Mathematics Subject Classification: Primary 60J80
Secondary 60F05

1. Introduction

A Galton–Watson branching process in varying environment (BPVE) $(X_n,\, n \geq 0)$ is defined as

$$X_0 = 1, \quad \text{and } X_n = \sum_{j=1}^{X_{n-1}} \xi_{n,j}, \quad n \in \mathbb{N} = \{1, 2, \ldots\}, \tag{1}$$

where $\{\xi_{n,j}\}_{n,j\in\mathbb{N}}$ are nonnegative independent random variables, such that for each n, $\{\xi_{n,j}\}_{j\in\mathbb{N}}$ are identically distributed, and let ξ_n denote a generic copy. We can interpret X_n as the size of the nth generation of a population and $\xi_{n,j}$ represents the number of offspring produced by the jth individual in generation $n-1$. These processes are natural extensions of usual homogeneous Galton–Watson processes, where the offspring distribution does not change with the generation.

The investigation of BPVE started in the early 1970's with the work of Church [4] and Fearn [6]. There is a recent increasing interest again in these processes, which was triggered by Kersting [15], who obtained necessary and sufficient condition for almost sure extinction for regular processes. In [15] he introduced the shape function of a generating function (g.f.), which turned out to be the appropriate tool for the

* Postal address: Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, 6720 Szeged, Hungary
** Postal address: Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, 6720 Szeged, Hungary
* Email address: kevei@math.u-szeged.hu
** Email address: kubatovics@server.math.u-szeged.hu

analysis. Moreover, in [15] Yaglom-type limit theorems (that is conditioned on non-extinction) were obtained. Under different regularity conditions Bhattacharya and Perlman [2] also proved Yaglom-type limit theorems both in the discrete and in the continuous time setting, extending the results of Jagers [14]. For multitype processes these questions were investigated by Dolgopyat et al. [5]. Cardona-Tobón and Palau [3] obtained probabilistic proofs of the results in [15] using spine decomposition. On general theory of BPVE we refer to the recent monograph by Kersting and Vatutin [16], and to the references therein.

Here we are interested in *branching processes in nearly degenerate varying environment.* Let $f_n(s) = \mathbb{E}s^{\xi_n}$, $s \in [0,1]$, denote the g.f. of the offspring distribution of the nth generation, and put $\overline{f}_n := f_n'(1) = \mathbb{E}\xi_n$ for the offspring mean. We assume the following conditions:

(C1) $\lim_{n\to\infty} \overline{f}_n = 1$, $\sum_{n=1}^\infty (1-\overline{f}_n)_+ = \infty$, $\sum_{n=1}^\infty (\overline{f}_n - 1)_+ < \infty$,

(C2) $\lim_{n\to\infty, \overline{f}_n<1} \frac{f_n''(1)}{1-\overline{f}_n} = \nu \in [0,\infty)$, and $\frac{f_n''(1)}{|1-\overline{f}_n|}$ is bounded,

(C3) if $\nu > 0$, then $\lim_{n\to\infty, \overline{f}_n<1} \frac{f_n'''(1)}{1-\overline{f}_n} = 0$, and $\frac{f_n'''(1)}{|1-\overline{f}_n|}$ is bounded,

where here, and later on $\lim_{n\to\infty, \overline{f}_n<1}$ means that the convergence holds along the subsequence $\{n : \overline{f}_n < 1\}$, and $a_+ = \max\{a, 0\}$. Condition (C1) means that the process is nearly critical and the convergence $\overline{f}_n \to 1$ is not too fast. Note that we also allow supercritical generations, that is, with $\overline{f}_n > 1$. However, subcritical regimes dominate since by (C1) $\mathbb{E}X_n = \prod_{j=1}^n \overline{f}_j \to 0$, thus the process dies out almost surely.

Condition (C2) implies that $f_n''(1) \to 0$, thus $f_n(s) \to s$, so the branching mechanism converges to the degenerate branching. Therefore, this nearly critical model does not have a natural homogeneous counterpart. To obtain nontrivial limit theorems we can condition on non-extinction, or add immigration. Note that condition (C2) implies that the offspring distributions have finite second moment, which is assumed throughout the paper.

Conditioned on $X_n > 0$, in our Theorem 1 we prove that X_n converges in distribution without normalization to a geometric distribution. Although, the process is nearly critical, in this sense its behavior is similar to a homogeneous subcritical process, where no normalization is needed to obtain a limit, see Yaglom's theorem [21] (or in Athreya and Ney [1, Theorem I.8.1, Corollary I.8.1]). For a regular BPVE (in the sense of [15]), necessary and sufficient conditions were obtained for the tightness of X_n (without normalization) conditioned on $X_n > 0$ in [15, Corollary 2]. However, to the best of our knowledge, our result is the first proper limit theorem without normalization. Similarly to Yaglom's theorem in the homogeneous critical case (see [21] or [1, Theorem I.9.2]), exponential limits for properly normalized processes conditioned on non-extinction were obtained in several papers, see e.g. [14], [2], [15].

Allowing immigration usually leads to similar behavior as conditioning on non-extinction. A branching process in varying environment with immigration $(Y_n,\, n \geq 0)$ is defined as

$$Y_0 = 0, \quad \text{and } Y_n = \sum_{j=1}^{Y_{n-1}} \xi_{n,j} + \varepsilon_n, \quad n \in \mathbb{N}, \tag{2}$$

where $\{\xi_n, \xi_{n,j}, \varepsilon_n\}_{n,j\in\mathbb{N}}$ are nonnegative, independent random variables such that

$\{\xi_n, \xi_{n,j}\}_{j\in\mathbb{N}}$ are identically distributed. As before $\xi_{n,j}$ is the number of offspring of the jth individual in generation $n-1$ and ε_n is the number of immigrants.

The study of nearly degenerate BPVE with immigration was initiated by Györfi et al. [12], where it was assumed that the offspring have Bernoulli distribution. The most interesting finding in [12] is that the slow dying out effect can be balanced by a slow immigration rate to obtain nontrivial distributional limit without normalization. In this case the resulting process is an inhomogeneous first-order integer-valued autoregressive process (INAR(1)). INAR processes are important in various fields of applied probability, for theory and applications we refer to the survey of Weiß [20]. The setup of [12] was extended by Kevei [17], allowing more general offspring distributions. For INAR(1) processes the condition $\overline{f}_n < 1$ is automatic (being a probability), while in the more general setup of [17] it was assumed. Our conditions (C1)–(C2) do allow supercritical generations, i.e. $\overline{f}_n > 1$, with dominating subcritical regime. The multitype case was studied by Györfi et al. [11].

In general, Galton–Watson processes with immigration in varying environment (that is, with time-dependent immigration) are less studied. Gao and Zhang [8] proved central limit theorem and law of iterated logarithm. Ispány [13] obtained diffusion approximation in the strongly critical case, that is, when instead of (C1) condition $\sum_{n=1}^{\infty} |1-\overline{f}_n| < \infty$ holds. González et al. [9] investigated a.s. extinction, and obtained limit theorems for the properly normalized process. Since we prove limit theorems for the process without normalization, both our results and assumptions are rather different from those in the mentioned papers.

Generalizing the results in [17], in Theorems 2 and 3 we show that under appropriate assumption on the immigration, the slow extinction trend and the slow immigration trend are balanced and we get a nontrivial compound-Poisson limit distribution without normalization.

The rest of the paper is organized as follows. Section 2 contains the main results. All the proofs are gathered together in Section 3. The proofs are rather technical, and based on analysis of the composite g.f. We rely on the shape function technique worked out in [15].

2. Main results

2.1. Yaglom-type limit theorems

Consider the BPVE (X_n) in (1). Condition (C1) implies that the process dies out almost surely. In [15] a BPVE process (X_n) is called *regular* if

$$\mathbb{E}[\xi_n^2 \mathbb{1}(\xi_n \geq 2)] \leq c\, \mathbb{E}[\xi_n \mathbb{1}(\xi_n \geq 2)]\, \mathbb{E}[\xi_n | \xi_n \geq 1]$$

holds for some $c > 0$ and for all n, where $\mathbb{1}$ stands for the indicator function. In our setup, if (C1)–(C3) are satisfied with $\nu > 0$ then the process is regular, while if $\nu = 0$ there are nonregular examples (e.g. $\mathbb{P}(\xi_n = 0) = n^{-1}$, $\mathbb{P}(\xi_n = 1) = 1 - n^{-1} - n^{-4}$, $\mathbb{P}(\xi_n = n) = n^{-4}$ works). For regular processes the results in [15] apply. According to Kersting's classification ([15, Proposition 1]) of *regular* BPVEs, a regular process satisfying (C1)-(C3) is subcritical. Indeed, $\mathbb{E}X_n \to 0$, and by Lemma 2 and (C2)

$$\lim_{n\to\infty} \overline{f}_{0,n} \sum_{k=1}^{n} \frac{f_k''(1)}{\overline{f}_{0,k-1}\overline{f}_k^2} = \nu.$$

Furthermore, [15, Corollary 2] states that the sequence $\mathcal{L}(X_n|X_n > 0)$ is tight if and only if $\sup_{n\geq 0} \mathbb{E}[X_n|X_n > 0] < \infty$. Here $\mathcal{L}(X_n|X_n > 0)$ stands for the law of X_n conditioned on non-extinction. In Lemma 4 below we show that the latter condition holds, in fact the limit exists. Therefore, the sequence of conditional laws is tight. In the next result, we prove that the limit distribution also exists.

