Part I
The simple Galton-Watson process: Classical approach

Chapter 1
Basic definitions and first results

1.1 The model

The simplest version of a population size model which sets out at the level of individuals may be informally described as follows: All individuals behave independently and have equally distributed random lifetimes and random numbers of children (offspring). In other words, if an individual v has lifetime T_v and produces N_v children, then the pairs $(T_\mathsf{v}, N_\mathsf{v})$ are drawn from a family of iid random variables. Population size may be measured either in real time, that is, by counting the number of living individuals at any time $t \in \mathbb{R}_\geq$, or over generations $n \in \mathbb{N}_0$, that is by counting the numbers of individuals with the same number of ancestors. In the following, the latter, genealogical perspective is adopted, which does not care about lifetimes and leads to the most classical model in theory of branching processes

Definition 1.1. Let $(Z_n)_{n\geq 0}$ be a sequence of integer-valued random variables, recursively defined by

$$Z_n = \sum_{k=1}^{Z_{n-1}} X_{n,k}, \quad n \geq 1, \tag{1.1}$$

where $\{X_{n,k} : n, k \geq 1\}$ forms a family of iid integer-valued random variables with common distribution $(p_n)_{n\geq 0}$ and independent of Z_0. For each $n \geq 0$, the random variable Z_n is interpreted as the size of the n^{th} generation of a given population. Then $(Z_n)_{n\geq 0}$ is called a *simple Galton-Watson process* or just *Galton-Watson process (GWP)* with *offspring distribution* $(p_n)_{n\geq 0}$ and Z_0 ancestors.

As one can easily see, the distribution of $(Z_n)_{n\geq 0}$ is completely determined by two input parameters, the offspring distribution $(p_n)_{n\geq 0}$ and the (ancestral) distribution of Z_0. In fact, $(Z_n)_{n\geq 0}$ constitutes a temporally homogeneous Markov chain on $\mathbb{N}_0$ with transition matrix $(p_{ij})_{i,j\geq 0}$, given by

$$p_{ij} \;=\; \mathbb{P}(Z_{n+1}=j|Z_n=i) \;=\; p_j^{*i}, \tag{1.2}$$

where $(p_n^{*i})_{n\ge 0}$ denotes the i-fold convolution of $(p_n)_{n\ge 0}$, which equals the Dirac measure at 0 in the case $i=0$. In other words, the simple GWP may also be characterized as a Markov chain on the nonnegative integers with the special transition structure specified by (1.2).

As common in the theory of Markov chains, we may study $(Z_n)_{n\ge 0}$ within the following *standard model*: Let Z_0 and all $X_{n,k}$ be defined on a measurable space $(\Omega,\mathfrak{A})$ together with a family of probability measures $\{\mathbb{P}_i : i\in\mathbb{N}_0\}$ such that the $X_{n,k}$ are iid with common distribution $(p_n)_{n\ge 0}$ under each $\mathbb{P}_i$, while $\mathbb{P}_i(Z_0=i)=1$. In view of the recursive definition (1.1) it is useful to define the filtration

$$\mathscr{F}_0 := \sigma(Z_0) \quad \text{and} \quad \mathscr{F}_n := \sigma(Z_0,\{X_{j,k} : 1\le j\le n, k\ge 1\}) \quad \text{for } n\ge 1$$

and to note that $(Z_n)_{n\ge 0}$ is Markov-adapted to $(\mathscr{F}_n)_{n\ge 0}$, that is, each Z_n is $\mathscr{F}_n$-measurable and

$$\mathbb{P}(Z_{n+1}=j|\mathscr{F}_n) \;=\; \mathbb{P}(Z_{n+1}=j|Z_n) \quad \text{a.s.}$$

where ”a.s.” means ”$\mathbb{P}_i$-a.s. for all $i\in\mathbb{N}_0$”.

Due to the independence assumptions, it is intuitively clear and confirmed by the next lemma that a GWP with i ancestors is just the sum of i independent GWP’s with one ancestor. This worth to be pointed out because it will allow us in many places to reduce our further analysis to the case where $Z_0=1$.