The random variable V has geometric distribution with parameter $p \in (0,1]$, $V \sim \mathrm{Geom}(p)$, if $\mathbb{P}(V = k) = (1-p)^{k-1}p$, $k = 1, 2, \ldots$.

Theorem 1. *Assume that (C1)–(C3) are satisfied. Then, for the BPVE* (X_n) *in* (1)

$$\mathcal{L}(X_n|X_n > 0) \xrightarrow{\mathcal{D}} \mathrm{Geom}\left(\frac{2}{2+\nu}\right) \quad \text{as } n \to \infty,$$

where $\xrightarrow{\mathcal{D}}$ *denotes convergence in distribution.*

The result holds with $\nu = 0$, in which case the sequence of random variables converges in probability to 1.

Example 1. Let $f_n(s) = f_n[0] + f_n[1]s + f_n[2]s^2$, then $\overline{f}_n = f_n[1] + 2f_n[2]$, $f_n''(1) = 2f_n[2]$, and $f_n'''(1) = 0$. Assuming that $f_n[0] + f_n[2] \to 0$, $f_n[0] > f_n[2]$, $\sum_{n=1}^{\infty}(f_n[0] - f_n[2]) = \infty$ and $2f_n[2]/(f_n[0] - f_n[2]) \to \nu \in [0,\infty)$ as $n \to \infty$, the conditions of Theorem 1 are fulfilled.

2.2. Nearly degenerate branching processes with immigration

Recall (Y_n) the BPVE with immigration from (2), and introduce the factorial moments of the immigration

$$m_{n,k} := \mathbb{E}[\varepsilon_n(\varepsilon_n - 1)\cdots(\varepsilon_n - k + 1)], \quad k \in \mathbb{N}.$$

Theorem 2. *Assume that (C1)–(C3) are satisfied and*

(C4) $\lim_{n\to\infty, \overline{f}_n<1} \frac{m_{n,k}}{k!(1-\overline{f}_n)} = \lambda_k$, $k = 1, 2, \ldots, K$, *with* $\lambda_K = 0$, *and for each* $k \leq K$ *the sequence* $(m_{n,k}/|1-\overline{f}_n|)_n$ *is bounded.*

Then, for the BPVE with immigration (Y_n) *in* (2)

$$Y_n \xrightarrow{\mathcal{D}} Y \quad \text{as } n \to \infty,$$

where the random variable Y *has compound-Poisson distribution, and for its g.f.* f_Y

$$\log f_Y(s) = \begin{cases} -\sum_{k=1}^{K-1} \frac{2^k \lambda_k}{\nu^k} \left(\log\left(1 + \frac{\nu}{2}(1-s)\right) + \sum_{i=1}^{k-1} \frac{\nu^i}{i 2^i}(s-1)^i \right), & \nu > 0, \\ \sum_{k=1}^{K-1} \frac{\lambda_k}{k}(s-1)^k, & \nu = 0. \end{cases}$$

Note that f_Y is continuous in $\nu \in [0,\infty)$. Under further assumptions we might allow infinitely many nonzero λ's.

Theorem 3. *Assume that (C1)–(C3) are satisfied and*

(C4') $\lim_{n\to\infty, \overline{f}_n<1} \frac{m_{n,k}}{k!(1-\overline{f}_n)} = \lambda_k$, $k = 1, 2, \ldots$, *such that* $\limsup_{n\to\infty} \lambda_n^{1/n} \leq 1$, *and* $(m_{n,k}/|1-\overline{f}_n|)_n$ *is bounded for each* k.

Then, for the BPVE with immigration (Y_n) *in* (2)

$$Y_n \xrightarrow{\mathcal{D}} Y \quad \text{as } n \to \infty,$$

where the random variable Y *has compound-Poisson distribution, and for its g.f.* f_Y

$$\log f_Y(s) = \begin{cases} -\sum_{k=1}^{\infty} \frac{2^k \lambda_k}{\nu^k} \left(\log\left(1 + \frac{\nu}{2}(1-s)\right) + \sum_{i=1}^{k-1} \frac{\nu^i}{i 2^i} (s-1)^i \right), & \nu > 0, \\ \sum_{k=1}^{\infty} \frac{\lambda_k}{k} (s-1)^k, & \nu = 0. \end{cases} \tag{3}$$

Remark 1. If $K = 2$ in Theorem 2 then Y has (generalized) negative binomial distribution with parameters $\frac{2\lambda_1}{\nu}$ (not necessarily integer) and $\frac{2}{2+\nu}$ as shown in Theorem 5 of [17]. In particular, Theorems 2 and 3 are generalizations of [17, Theorem 5].

If $\nu = 0$ in condition (C2), then the results in Theorems 2 and 3 follow from [17, Theorem 4], when $\overline{f}_n < 1$ for all n. Our results are new even in this special case.

Remark 2. Conditions (C4) and (C4') mean that the immigration is of the proper order, that is $\lim_{n\to\infty, \overline{f}_n < 1} \frac{\mathbb{P}(\varepsilon_n = j)}{1 - \overline{f}_n}$ exists for each $j \geq 1$, see Theorem 4.

Conditions (C4) and (C4') are not comparable in the sense that none of them implies the other. Indeed, if (C4) holds with $K = 3$, then it is possible that 4th moments does not exists, while in (C4') all moments have to be finite.

One can construct examples for which $\lambda_2 = 0$, $\lambda_3 = 1$, $\lambda_4 = \infty$. In this case (C4) in Theorem 2 holds with $K = 2$. On the other hand, it is easy to show that if $\lambda_{n_1} = \lambda_{n_2} = 0$, for some $n_1 < n_2$, then $\lambda_n = 0$ for $n_1 < n < n_2$.

The g.f. f_Y of the limit has a rather complicated form. In the proof, showing the pointwise convergence of the g.f.s, we prove that the accompanying laws converge in distribution to Y. Since the accompanying laws are compound-Poisson, this implies (see e.g. Steutel and van Harn [19, Proposition 2.2]) that the limit Y is compound-Poisson too. That is, $Y = \sum_{i=1}^{N} Z_i$, where $Z, Z_1, Z_2, \ldots$ are i.i.d. nonnegative integer valued random variables, and independently N has Poisson distribution. This is an important class of distributions, since it is exactly the class of infinitely divisible distributions on the nonnegative integers, see e.g. [19, Theorem II.3.2]. Interestingly, from the form of the g.f. f_Y it is difficult to deduce that it is compound-Poisson, because the logarithm is given as a power series in $1-s$, while for the compound-Poisson this is a power series in s. Under some conditions on the sequence (λ_n) we can rewrite $\log f_Y$ to a series expansion in s, which allows us to understand the structure of the limit.

Theorem 4. *Assume one of the following:*

(i) conditions of Theorem 2 hold; or

(ii) conditions of Theorem 3 hold and $\limsup_{n\to\infty} \lambda_n^{1/n} \leq 1/2$.

Then the limiting g.f. in Theorems 2 and 3 can be written as

$$f_Y(s) = \exp\left\{ \sum_{n=1}^{\infty} A_n (s^n - 1) \right\},$$

where

$$A_n = \begin{cases} \frac{\nu^n}{n(2+\nu)^n} \sum_{j=1}^{\infty} q_j \left[\left(1 + \frac{2}{\nu}\right)^{\min(j,n)} - 1 \right], & \text{if } \nu > 0, \\ \frac{1}{n} \sum_{j=n}^{\infty} q_j, & \text{if } \nu = 0, \end{cases}$$

and

$$q_j = \lim_{n\to\infty, \overline{f}_n < 1} \frac{\mathbb{P}(\varepsilon_n = j)}{1-\overline{f}_n}, \quad j = 1, 2, \ldots.$$

Remark 3. For $\nu = 0$, the latter formula was obtained in [12, Remark 2].

In fact, the last limit exists without any extra conditions on the $(\lambda_n)_n$ sequence, see Lemma 8. Furthermore, it is clear that under (C4) $q_k = 0$ for $k \geq K$.

The form of f_Y also implies that the limit has the representation $Y = \sum_{n=1}^{\infty} n N_n$, where (N_n) are independent Poisson random variables, such that N_n has parameter A_n.

3. Proofs

3.1. Preparation

In the proofs we analyze the g.f. of the underlying processes. To prove distributional convergence for nonnegative integer valued processes, it is enough to prove the pointwise convergence of the g.f., and show that the limit is a g.f. as well, see Feller [7, p.280].