Lemma 1.2. *Every GWP $(Z_n)_{n\ge 0}$ with k ancestors is the sum of k independent GWP’s $(Z_n^{(i)})_{n\ge 0}$, $1\le i\le k$, with one ancestor and the same offspring distribution. If $(Z_n)_{n\ge 0}$ is given in a standard model, this may also be stated as*

$$\mathbb{P}_k((Z_n)_{n\ge 0}\in\cdot) \;=\; \mathbb{P}_1((Z_n)_{n\ge 0}\in\cdot)^{*k}$$

for all $k\ge 0$.

1.2 The model behind: genealogical structure and random trees

Behind any GWP with one ancestor there is a more informative model describing the genealogy of the considered population in terms of a suitable random tree, called *Galton-Watson tree*. This is done by assigning a suitable label to each individual that reflects the ancestral structure and by interpreting each label as a node in the tree. A common way is via embedding in the so-called *Ulam-Harris tree* with vertex set

$$\mathbb{V} \;=\; \bigcup_{n\ge 0}\mathbb{N}^n,$$

where $\mathbb{N}^0 := \{\varnothing\}$ consists of the root. Each vertex $\mathsf{v} = (\mathsf{v}_1, ..., \mathsf{v}_n) \in \mathbb{V} \setminus \{\varnothing\}$, shortly written as $\mathsf{v}_1...\mathsf{v}_n$ hereafter, is connected to the root via the unique shortest path

$$\varnothing \to \mathsf{v}_1 \to \mathsf{v}_1\mathsf{v}_2 \to ... \to \mathsf{v}_1...\mathsf{v}_n. \tag{1.3}$$

Let $|\mathsf{v}|$ denote the length (generation) of v, thus $|\mathsf{v}_1...\mathsf{v}_n| = n$ and particularly $|\varnothing| = 0$. Further let $\mathsf{uv} = \mathsf{u}_1...\mathsf{u}_m\mathsf{v}_1...\mathsf{v}_n$ denote the concatenation of two vectors $\mathsf{u} = \mathsf{u}_1...\mathsf{u}_m$ and $\mathsf{v} = \mathsf{v}_1...\mathsf{v}_n$.

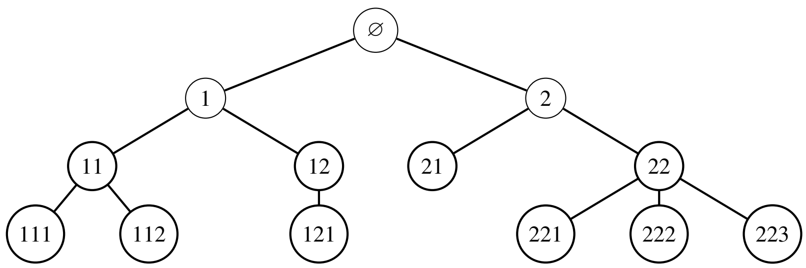

Fig. 1.1 A Galton-Watson tree with Ulam-Harris labeling.

The *Galton-Watson tree* $\boldsymbol{GW}$ associated with $(Z_n)_{n\ge 0}$ may now be defined in the following inductive manner as a random subtree of $\mathbb{V}$: First, the ancestor gets the label $\varnothing$. Second, given that an individual of the population has been labeled by v and gives birth to X_v children, label its offspring by $\mathsf{v}1, ..., \mathsf{v}X_\mathsf{v}$ [☞ Figure 1.1]. Then it is clear that any individual of generation n gets a label $\mathsf{v} = \mathsf{v}_1...\mathsf{v}_n$ with $|\mathsf{v}| = n$, thus

$$Z_n = |\{\mathsf{v} \in \boldsymbol{GW} : |\mathsf{v}| = n\}|,$$

and that the shortest path (1.3) from the root $\varnothing$ to v provides us with the ancestral line of v where $\mathsf{v}_1...\mathsf{v}_{n-1}$ denotes its mother, $\mathsf{v}_1...\mathsf{v}_{n-2}$ its grandmother, etc. The offspring numbers X_v are iid with common distribution $(p_n)_{n\ge 0}$ and Z_n may also be written as

$$Z_n = \sum_{\mathsf{v}\in\boldsymbol{GW}:|\mathsf{v}|=n-1} X_\mathsf{v}$$

for each $n \ge 1$.

We refrain at this point from a precise formalization of random trees as random elements on a suitable probability space and refer instead to Chapter 4.

1.3 Generating functions and moments

We have already pointed out that, apart from the number of ancestors, the distribution of a GWP $(Z_n)_{n\ge 0}$ is completely determined by its offspring distribution $(p_n)_{n\ge 0}$, which in turn is uniquely determined by its *generating function (gf)*

$$f(s) \;=\; \sum_{n\geq 0} p_n s^n, \quad s \in [-1,1].$$

As we will see, many of the subsequent results are based on a thorough analysis of gf's. Their most important properties are therefore collected in Appendix A.1.

In the following, let $(Z_n)_{n\geq 0}$ be given in a standard model in which expectation and variance under $\mathbb{P}_i$ are denoted as $\mathbb{E}_i$ and $\mathbb{V}\mathrm{ar}_i$, respectively. Instead of $\mathbb{P}_1, \mathbb{E}_1$ and $\mathbb{V}\mathrm{ar}_1$ we also write $\mathbb{P}, \mathbb{E}$ and $\mathbb{V}\mathrm{ar}$, respectively. Obviously,

$$f(s) \;=\; \mathbb{E}_i s^{X_{n,k}}$$

for all $i \in \mathbb{N}_0$ and $k,n \in \mathbb{N}$, in particular $f(s) = \mathbb{E}s^{Z_1}$.

Our first task is to determine, for each $n \in \mathbb{N}_0$, the gf of Z_n under $\mathbb{P}$ as a function of f. This will be accomplished by the next lemma and also provide us with $\mathbb{E}Z_n$ and $\mathbb{V}\mathrm{ar}\, Z_n$ via differentiation.

Lemma 1.3. *Let $T, X_1, X_2, \ldots$ be independent, integer-valued random variables such that $X_1, X_2, \ldots$ are further identically distributed with common gf f, mean μ and variance σ^2 (if it exists). Let g be the gf of T. Then*

$$Y \;=\; \sum_{k=1}^{T} X_j \quad [:= 0 \text{ on } \{T = 0\}]$$

has gf $\mathbb{E}s^Y = g \circ f(s)$ and mean

$$\mathbb{E}Y \;=\; \mu\, \mathbb{E}T. \tag{1.4}$$

Furthermore, if μ and $\mathbb{E}T$ are both finite (thus σ^2 exists), then

$$\mathbb{V}\mathrm{ar}\, Y \;=\; \sigma^2\, \mathbb{E}T + \mu^2\, \mathbb{V}\mathrm{ar}\, T. \tag{1.5}$$

Proof. Under the stated conditions and by making use of the multiplication lemma for gf's, we infer for each $s \in [-1,1]$

$$\begin{aligned}
\mathbb{E}s^Y \;&=\; \mathbb{P}(T=0) + \sum_{k\geq 1} \int_{\{T=k\}} s^{X_1+\ldots+X_k}\, d\mathbb{P} \\
&=\; \mathbb{P}(T=0) + \sum_{k\geq 1} \mathbb{P}(T=k)\, \mathbb{E}s^{X_1+\ldots+X_k} \\
&=\; \sum_{k\geq 0} \mathbb{P}(T=k)\, f(s)^k \\
&=\; \mathbb{E}f(s)^T \\
&=\; g(f(s)),
\end{aligned}$$

which is the first assertion. The remaining ones (1.4) and (1.5) are now easily obtained via differentiation of $g(f(s))$ and by letting $s \uparrow 1$. The details are left to the reader. $\square$