Recall that $f_n(s) = \mathbb{E}s^{\xi_n}$ stands for the offspring g.f. in generation n. For the composite g.f. introduce the notation $f_{n,n}(s) = s$, for $j < n$

$$f_{j,n}(s) := f_{j+1} \circ \ldots \circ f_n(s),$$

and for the corresponding means $\overline{f}_{n,n} = 1$,

$$\overline{f}_{j,n} := \overline{f}_{j+1} \cdots \overline{f}_n, \quad j < n.$$

Then it is well-known that $\mathbb{E}s^{X_n} = f_{0,n}(s)$ and $\mathbb{E}X_n = \overline{f}_{0,n}$.

For a g.f. f, with mean $\overline{f}$ and $f''(1) < \infty$, define the *shape function* as

$$\varphi(s) = \frac{1}{1-f(s)} - \frac{1}{\overline{f}(1-s)}, \ 0 \leq s < 1, \quad \varphi(1) = \frac{f''(1)}{2\,(\overline{f})^2}. \tag{4}$$

Let φ_j be the shape function of f_j. By definition of $f_{j,n}$

$$\frac{1}{1-f_{j,n}(s)} = \frac{1}{\overline{f}_{j+1}(1-f_{j+1,n}(s))} + \varphi_{j+1}(f_{j+1,n}(s)).$$

Therefore, iteration gives ([15, Lemma 5], [16, Proposition 1.3])

$$\frac{1}{1-f_{j,n}(s)} = \frac{1}{\overline{f}_{j,n}(1-s)} + \varphi_{j,n}(s), \tag{5}$$

where

$$\varphi_{j,n}(s) := \sum_{k=j+1}^{n} \frac{\varphi_k(f_{k,n}(s))}{\overline{f}_{j,k-1}}. \tag{6}$$

The latter formulas show the usefulness of the shape function. The next statement gives precise upper and lower bounds on the shape function.

Lemma 1. ([15, Lemma 1], [16, Proposition 1.4].) *Assume* $0 < \overline{f} < \infty$, $f''(1) < \infty$ *and let* $\varphi(s)$ *be the shape function of* f. *Then, for* $0 \leq s \leq 1$,

$$\frac{1}{2}\varphi(0) \leq \varphi(s) \leq 2\varphi(1).$$

For further properties of shape functions we refer to [16, Chapter 1] and to [15].

We frequently use the following extension of Lemma 5 from [12], which is a version of the Silverman–Toeplitz theorem.

Lemma 2. *Let* $(\overline{f}_n)_{n\in\mathbb{N}}$ *be a sequence of positive real numbers satisfying (C1), and define*

$$a_{n,j}^{(k)} = (1-\overline{f}_j) \prod_{i=j+1}^{n} \overline{f}_i^k = (1-\overline{f}_j)\overline{f}_{j,n}^k, \quad n,j,k \in \mathbb{N}, \quad j \leq n-1,$$

and $a_{n,n}^{(k)} = 1 - \overline{f}_n$. *If* $(x_n)_{n\in\mathbb{N}}$ *is bounded and* $\lim_{n\to\infty, \overline{f}_n<1} x_n = x \in \mathbb{R}$, *then for all* $k \in \mathbb{N}$

$$\begin{aligned} \lim_{n\to\infty} \sum_{j=1}^{n} a_{n,j}^{(k)} x_j &= \frac{x}{k}, \\ \lim_{n\to\infty} \sum_{j=1}^{n} |a_{n,j}^{(k)}| x_j &= \frac{x}{k}. \end{aligned} \tag{7}$$

Proof. Let us define

$$A = \{j : \overline{f}_j > 1\}, \quad A_n = A \cap \{1, \ldots, n\}.$$

First we show that the following conditions (similar to that of the Silverman–Toeplitz theorem) hold:

(i) $\lim_{n\to\infty} a_{n,j}^{(k)} = 0$, for all $j \in \mathbb{N}$,

(ii) $\lim_{n\to\infty} \sum_{j=1}^{n} a_{n,j}^{(k)} = \frac{1}{k}$,

(iii) $\sup_{n\in\mathbb{N}} \sum_{j=1}^{n} \left| a_{n,j}^{(k)} \right| < \infty$,

(iv) $\lim_{n\to\infty} \sum_{j\in A_n} |a_{n,j}^{(k)}| = 0$.

It is easy to see that (i) holds, as by (C1)

$$|a_{n,j}^{(k)}| = |1-\overline{f}_j| \prod_{\ell=j+1}^{n} \overline{f}_\ell^k \leq |1-\overline{f}_j| \exp\Big\{ -k \sum_{\ell=j+1}^{n} (1-\overline{f}_\ell) \Big\} \to 0.$$

Here, and later on any nonspecified limit relation is meant as $n \to \infty$. Since

$$0 \leq \overline{f}_{0,n} \leq \exp\Big\{ -\sum_{j=1}^{n} (1-\overline{f}_j) \Big\} \to 0,$$

we have

$$\sum_{j=1}^{n} a_{n,j}^{(1)} = \sum_{j=1}^{n} (1-\overline{f}_j)\overline{f}_{j,n} = \sum_{j=1}^{n} (\overline{f}_{j,n} - \overline{f}_{j-1,n}) = 1 - \overline{f}_{0,n} \to 1,$$

which is (ii) for $k=1$. Furthermore,

$$\sum_{j=1}^{n} |a_{n,j}^{(1)}| = \sum_{j=1}^{n} |1-\overline{f}_j| \cdot \overline{f}_{j,n} = \sum_{j=1}^{n} \left[(1-\overline{f}_j) + 2(1-\overline{f}_j)_- \right] \overline{f}_{j,n} < \infty,$$

that is (iii) holds for $k=1$. Note that $\sup_{j,n} \overline{f}_{j,n} < \infty$ by (C1). To see (iv), we have

$$\sum_{j \in A_n} |a_{n,j}^{(k)}| = \sum_{j=1}^{n} (\overline{f}_j - 1)_+ \overline{f}_{j,n}^k \to 0,$$

since $\lim_{n\to\infty} \overline{f}_{j,n} = 0$, and $((\overline{f}_j - 1)_+ \sup_{\ell,n} \overline{f}_{\ell,n}^k)_j$ is an integrable majorant, so Lebesgue's dominated convergence theorem applies.

Before proving (ii) and (iii) for $k \geq 2$ we show (7) for $k=1$. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence with the required properties. Then,

$$\sum_{j=1}^{n} a_{n,j}^{(1)} x_j - x = \sum_{j=1}^{n} a_{n,j}^{(1)} (x_j - x) + x \left(\sum_{j=1}^{n} a_{n,j}^{(1)} - 1 \right),$$

where the second term tends to 0 by (ii). For the first term we have

$$\begin{aligned} \left| \sum_{j=1}^{n} a_{n,j}^{(1)} (x_j - x) \right| &= \left| \sum_{j \in A_n} a_{n,j}^{(1)} (x_j - x) + \sum_{j \notin A_n} a_{n,j}^{(1)} (x_j - x) \right| \\ &\leq 2 \sup_k |x_k| \cdot \sum_{j \in A_n} |a_{n,j}^{(1)}| + \sum_{j \notin A_n} |a_{n,j}^{(1)} (x_j - x)|, \end{aligned}$$

where the first term tends to 0 by (iv), and the second tends to 0 by (i) and (iii). Thus the first line in (7) holds for $k=1$. Furthermore, since $a_{j,n}^{(k)} \leq 0$ if and only if $j \in A_n$, conditions (i)–(iv) remains true for $|a_{n,j}^{(k)}|$, thus

$$\lim_{n\to\infty} \sum_{j=1}^{n} |a_{n,j}^{(1)}| \, x_j = x. \tag{8}$$

Next we prove (ii) and (iii) for arbitrary $k \geq 2$. By the binomial theorem

$$1 - \overline{f}_{0,n}^k = k \sum_{j=1}^{n} a_{n,j}^{(k)} + \sum_{i=2}^{k} (-1)^{i+1} \binom{k}{i} \sum_{j=1}^{n} (1-\overline{f}_j)^{i-1} a_{n,j}^{(k)}. \tag{9}$$

Moreover, for $i \geq 2$

$$\begin{aligned} \sum_{j=1}^{n} \left| (1-\overline{f}_j)^{i-1} a_{n,j}^{(k)} \right| &= \sum_{j=1}^{n} \left| 1-\overline{f}_j \right|^i \overline{f}_{j,n}^k \\ &\leq (\sup_{\ell,n} \overline{f}_{\ell,n})^{k-1} \sum_{j=1}^{n} |a_{n,j}^{(1)}| |1-\overline{f}_j|^{i-1} \to 0, \end{aligned}$$

by (8). Thus all the terms in the second sum on the RHS of (9) tend to 0. Since $\overline{f}_{0,n} \to 0$, (ii) follows. Then (iii) follows as for $k = 1$. This completes the proof of (ii) and (iii) for $k \geq 2$. Then (7) for $k \geq 2$ follows exactly as for $k = 1$. □

Lemma 3. *Let φ_n be the shape function of f_n. Then under the conditions of Theorem 1 with $\nu > 0$*

$$\lim_{n\to\infty, \overline{f}_n < 1} \sup_{s\in[0,1]} \frac{|\varphi_n(1) - \varphi_n(s)|}{1 - \overline{f}_n} = 0,$$

and the sequence $\sup_{s\in[0,1]} \frac{|\varphi_n(1)-\varphi_n(s)|}{|1-\overline{f}_n|}$ *is bounded.*

Proof. To ease notation we suppress the lower index n. By the Taylor expansion

$$f(s) = 1 + \overline{f}\,(s-1) + \frac{1}{2} f''(1)(s-1)^2 + \frac{1}{6} f'''(t)(s-1)^3,$$

for some $t \in (s, 1)$. Thus, recalling (4) the shape function can be written in the form

$$\varphi(s) = \frac{\overline{f}\,(1-s) - 1 + f(s)}{\overline{f}\,(1-s)(1-f(s))} = \frac{f''(1)}{2\overline{f}^2} \frac{1 - \frac{f'''(t)}{3f''(1)}(1-s)}{1 - \frac{f''(1)}{2\overline{f}}(1-s) + \frac{f'''(t)}{6\overline{f}}(1-s)^2}.$$

Therefore,

$$\frac{\varphi(1) - \varphi(s)}{1 - \overline{f}} = \frac{f''(1)}{2\overline{f}^2(1-\overline{f})} \left(1 - \frac{1 - \frac{f'''(t)}{3f''(1)}(1-s)}{1 - \frac{f''(1)}{2\overline{f}}(1-s) + \frac{f'''(t)}{6\overline{f}}(1-s)^2} \right). \tag{10}$$

Using the assumptions and the monotonicity of f''', uniformly in $s \in (0, 1]$

$$\lim_{n\to\infty, \overline{f}_n<1} \left[\frac{f_n'''(t)}{f_n''(1)} + \frac{f_n''(1)}{\overline{f}_n} + \frac{f_n'''(t)}{\overline{f}_n} \right] = 0,$$

thus convergence on $\{n : \overline{f}_n < 1\}$ follows.

The boundedness also follows from (10), since if $\overline{f}_n > 1$ then

$$f_n''(1) = \mathbb{E}(\xi_n(\xi_n - 1)) \geq \mathbb{E}(\xi_n - 1) = \overline{f}_n - 1,$$

and $f_n'''(1)/|1 - \overline{f}_n|$ is bounded by assumption (C3). □

3.2. Proof of Theorem 1

Lemma 4. *Under the conditions of Theorem 1, for $s \in [0, 1)$*

$$\lim_{n\to\infty} \frac{\overline{f}_{0,n}}{1 - f_{0,n}(s)} = \frac{1}{1-s} + \frac{\nu}{2}.$$

Proof. Recalling (5) and (6) we have to show that

$$\overline{f}_{0,n}\varphi_{0,n}(s) = \sum_{j=1}^{n} \overline{f}_{j-1,n}\varphi_j(f_{j,n}(s)) \to \frac{\nu}{2}.$$

First let $\nu = 0$. Using Lemmas 1 and 2,

$$\overline{f}_{0,n}\varphi_{0,n}(s) \leq \sum_{j=1}^{n} \overline{f}_{j,n} \frac{f_j''(1)}{\overline{f}_j} = \sum_{j=1}^{n} |a_{n,j}^{(1)}| \frac{f_j''(1)}{\overline{f}_j |1-\overline{f}_j|} \to 0.$$

If $\overline{f}_j = 1$ then by (C2) necessarily $f_j''(1) = 0$, and $0/0$ in these type of sums are meant as 0.

For $\nu \in (0,\infty)$ write

$$\sum_{j=1}^{n} \overline{f}_{j-1,n}\varphi_j(f_{j,n}(s)) = \sum_{j=1}^{n} \overline{f}_{j-1,n}\varphi_j(1) - \sum_{j=1}^{n} \overline{f}_{j-1,n}(\varphi_j(1) - \varphi_j(f_{j,n}(s))). \tag{11}$$

By Lemma 2 for the first term we have

$$\sum_{j=1}^{n} \overline{f}_{j-1,n}\varphi_j(1) = \frac{1}{2}\sum_{j=1}^{n} a_{n,j}^{(1)} \frac{1}{\overline{f}_j} \frac{f_j''(1)}{1-\overline{f}_j} \to \frac{\nu}{2}. \tag{12}$$

For the second term in (11),

$$\left| \sum_{j=1}^{n} \overline{f}_{j-1,n}(\varphi_j(1) - \varphi_j(f_{j,n}(s))) \right| \leq \sum_{j=1}^{n} a_{n,j}^{(1)} \overline{f}_j \sup_{s\in[0,1]} \frac{|\varphi_j(1) - \varphi_j(s)|}{|1-\overline{f}_j|} \to 0,$$

according to Lemmas 2 and 3. Combining with (11) and (12) the statement follows. □

Proof of Theorem 1. We prove the convergence of the conditional g.f. For $s \in (0,1)$ we have by the previous lemma

$$\mathbb{E}[s^{X_n} | X_n > 0] = \frac{f_{0,n}(s) - f_{0,n}(0)}{1 - f_{0,n}(0)} = 1 - \frac{1 - f_{0,n}(s)}{1 - f_{0,n}(0)} \to \frac{2}{2+\nu} \frac{s}{1 - \frac{\nu}{\nu+2}s},$$

where the limit is the g.f. of the geometric distribution with parameter $\frac{2}{2+\nu}$. □

3.3. Proofs of subsection 2.2

We need a strange version of the Riemann approximation, where the points in the partition are not necessarily increasing. It follows immediately from the uniform continuity.

Lemma 5. *Assume that* $x_{n,j} \in [0,1]$, $j = 0,\ldots,n$, $n \in \mathbb{N}$, *such that* $x_{n,0} = 0$, $x_{n,n} = 1$, $\lim_{n\to\infty} \sup_j |x_{n,j} - x_{n,j-1}| = 0$, *and* $\sup_n \sum_{j=1}^n |x_{n,j} - x_{n,j-1}| < \infty$. *If* f *is continuous on* $[0,1]$ *then*

$$\lim_{n\to\infty} \sum_{j=1}^{n} f(u_{n,j})(x_{n,j} - x_{n,j-1}) = \int_0^1 f(x)\,\mathrm{d}x,$$

with $u_{n,j} \in [\min(x_{n,j-1}, x_{n,j}), \max(x_{n,j-1}, x_{n,j})]$.

We need a simple lemma on an alternating sum involving binomial coefficients.

Lemma 6. *For $x \in \mathbb{R}$*

$$\sum_{i=1}^{k-1} \binom{k-1}{i} (-1)^i \frac{(1+x)^i - 1}{i} = \sum_{i=1}^{k-1} (-1)^i \frac{x^i}{i}.$$

Proof. Both sides are polynomials of x, and equality holds for $x = 0$, therefore it is enough to show that the derivatives are equal. Differentiating the LHS we obtain

$$\sum_{i=1}^{k-1} \binom{k-1}{i} (-1)^i (1+x)^{i-1} = \frac{1}{1+x} \left[(1 - (1+x))^{k-1} - 1 \right].$$

Differentiating the RHS and multiplying by $(1+x)$, the statement follows. □

The next statement is an easy consequence of the Taylor expansion of the g.f.

Lemma 7. ([12, Lemma 6].) *Let ε be a nonnegative integer-valued random variable with factorial moments*

$$m_k := \mathbb{E}[\varepsilon(\varepsilon-1)\cdots(\varepsilon-k+1)], \quad k \in \mathbb{N},$$

$m_0 := 1$ and with g.f. $h(s) = \mathbb{E}s^\varepsilon$, $s \in [-1,1]$. If $m_\ell < \infty$ for some $\ell \in \mathbb{N}$ with $|s| \le 1$, then

$$h(s) = \sum_{k=0}^{\ell-1} \frac{m_k}{k!}(s-1)^k + R_\ell(s), \quad |R_\ell(s)| \le \frac{m_\ell}{\ell!}|s-1|^\ell.$$

Proof of Theorem 2. Recall that $f_n(s) = \mathbb{E}s^{\xi_n}$, and let $g_n(s) = \mathbb{E}s^{Y_n}$ and $h_n(s) = \mathbb{E}s^{\varepsilon_n}$. Then the branching property gives that (see e.g. [9, Proposition 1])

$$g_n(s) = \prod_{j=1}^n h_j(f_{j,n}(s)).$$

We prove the convergence of the g.f., that is

$$\lim_{n\to\infty} g_n(s) = f_Y(s), \quad s \in [0,1].$$

Fix $s \in [0,1)$. Let us introduce

$$\widehat{g}_n(s) = \prod_{j=1}^n e^{h_j(f_{j,n}(s))-1}.$$

Note that $\widehat{g}_n$ is a kind of accompanying law, its distribution is compound-Poisson, therefore its limit is compound-Poisson too, see [19, Proposition 2.2]. By convexity of the g.f.

$$f_{j,n}(s) \ge 1 + \overline{f}_{j,n}(s-1). \tag{13}$$

If $|a_k|, |b_k| < 1$, $k = 1, \ldots, n$ then

$$\left| \prod_{k=1}^n a_k - \prod_{k=1}^n b_k \right| \le \sum_{k=1}^n |a_k - b_k|. \tag{14}$$

Consequently, using (14), the inequalities $|e^u - 1 - u| \le u^2$ for $|u| \le 1$, and $0 \le 1 - h_j(s) \le m_{j,1}(1-s)$, and (13), we have

$$\begin{aligned} |g_n(s) - \widehat{g}_n(s)| &\le \sum_{j=1}^n |e^{h_j(f_{j,n}(s))-1} - h_j(f_{j,n}(s))| \\ &\le \sum_{j=1}^n (h_j(f_{j,n}(s)) - 1)^2 \le \sum_{j=1}^n \frac{m_{j,1}^2}{|1-\overline{f}_j|} |a_{n,j}^{(2)}| \to 0, \end{aligned} \tag{15}$$

where we used Lemma 2, since $\lim_{n\to\infty, \overline{f}_n < 1} \frac{m_{n,1}}{1-\overline{f}_n} = \lambda_1$ implies $\lim_{n\to\infty, \overline{f}_n<1} \frac{m_{n,1}^2}{1-\overline{f}_n} = 0$.