The gf of Z_n as well as its mean and its variance are now directly obtained by means of Lemma 1.3. Let

$$\mathsf{m} \;=\; \mathbb{E}Z_1 \;=\; \sum_{k\ge 1} kp_k \;=\; f'(1)$$

denote the expected number of children per individual, called *offspring mean* or *reproduction mean*. It always exists, but may be infinite. If m is finite, we may further define

$$\sigma^2 \;=\; \mathbb{V}\mathrm{ar}\, Z_1 \;=\; \sum_{k\ge 1} k^2 p_k - \mathsf{m}^2 \;=\; f''(1) + f'(1)(1-f'(1)), \tag{1.6}$$

called *offspring* or *reproduction variance*.

Proposition 1.4. *For each $n \in \mathbb{N}$, the gf f_n of Z_n under $\mathbb{P}$ is given by the n-fold composition of f, i.e.*

$$f_n \;=\; f^{\circ n} \;:=\; f \circ \ldots \circ f. \tag{1.7}$$

Moreover, Z_n has mean $\mathbb{E}Z_n = \mathsf{m}^n$ and variance

$$\mathbb{V}\mathrm{ar}\, Z_n \;=\; \begin{cases} \dfrac{\sigma^2 \mathsf{m}^{n-1}(\mathsf{m}^n - 1)}{\mathsf{m}-1}, & \text{if } \mathsf{m} \ne 1, \\ n\sigma^2, & \text{if } \mathsf{m} = 1, \end{cases} \tag{1.8}$$

provided that m is finite.

Proof. Since $X_{n,1}, X_{n,2}, \ldots$ are iid with common gf f and independent of Z_{n-1} and since $Z_n = \sum_{k=1}^{Z_{n-1}} X_{n,k}$, we infer from Lemma 1.3 that $f_n = f_{n-1} \circ f$ for each $n \in \mathbb{N}$ and thereupon (1.7). The same lemma further gives $\mathbb{E}Z_n = \mathsf{m}\,\mathbb{E}Z_{n-1}$ for each n and thus $\mathbb{E}Z_n = \mathsf{m}^n$. Assuming $\mathsf{m} < \infty$ and defining $W_n = \mathsf{m}^{-n} Z_n$, a straightforward calculation yields

$$\mathbb{V}\mathrm{ar}\, W_n \;=\; \frac{\sigma^2}{\mathsf{m}^{n+1}} + \mathbb{V}\mathrm{ar}\, W_{n-1} \tag{1.9}$$

for each $n \in \mathbb{N}$. To derive (1.8) from this recursion is left as an exercise. $\square$

Problems

Problem 1.5. Complete the proof of Prop. 1.4 by showing (1.9) and (1.8). Give also a closed form expression for $\mathbb{V}\mathrm{ar}\, W_n$.

Problem 1.6. Show that Lemma 1.3 may be generalized as follows: If ceteris paribus $X_1, X_2, \ldots$ take values in $\mathbb{R}_\geq$ and have Laplace transform (LT) φ, then the LT ψ of Y satisfies $\psi = g \circ \varphi$.

1.4 Back to the genesis: the extinction probability

We turn to the problem that leads back to the early days of branching processes, viz. the computation of the probability that a population described by a GWP eventually dies out. The solution presented here is essentially the historical one given by Watson and based on gf's.