Therefore, we need to show that

$$\lim_{n\to\infty} \sum_{j=1}^n (h_j(f_{j,n}(s)) - 1) = \log f_Y(s). \tag{16}$$

By Lemma 7,

$$h_j(s) = \sum_{k=0}^{K-1} \frac{m_{j,k}}{k!} (s-1)^k + R_{j,K}(s), \tag{17}$$

where $|R_{j,K}(s)| \le \frac{m_{j,K}}{K!}(1-s)^K$, thus

$$\begin{aligned} \sum_{j=1}^n (h_j(f_{j,n}(s)) - 1) &= \sum_{j=1}^n \left[\sum_{k=1}^{K-1} \frac{m_{j,k}}{k!} (f_{j,n}(s) - 1)^k + R_{j,K}(f_{j,n}(s)) \right] \\ &= \sum_{k=1}^{K-1} (-1)^k \sum_{j=1}^n \frac{m_{j,k}}{k!(1-\overline{f}_j)} a_{n,j}^{(k)} \left(\frac{1 - f_{j,n}(s)}{\overline{f}_{j,n}} \right)^k + \sum_{j=1}^n R_{j,K}(f_{j,n}(s)). \end{aligned} \tag{18}$$

By (13) and Lemma 2,

$$\begin{aligned} \left| \sum_{j=1}^n R_{j,K}(f_{j,n}(s)) \right| &\le \sum_{j=1}^n \frac{m_{j,K}}{K!} |f_{j,n}(s) - 1|^K \le \sum_{j=1}^n \frac{m_{j,K}}{K!} \overline{f}_{j,n}^K (1-s)^K \\ &= (1-s)^K \sum_{j=1}^n \frac{m_{j,K}}{K! |1-\overline{f}_j|} |a_{n,j}^{(K)}| \to 0. \end{aligned} \tag{19}$$

Till this point everything works for $\nu \ge 0$. Now assume that $\nu > 0$. Then, by Lemmas 2 and 3,

$$\left| \sum_{i=j+1}^n a_{n,i}^{(1)} \overline{f}_i \frac{\varphi_i(1) - \varphi_i(f_{i,n}(s))}{1 - \overline{f}_i} \right| \le \sup_{k\ge 1} \overline{f}_k \sum_{i=1}^n |a_{n,i}^{(1)}| \sup_{t\in[0,1]} \frac{|\varphi_i(1) - \varphi_i(t)|}{|1-\overline{f}_i|} \to 0, \tag{20}$$

and similarly

$$\left| \sum_{i=j+1}^n a_{n,i}^{(1)} \overline{f}_i \frac{\varphi_i(1)}{1-\overline{f}_i} - \frac{\nu}{2} \sum_{i=j+1}^n a_{n,i}^{(1)} \right| \le \sum_{i=1}^n |a_{n,i}^{(1)}| \left| \frac{1}{\overline{f}_i} \frac{f_i''(1)}{2(1-\overline{f}_i)} - \frac{\nu}{2} \right| \to 0. \tag{21}$$

Putting

$$\varepsilon_{j,n} = \sum_{i=j+1}^{n} a_{n,i}^{(1)} \overline{f}_i \frac{\varphi_i(f_{i,n}(s))}{1-\overline{f}_i} - \frac{\nu}{2}(1-\overline{f}_{j,n}), \tag{22}$$

by (5) and (6)

$$\begin{aligned} \frac{\overline{f}_{j,n}}{1-f_{j,n}(s)} &= \frac{1}{1-s} + \sum_{i=j+1}^{n} a_{n,i}^{(1)} \overline{f}_i \frac{\varphi_i(f_{i,n}(s))}{1-\overline{f}_i} \\ &= \frac{1}{1-s} + \frac{\nu}{2}(1-\overline{f}_{j,n}) + \varepsilon_{j,n}. \end{aligned} \tag{23}$$

Noting that $\sum_{i=j+1}^{n} a_{n,i}^{(1)} = (1-\overline{f}_{j,n})$, (20), (21) and the triangle inequality imply for $\varepsilon_{j,n}$ in (22) that

$$\max_{j\le n} |\varepsilon_{j,n}| = \max_{j\le n} \left| \frac{\overline{f}_{j,n}}{1-f_{j,n}(s)} - \frac{1}{1-s} - \frac{\nu}{2}(1-\overline{f}_{j,n}) \right| \to 0. \tag{24}$$

The latter further implies that

$$\limsup_{n\to\infty} \max_{j\le n} \frac{1-f_{j,n}(s)}{\overline{f}_{j,n}} \le 1-s, \tag{25}$$

and that for n large enough by the mean value theorem and (23)

$$\left| \left(\frac{\overline{f}_{j,n}}{1-f_{j,n}(s)} \right)^{-k} - \left(\frac{1}{1-s} + \frac{\nu}{2}(1-\overline{f}_{j,n}) \right)^{-k} \right| \le k\varepsilon_{j,n}.$$

Thus, by (25)

$$\begin{aligned} &\left| \sum_{j=1}^{n} a_{n,j}^{(k)} \left[\frac{m_{j,k}}{k!(1-\overline{f}_j)} \left(\frac{1-f_{j,n}(s)}{\overline{f}_{j,n}} \right)^{k} - \lambda_k \left(\frac{1}{1-s} + \frac{\nu}{2}(1-\overline{f}_{j,n}) \right)^{-k} \right] \right| \\ &\le \sum_{j=1}^{n} |a_{n,j}^{(k)}| \left[\left| \frac{m_{j,k}}{k!(1-\overline{f}_j)} - \lambda_k \right| \left(\frac{1-f_{j,n}(s)}{\overline{f}_{j,n}} \right)^{k} \right. \\ &\qquad \left. + \lambda_k \left| \left(\frac{1-f_{j,n}(s)}{\overline{f}_{j,n}} \right)^{k} - \left(\frac{1}{1-s} + \frac{\nu}{2}(1-\overline{f}_{j,n}) \right)^{-k} \right| \right] \\ &\le \sum_{j=1}^{n} |a_{n,j}^{(k)}| \left| \frac{m_{j,k}}{k!(1-\overline{f}_j)} - \lambda_k \right| + \lambda_k \sum_{j=1}^{n} |a_{n,j}^{(k)}| k \max_{j\le n} \varepsilon_{j,n} \to 0, \end{aligned} \tag{26}$$

where the second inequality holds for n large enough.

Furthermore, by Lemma 5

$$\lim_{n\to\infty} \sum_{j=1}^{n} a_{n,j}^{(k)} \left(\frac{1}{1-s} + \frac{\nu}{2}(1-\overline{f}_{j,n}) \right)^{-k} = \int_0^1 \frac{y^{k-1}}{\left(\frac{1}{1-s} + \frac{\nu}{2}(1-y) \right)^k} \mathrm{d}y, \tag{27}$$

since the LHS is a Riemann approximation of the RHS corresponding to the partition $\{\overline{f}_{j,n}\}_{j=-1}^{n}$, with $\overline{f}_{-1,n} := 0$, and $\overline{f}_{j,n} - \overline{f}_{j-1,n} = a_{n,j}^{(1)} \to 0$ uniformly in j according to Lemma 2, while $\overline{f}_{0,n} \to 0$ by (C1). Changing variables $u = \nu(1-s)(1-y) + 2$ and using the binomial theorem,

$$\begin{aligned}
&\int_0^1 \frac{y^{k-1}}{\left(\frac{1}{1-s} + \frac{\nu}{2}(1-y)\right)^k} \,\mathrm{d}y = 2^k(1-s)^k \int_0^1 \frac{y^{k-1}}{(2+\nu(1-s)(1-y))^k} \,\mathrm{d}y \\
&= \frac{2^k}{\nu^k} \int_2^{\nu(1-s)+2} \frac{(2+\nu(1-s)-u)^{k-1}}{u^k} \,\mathrm{d}u \\
&= \frac{2^k}{\nu^k} \int_2^{\nu(1-s)+2} \frac{1}{u^k} \sum_{i=0}^{k-1} \left[\binom{k-1}{i} (2+\nu(1-s))^i (-u)^{k-1-i} \right] \mathrm{d}u \\
&= \frac{2^k}{\nu^k} \sum_{i=0}^{k-1} \left[\binom{k-1}{i} (2+\nu(1-s))^i (-1)^{k-1-i} \int_2^{\nu(1-s)+2} u^{-(i+1)} \,\mathrm{d}u \right] \\
&= (-1)^{k+1} \frac{2^k}{\nu^k} \left(\log\left(1 + \frac{\nu}{2}(1-s)\right) + \sum_{i=1}^{k-1} \binom{k-1}{i} (-1)^i \frac{(1+\frac{\nu}{2}(1-s))^i - 1}{i} \right) \\
&= (-1)^{k+1} \frac{2^k}{\nu^k} \left(\log\left(1 + \frac{\nu}{2}(1-s)\right) + \sum_{i=1}^{k-1} (-1)^i \frac{\nu^i}{i2^i} (1-s)^i \right),
\end{aligned} \tag{28}$$

where the last equality follows from Lemma 6.

Substituting back into (18), by (19), (26), (27), and (28) we obtain (16).

To finish the proof we need to handle the $\nu = 0$ case. Then the calculations are easier. We can still define $\varepsilon_{j,n}$ as in (22), and (21) remains true. Applying Lemma 1, we see that (24) and (25) hold true as well, therefore (26) follows, with $\nu = 0$ everywhere. There is no need for the Riemann approximation, (16) follows from Lemma 2. □

Proof of Theorem 3. Here we consider only the $\nu > 0$ case. For $\nu = 0$ the calculations are easier, and follow similarly as in the previous proof.

Similarly to the proof of Theorem 2, (15) holds, so it is enough to prove that

$$\sum_{j=1}^{n} (h_j(f_{j,n}(s)) - 1) \to \log f_Y(s).$$

Fix $s \in (0,1)$. By Taylor's theorem, for some $\xi \in (0, \nu(1-s)/2)$

$$\log\left(1 + \frac{\nu}{2}(1-s)\right) + \sum_{i=1}^{k-1} (-1)^i \frac{\nu^i}{i2^i} (1-s)^i = \frac{\nu^k (1-s)^k}{2^k} \frac{1}{(1+\xi)^k} \frac{1}{k},$$

so

$$\frac{2^k}{\nu^k} \left| \log\left(1 + \frac{\nu}{2}(1-s)\right) + \sum_{i=1}^{k-1} (-1)^i \frac{\nu^i}{i2^i} (1-s)^i \right| \le \frac{(1-s)^k}{k}.$$

Therefore, for any $\varepsilon > 0$ there exists ℓ large enough such that

$$\begin{aligned} &\left| \sum_{k=\ell}^{\infty} \frac{2^k \lambda_k}{\nu^k} \left(\log\left(1 + \frac{\nu}{2}(1-s)\right) + \sum_{i=1}^{k-1} (-1)^i \frac{\nu^i}{i2^i} (1-s)^i \right) \right| \\ &\leq \sum_{k=\ell}^{\infty} \frac{\lambda_k (1-s)^k}{k} < \varepsilon. \end{aligned} \tag{29}$$

By Lemma 7, (17) holds with $K = \ell$ therefore

$$\begin{aligned} \sum_{j=1}^{n} (h_j(f_{j,n}(s)) - 1) &= \sum_{j=1}^{n} \left[\sum_{k=1}^{\ell-1} \frac{m_{j,k}}{k!} (f_{j,n}(s) - 1)^k + R_{j,\ell}(f_{j,n}(s)) \right] \\ &= \sum_{k=1}^{\ell-1} \sum_{j=1}^{n} \frac{m_{j,k}}{k!} (f_{j,n}(s) - 1)^k + \sum_{j=1}^{n} R_{j,\ell}(f_{j,n}(s)). \end{aligned}$$

Moreover, by (13) and Lemma 2,

$$\begin{aligned} \left| \sum_{j=1}^{n} R_{j,\ell}(f_{j,n}(s)) \right| &\leq \sum_{j=1}^{n} \frac{m_{j,\ell}}{\ell!} |f_{j,n}(s) - 1|^\ell \leq \sum_{j=1}^{n} \frac{m_{j,\ell}}{\ell!} \overline{f}_{j,n}^{\ell} (1-s)^\ell \\ &\to \frac{(1-s)^\ell}{\ell} \lambda_\ell \leq \varepsilon. \end{aligned} \tag{30}$$

Summarizing,

$$\begin{aligned} &\left| \sum_{j=1}^{n} (h_j(f_{j,n}(s)) - 1) - \log f_Y(s) \right| \\ &\leq \left| \sum_{k=1}^{\ell-1} \sum_{j=1}^{n} \frac{m_{j,k}}{k!} (f_{j,n}(s) - 1)^k \right. \\ &\quad \left. + \sum_{k=1}^{\ell-1} \frac{2^k \lambda_k}{\nu^k} \left(\log\left(1 + \frac{\nu}{2}(1-s)\right) - \sum_{i=1}^{k-1} (-1)^{i+1} \frac{\nu^i}{i2^i} (1-s)^i \right) \right| \\ &\quad + \left| \sum_{k=\ell}^{\infty} \frac{2^k \lambda_k}{\nu^k} \left(\log\left(1 + \frac{\nu}{2}(1-s)\right) - \sum_{i=1}^{k-1} (-1)^{i+1} \frac{\nu^i}{i2^i} (1-s)^i \right) \right| \\ &\quad + \left| \sum_{j=1}^{n} R_{j,\ell}(f_{j,n}(s)) \right|, \end{aligned}$$

where the first term on the RHS converges to 0 by the previous result, while the second and third are small for n large by (29) and (30). As $\varepsilon > 0$ is arbitrary, the proof is complete. $\square$

3.4. Proof of Theorem 4

Before the proof we need two auxiliary lemmas.

Lemma 8. *If condition (C4) of Theorem 2 or (C4') of Theorem 3 hold then* $q_i = \lim_{n\to\infty, \overline{f}_n<1} \frac{\mathbb{P}(\varepsilon_n=i)}{1-\overline{f}_n}$ *exists for each* $i = 1, 2, \ldots$.

Proof. Without loss of generality we assume that $\overline{f}_n < 1$ for each n. Otherwise, simply consider everything on the subsequence $\{n : \overline{f}_n < 1\}$.

Let $h_n[i] = \mathbb{P}(\varepsilon_n = i)$. If (C4) holds, there are only finitely many λ's, and the statement follows easily by backward induction. However, if (C4') holds, a more involved argument is needed, which works in both cases.

The kth moment of a random variable can be expressed in terms of its factorial moments as

$$\mu_{n,k} := \mathbb{E}\varepsilon_n^k = \sum_{i=1}^k \left\{ {k \atop i} \right\} m_{n,i},$$

where $\left\{ {k \atop i} \right\} = \frac{1}{i!} \sum_{j=0}^i (-1)^j \binom{i}{j} (i-j)^k$ denotes the Stirling number of the second kind. On Stirling numbers we refer to Graham et al. [10, Section 6.1]. Therefore,

$$\begin{aligned} \sum_{i=1}^\infty \frac{i^k h_n[i]}{1-\overline{f}_n} = \frac{\mu_{n,k}}{1-\overline{f}_n} &= \sum_{i=1}^k \left\{ {k \atop i} \right\} \frac{m_{n,i}}{1-\overline{f}_n} \\ &\to \sum_{i=1}^k \left\{ {k \atop i} \right\} i! \lambda_i =: \mu_k. \end{aligned} \tag{31}$$

Hence, the sequence $(h_n[i]/(1-\overline{f}_n))_{n\in\mathbb{N}}$ is bounded in n for all i. Therefore, any subsequence contains a further subsequence (n_ℓ) such that

$$\lim_{\ell\to\infty} \frac{h_{n_\ell}[i]}{1-\overline{f}_{n_\ell}} = q_i, \quad \forall i \in \mathbb{N}. \tag{32}$$

To prove the statement we have to show that the sequence (q_i) is unique, it does not depend on the subsequence. Note that the sequence $(q_i)_{i\in\mathbb{N}}$ is not necessarily a probability distribution.