So let $(Z_n)_{n\geq 0}$ be a GWP in a standard model with offspring distribution $(p_n)_{n\geq 0}$ having gf f. The extinction probability under $\mathbb{P}_i$ is denoted as $q(i)$, in the case $i = 1$ also as q. Since the event of extinction is an absorbing event, that is, $Z_n = 0$ implies $Z_{n+k} = 0$ for all $k \geq 1$, we infer

$$q(i) \;=\; \mathbb{P}_i\left(\bigcup_{k\geq 0}\{Z_k = 0\}\right) \;=\; \lim_{n\to\infty} \mathbb{P}_i\left(\bigcup_{k=1}^{n}\{Z_k = 0\}\right) \;=\; \lim_{n\to\infty} \mathbb{P}_i(Z_n = 0). \quad (1.10)$$

Furthermore, by Lemma 1.2, we have

$$\mathbb{P}_i(Z_n = 0) \;=\; \mathbb{P}_i(Z_n^{(j)} = 0 \text{ for each } 1 \leq j \leq i) \;=\; \mathbb{P}(Z_n = 0)^i$$

for all $i, n \geq 1$ and thus $q(i) = q^i$. Therefore it suffices to find q, the extinction probability of $(Z_n)_{n\geq 0}$ under $\mathbb{P} = \mathbb{P}_1$.

Because of the trivial implications

$$\begin{aligned} p_0 = 0 \quad &\Longrightarrow \quad q = 0, \\ p_0 + p_1 = 1,\ 0 < p_0 < 1 \quad &\Longrightarrow \quad q = 1, \end{aligned}$$

we may confine ourselves hereafter to the case where

$$0 < p_0 \leq p_0 + p_1 < 1. \quad (1.11)$$

Keeping the notation of the previous section, we are now able to prove the following result.

Theorem 1.7. *The extinction probability q is given as the smallest fixed point of f in $[0,1]$, i.e., the smallest solution to the equation $f(s) = s$ in this interval. Provided that (1.11), f possesses exactly one fixed point in $[0,1)$ if $\mathsf{m} > 1$, and none if $\mathsf{m} \leq 1$. Hence $p_0 < q < 1$ in the first case and $q = 1$ in the second one.*

So we have the intuitive result that a nontrivial GWP (i.e. $p_1 \neq 1$) has a positive chance of survival iff the mean number of offspring per individual is bigger than one.

Proof. By (1.10), $q = \lim_{n\to\infty} \mathbb{P}(Z_n = 0) = \lim_{n\to\infty} f_n(0)$. Since f is continuous on $[0,1]$ and $f_{n+1} = f \circ f_n$ for all $n \geq 1$, it follows that

$$f(q) \;=\; \lim_{n\to\infty} f \circ f_n(0) \;=\; \lim_{n\to\infty} f_{n+1}(0) \;=\; q,$$

i.e., the fixed point property of q. Let $a \in [0,1]$ be an arbitrary fixed point of f. Using the monotonicity of all f_n on $[0,1]$, we obtain

$$a \;=\; f(a) \;=\; f_n(a) \;\geq\; f_n(0)$$

for all $n \geq 1$ and therefore $a \geq \lim_{n\to\infty} f_n(0) = q$. So we have shown that q is indeed the smallest fixed point of f in $[0,1]$.

From now on assume (1.11) in which case f must be strictly increasing and strictly convex on $[0,1]$. We define $g(s) = f(s) - s$ and note that $g(0) = p_0 > 0$ and $g(1) = 0$.

Since $f'(s) < f'(1) = \mathsf{m}$ for all $s \in [0,1)$, we have $g'(s) = f'(s) - 1 < 0$ for all $s \in [0,1)$ if $\mathsf{m} \leq 1$.

Now suppose that $\mathsf{m} > 1$ in which case $f(s)$ grows faster than s in a left neighborhood of 1, for $g'(1) = \mathsf{m} - 1 > 0$ and g' is continuous. On the other hand, $f(s) > s$ hold strue in a right neighborhood of 0, for $g(0) > 0$ and g is continuous. By an appeal to the intermediate value theorem, there exists at least one fixed point s_1 of f in $(0,1)$. If there were a second one s_2, wlog $s_1 < s_2 < 1$, we would infer $g(s_1) = g(s_2) = g(1) = 0$ and thereupon the existence of $a,b \in (0,1)$, $s_1 < a < s_2 < b < 1$, with $g'(a) = g'(b) = 0$ or, equivalently, $f'(a) = f'(b)$ (Rolle's theorem), a contradiction to the strict convexity of f. Hence q is the unique fixed point of f in $(0,1)$ and further satisfies $q > p_0$ because

$$q \;\geq\; \mathbb{P}(Z_1 = 0) + \mathbb{P}(Z_1 > 0, Z_2 = 0) \;\geq\; p_0 + p_n p_0^n$$

for any $n \in \mathbb{N}$ so large that $p_n > 0$. This completes the proof of the theorem. □