By Fatou's lemma and (31)

$$\mu_k = \lim_{\ell\to\infty} \frac{\mu_{n_\ell,k}}{1-\overline{f}_{n_\ell}} \geq \sum_{i=1}^\infty \lim_{\ell\to\infty} \frac{h_{n_\ell}[i]}{1-\overline{f}_{n_\ell}} i^k = \sum_{i=1}^\infty q_i i^k. \tag{33}$$

Let $\varepsilon > 0$ and $k \in \mathbb{N}$ be arbitrary. Write $n_\ell = n$, and put $K := \lfloor \frac{\mu_{k+1}}{\varepsilon} \rfloor + 1$, with $\lfloor \cdot \rfloor$ standing for the lower integer part. Then

$$\sum_{i=1}^\infty \frac{h_n[i]}{1-\overline{f}_n} i^{k+1} \geq \sum_{i=K+1}^\infty \frac{h_n[i]}{1-\overline{f}_n} i^k K,$$

and hence, by (31) and the definition of K

$$\limsup_{n\to\infty} \sum_{i=K+1}^\infty \frac{h_n[i]}{1-\overline{f}_n} i^k \leq \varepsilon.$$

Therefore, by (32)

$$\mu_k = \lim_{n\to\infty} \sum_{i=1}^{\infty} \frac{h_n[i]}{1-\overline{f}_n} i^k \le \limsup_{n\to\infty} \sum_{i=1}^{K} \frac{h_n[i]}{1-\overline{f}_n} i^k + \varepsilon \le \sum_{i=1}^{\infty} q_i i^k + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary by equation (33)

$$\mu_k = \sum_{i=1}^{\infty} q_i i^k. \tag{34}$$

Using the Stieltjes moment problem we show that the sequence (μ_k) uniquely determines (q_i). In order to do that, it is enough to show that Carleman's condition ([18, Theorem 5.6.6.]) is fulfilled, that is, μ_k is not too large. Since $\limsup_{n\to\infty} \lambda_n^{1/n} \le 1$, for n large $\lambda_n \le 2^n$. Furthermore, by trivial upper bounds

$$\left\{ {k \atop i} \right\} i! = \sum_{\ell=0}^{i} (-1)^\ell \binom{i}{\ell} (i-\ell)^k \le i^k 2^i,$$

thus, by (31) for some $C > 0$

$$\mu_k = \sum_{i=0}^{k} \left\{ {k \atop i} \right\} i! \lambda_i \le C + \sum_{i=0}^{k} i^k 4^i \le C(4k)^k \le k^{2k},$$

for k large enough, showing that Carleman's condition holds. Hence the sequence $(q_i)_{i\in\mathbb{N}}$ is indeed unique, and the proof is complete. □

Lemma 9. *For any $L \ge 1$, $n \ge 1$, and $x \in \mathbb{R}$*

$$\sum_{j=1}^{L} \sum_{\ell=1}^{j} (-1)^{\ell+j} \binom{L+n}{j+n} \binom{j-\ell+n-1}{n-1} x^\ell = (1+x)^L - 1.$$

Proof. Changing the order of summation, it is enough to prove that for $L \ge 1$, $n \ge 1$, $1 \le \ell \le L$

$$\sum_{j=\ell}^{L} (-1)^{j+\ell} \binom{L+n}{j+n} \binom{j-\ell+n-1}{n-1} = \binom{L}{\ell}. \tag{35}$$

We prove (35) by induction on L. This holds for $L = 1$, and assuming for $L \ge 1$, for $L + 1$ we have by the induction hypothesis for (L, ℓ, n) and $(L, \ell - 1, n)$

$$\begin{aligned}
&\sum_{j=\ell}^{L+1} (-1)^{j+\ell} \binom{L+1+n}{j+n} \binom{j-\ell+n-1}{n-1} \\
&= \binom{L}{\ell} + \sum_{j=\ell}^{L+1} (-1)^{j+\ell} \binom{L+n}{j+n-1} \binom{j-\ell+n-1}{n-1} \\
&= \binom{L}{\ell} + \sum_{j=\ell-1}^{L} (-1)^{j+\ell-1} \binom{L+n}{j+n} \binom{j-(\ell-1)+n-1}{n-1} \\
&= \binom{L}{\ell} + \binom{L}{\ell-1} = \binom{L+1}{\ell},
\end{aligned}$$

as claimed. □

Proof of Theorem 4. In order to handle assumptions (i) and (ii) together, under (i) we use the notation $\lambda_k = 0$ for $k \geq K$. Then the limiting g.f. f_Y is given in (3).

First assume that $\nu > 0$. Substituting the Taylor series of the logarithm

$$\log\left(1+\frac{\nu}{2}(1-s)\right) = \log\left(1+\frac{\nu}{2}\right) - \sum_{n=1}^{\infty} n^{-1}\left(\frac{\nu}{2+\nu}\right)^n s^n, \quad s \in (0,1),$$

into (3), expanding $(1-s)^i$, and gathering together the powers of s, we have

$$\begin{aligned} f_Y(s) &= \exp\Bigg\{ -\sum_{k=1}^{\infty} \frac{2^k\lambda_k}{\nu^k}\Bigg(\log\left(1+\frac{\nu}{2}\right) - \sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{\nu}{2+\nu}\right)^n s^n \\ &\qquad + \sum_{i=1}^{k-1}(-1)^i \frac{\nu^i}{i2^i}\sum_{j=0}^{i}\binom{i}{j}(-s)^j\Bigg)\Bigg\} \\ &= \exp\Bigg\{ -\sum_{k=1}^{\infty}\frac{2^k\lambda_k}{\nu^k}\left(\log\left(1+\frac{\nu}{2}\right) + \sum_{i=1}^{k-1}(-1)^i\frac{\nu^i}{i2^i}\right) \\ &\qquad + \sum_{k=1}^{\infty}\frac{2^k\lambda_k}{\nu^k}\sum_{n=1}^{\infty} s^n\left(n^{-1}\left(\frac{\nu}{2+\nu}\right)^n - \sum_{i=n}^{k-1}(-1)^{i+n}\frac{\nu^i}{i2^i}\binom{i}{n}\right)\Bigg\}, \end{aligned} \tag{36}$$

where the empty sum is defined to be 0. We claim that the order of summation in (36) with respect to k and n can be interchanged. Indeed, this is clear under (i), since the sum in k is in fact finite.

Assume (ii). Then Taylor's theorem applied to the function $(1+x)^{-n}$, with $n \leq k-1$ gives

$$(1+x)^{-n} = n\sum_{\ell=n}^{k-1}(-1)^{\ell+n}\binom{\ell}{n}\frac{x^{\ell-n}}{\ell} + (-1)^{k-n}\binom{k-1}{k-n}(1+\xi)^{-k}x^{k-n},$$

with $\xi \in (0,x)$. Therefore,

$$\left|\sum_{i=n}^{k-1}(-1)^{i+n}\binom{i}{n}\frac{x^{i-n}}{i} - n^{-1}(1+x)^{-n}\right| \leq n^{-1}\binom{k-1}{k-n}x^{k-n}, \tag{37}$$

thus with $x = \nu/2$

$$\begin{aligned} &\sum_{k=1}^{\infty}\frac{2^k\lambda_k}{\nu^k}\sum_{n=1}^{k-1}s^n\left|\sum_{i=n}^{k-1}(-1)^{i+n}\binom{i}{n}\frac{\nu^i}{2^i i} - n^{-1}\left(\frac{\nu}{2+\nu}\right)^n\right| \\ &\leq \sum_{k=1}^{\infty}\frac{2^k\lambda_k}{\nu^k}\sum_{n=1}^{k-1}s^n\frac{\nu^n}{2^n}n^{-1}\binom{k-1}{k-n}\frac{\nu^{k-n}}{2^{k-n}} \\ &= \sum_{k=1}^{\infty}\lambda_k\sum_{n=1}^{k-1}s^n n^{-1}\binom{k-1}{k-n} \\ &\leq \sum_{k=1}^{\infty}\lambda_k(1+s)^k, \end{aligned}$$

which is finite for $s \in (0,1)$ if $\limsup \lambda_n^{1/n} \le 1/2$. The other part is easy to handle, as

$$\begin{aligned}\sum_{k=1}^{\infty} \frac{2^k \lambda_k}{\nu^k} \sum_{n=k}^{\infty} s^n n^{-1} \left(\frac{\nu}{2+\nu}\right)^n &\le \sum_{k=1}^{\infty} \frac{2^k \lambda_k}{\nu^k} s^k \left(\frac{\nu}{2+\nu}\right)^k \\ &= \sum_{k=1}^{\infty} \frac{2^k \lambda_k}{(2+\nu)^k} s^k,\end{aligned}$$

which is summable.

Summarizing, in both cases the order of summation in (36) can be interchanged, and doing so we obtain

$$f_Y(s) = \exp\left\{ -A_0 + \sum_{n=1}^{\infty} A_n s^n \right\},$$

where

$$\begin{aligned} A_0 &= \sum_{k=1}^{\infty} \frac{2^k \lambda_k}{\nu^k} \left(\sum_{i=1}^{k-1} (-1)^i \frac{\nu^i}{i 2^i} + \log\left(1 + \frac{\nu}{2}\right) \right) \\ A_n &= \sum_{k=1}^{\infty} \frac{2^k \lambda_k}{\nu^k} \left(\left(\frac{\nu}{2+\nu} \right)^n n^{-1} - \sum_{i=n}^{k-1} (-1)^{i+n} \binom{i}{n} \frac{\nu^i}{i 2^i} \right). \end{aligned} \tag{38}$$

Recall the notation from the proof of Lemma 8. By (31), using the inversion formula for Stirling numbers of the first and second kind we have, see e.g. [10, Exercise 12, p. 310],

$$k! \lambda_k = \sum_{i=0}^{k} (-1)^{k+i} \begin{bmatrix} k \\ i \end{bmatrix} \mu_k,$$

where $\begin{bmatrix} k \\ i \end{bmatrix}$ stands for the Stirling number of the first kind. Substituting (34), and using that

$$\sum_{i=0}^{k} (-1)^{k+i} \begin{bmatrix} k \\ i \end{bmatrix} j^k = j(j-1)\cdots(j-k+1)$$

see, e.g. [10, p.263, (6.13)], we obtain the formula

$$\lambda_k = \sum_{j=k}^{\infty} \binom{j}{k} q_j. \tag{39}$$

We also see that $\lambda_k = 0$ implies $q_j = 0$ for all $j \ge k$. Substituting (39) back into (38) we claim that the order of summation with respect to k and j can be interchanged.

This is again clear under (i), while under (ii), by (37)

$$\begin{aligned}
&\sum_{k=n+1}^{\infty} \sum_{j=k}^{\infty} \frac{2^k}{\nu^k} q_j \binom{j}{k} \left| \sum_{i=n}^{k-1} (-1)^{i+n} \binom{i}{n} \frac{\nu^i}{i 2^i} - n^{-1} \frac{\nu^n}{(2+\nu)^n} \right| \\
&\leq \sum_{k=n+1}^{\infty} \sum_{j=k}^{\infty} \frac{2^k}{\nu^k} q_j \binom{j}{k} \frac{\nu^n}{2^n} n^{-1} \binom{k-1}{k-n} \frac{\nu^{k-n}}{2^{k-n}} \\
&= \sum_{k=n+1}^{\infty} \lambda_k n^{-1} \binom{k-1}{k-n} < \infty,
\end{aligned}$$

since

$$n^{-1} \binom{k-1}{k-n} = k^{-1} \binom{k}{n} \leq k^n.$$

Thus the order of summation is indeed interchangeable, and we obtain

$$\begin{aligned}
A_n = \sum_{j=1}^{\infty} q_j \Bigg\{ &\left[\left(1 + \frac{2}{\nu}\right)^j - 1 \right] \left(\frac{\nu}{2+\nu} \right)^n n^{-1} \\
&- \sum_{k=n+1}^{j} \frac{2^k}{\nu^k} \binom{j}{k} \sum_{i=n}^{k-1} (-1)^{i+n} \binom{i}{n} \frac{\nu^i}{i 2^i} \Bigg\} \\
=: \sum_{j=1}^{\infty} &q_j B(n,j).
\end{aligned} \tag{40}$$

We claim that

$$B(n,j) = \frac{\nu^n}{(2+\nu)^n n} \left[\left(1 + \frac{2}{\nu}\right)^{n \wedge j} - 1 \right]. \tag{41}$$

This is clear for $j \leq n$. Writing $\ell = k - i$ in the summation and using Lemma 9 with $L = j - n$ we obtain

$$\begin{aligned}
&\sum_{k=n+1}^{j} \sum_{i=n}^{k-1} (-1)^i \binom{j}{k} \binom{i-1}{n-1} x^{k-i} \\
&= \sum_{k=n+1}^{j} \sum_{\ell=1}^{k-n} (-1)^{k+\ell} \binom{j}{k} \binom{k-\ell-1}{n-1} x^\ell \\
&= (-1)^n \left[(1+x)^{j-n} - 1 \right].
\end{aligned}$$

Using that $n\binom{i}{n} = i\binom{i-1}{n-1}$, and substituting back into (40) with $x = 2/\nu$ we have

$$B(n,j) = n^{-1} \left[1 - \left(1 + \frac{2}{\nu}\right)^{j-n} \right] + \left[\left(1 + \frac{2}{\nu}\right)^j - 1 \right] \left(\frac{\nu}{2+\nu} \right)^n n^{-1},$$

which after simplification gives (41).

Now let $\nu = 0$. In this case the calculations are again simpler. Expanding $(s-1)^i$ and changing the order of summation

$$\log f_Y(s) = \sum_{k=1}^{\infty} (-1)^k \frac{\lambda_k}{k} + \sum_{n=1}^{\infty} s^n \sum_{k=n}^{\infty} (-1)^{n+k} \frac{\lambda_k}{k} \binom{k}{n} =: -A_0 + \sum_{n=1}^{\infty} s^n A_n.$$

Substituting (39) and changing the order of summation again we get

$$A_n = \sum_{j=n}^{\infty} q_j \sum_{k=n}^{j} \binom{k}{n} \binom{j}{k} (-1)^{k+n} \frac{1}{k}. \tag{42}$$

The changes of the order of summation can be justified as in the previous case. For the coefficient of q_j we have

$$\sum_{k=n}^{j} \binom{k}{n} \binom{j}{k} (-1)^{k+n} \frac{1}{k} = \binom{j}{n} \sum_{k=n}^{j} \frac{1}{k} \binom{j-n}{k-n} (-1)^{k+n} = \frac{1}{n},$$

where we used that for $j \geq n$

$$\sum_{k=0}^{j-n} (-1)^k \frac{j}{k+n} \binom{j-n}{k} = \frac{1}{\binom{j-1}{n-1}}.$$

The latter formula follows by induction on j. Substituting back into (42), the proof is complete. $\square$

Acknowledgement

We are grateful to Mátyás Barczy for discussions on the topic. We thank the anonymous referees for useful comments and remarks, and in particular for suggesting to allow supercritical generations, which lead to more general results.

Funding information

This research was supported by the Ministry of Innovation and Technology of Hungary from the National Research, Development and Innovation Fund, project no. TKP2021-NVA-09. Kevei's research was partially supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] Athreya, K. B. and Ney, P. E. (1972). *Branching processes.* Die Grundlehren der mathematischen Wissenschaften, Band 196. Springer-Verlag, New York-Heidelberg.

[2] Bhattacharya, N. and Perlman, M. (2017). Time-inhomogeneous branching processes conditioned on non-extinction. *arXiv:1703.00337*.

[3] Cardona-Tobón, N. and Palau, S. (2021). Yaglom's limit for critical Galton-Watson processes in varying environment: a probabilistic approach. *Bernoulli* **27,** 1643–1665.

[4] Church, J. D. (1971). On infinite composition products of probability generating functions. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **19,** 243–256.

[5] Dolgopyat, D., Hebbar, P., Koralov, L. and Perlman, M. (2018). Multi-type branching processes with time-dependent branching rates. *J. Appl. Probab.* **55,** 701–727.

[6] Fearn, D. H. (1972). Galton-Watson processes with generation dependence. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. IV: Biology and health.* Univ. California Press, Berkeley, Calif. pp. 159–172.

[7] Feller, W. (1968). *An introduction to probability theory and its applications. Vol. I* third ed. John Wiley & Sons, Inc., New York-London-Sydney.

[8] Gao, Z. and Zhang, Y. (2015). Limit theorems for a Galton-Watson process with immigration in varying environments. *Bull. Malays. Math. Sci. Soc.* **38,** 1551–1573.

[9] González, M., Kersting, G., Minuesa, C. and del Puerto, I. (2019). Branching processes in varying environment with generation-dependent immigration. *Stoch. Models* **35,** 148–166.

[10] Graham, R. L., Knuth, D. E. and Patashnik, O. (1994). *Concrete mathematics* second ed. Addison-Wesley Publishing Company, Reading, MA. A foundation for computer science.

[11] Györfi, L., Ispány, M., Kevei, P. and Pap, G. (2015). Asymptotic behavior of multitype nearly critical Galton-Watson processes with immigration. *Theory Probab. Appl.* **59,** 590–610.

[12] Györfi, L., Ispány, M., Pap, G. and Varga, K. (2007). Poisson limit of an inhomogeneous nearly critical INAR(1) model. *Acta Sci. Math. (Szeged)* **73,** 789–815.

[13] Ispány, M. (2016). Some asymptotic results for strongly critical branching processes with immigration in varying environment. In *Branching processes and their applications.* vol. 219 of *Lect. Notes Stat.* Springer, [Cham] pp. 77–95.

[14] Jagers, P. (1974). Galton-Watson processes in varying environments. *J. Appl. Probab.* **11,** 174–178.

[15] Kersting, G. (2020). A unifying approach to branching processes in a varying environment. *J. Appl. Probab.* **57,** 196–220.

[16] Kersting, G. and Vatutin, V. (2017). *Discrete Time Branching Processes in Random Environment.* ISTE Ltd and John Wiley & Sons, Inc.

[17] Kevei, P. (2011). Asymptotics of nearly critical Galton-Watson processes with immigration. *Acta Sci. Math. (Szeged)* **77,** 681–702.

[18] Simon, B. (2015). *Real analysis.* A Comprehensive Course in Analysis, Part 1. American Mathematical Society, Providence, RI.

[19] Steutel, F. W. and van Harn, K. (2004). *Infinite divisibility of probability distributions on the real line* vol. 259 of *Monographs and Textbooks in Pure and Applied Mathematics.* Marcel Dekker, Inc., New York.

[20] Weiss, C. H. (2008). Thinning operations for modeling time series of counts—a survey. *AStA Adv. Stat. Anal.* **92,** 319–341.

[21] Yaglom, A. M. (1947). Certain limit theorems of the theory of branching random processes. *Doklady Akad. Nauk SSSR (N.S.)* **56,** 795–798.