Figure 1.2 illustrates the situation in each of the cases $\mathsf{m} < 1$, $\mathsf{m} = 1$ and $\mathsf{m} > 1$. In the last case, it seems natural to find q by solving the equation $f(s) = s$ if f is explicitly known. Unfortunately, this can only be done in a few special cases two of which are considered in Problems 1.10, 1.12 and the next section. On the other hand, approximative computation of q via iteration of f provides us with an alternative that is easy and even fast, as shown by the following corollary and illustrated in Figure 1.3.

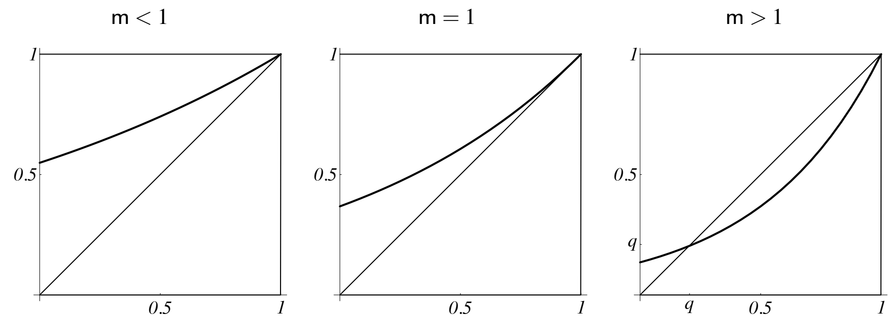

Fig. 1.2 The generating function f in the cases $\mathsf{m} <,=,> 1$.

Corollary 1.8. *Suppose that* (1.11) *holds true. Then* $f_n(s) \nearrow q$ *uniformly for* $s \in [0,q]$ *and* $f_n(s) \searrow q$ *for* $s \in [q,1)$. *The last convergence is uniform in* s *on any compact subset of* $[q,1)$. *Furthermore,*

$$0 \ < \ q - f_n(s) \ \leq \ f'(q)^n \tag{1.12}$$

for all $n \in \mathbb{N}$ *and* $s \in [0,q)$. *If* $\mathsf{m} \neq 1$, *then* $f'(q) < 1$.

Proof. Since f is increasing on $[0,1]$, $s < f(s) < f(q) = q$ holds true for $s \in [0,q)$. Now iterate to obtain

$$s < f_1(s) < f_2(s) < \ldots < f_n(s) < f_n(q) = q$$

for all $n \geq 1$ and thereby $f_n(s) \nearrow \hat{q} \leq q$. But $f_n(0) \leq f_n(s)$ together with $f_n(0) = \mathbb{P}(Z_n = 0) \nearrow q$ ensures $\hat{q} = q$ as well as the asserted uniform convergence on $[0,q]$.

If $s \in [q,1)$, thus $\mathsf{m} > 1$, a similar argument shows that $f_n(s) \searrow \hat{q} \in [q,1)$. But $\hat{q}$ must be a fixed point of f and thus $\hat{q} = q$ by Theorem 1.7, because

$$\hat{q} \ = \ \lim_{n\to\infty} f_{n+1}(s) \ = \ f\left(\lim_{n\to\infty} f_n(s)\right) \ = \ f(\hat{q}).$$

Furthermore, $0 \leq f_n(s) - q \leq f_n(t) - q$ for all $n \geq 1$, $t \in (q,1)$ and $s \in [q,t]$ implies that the convergence is uniform on any compact subset of $[q,1)$.

Finally, the convexity of f implies that

$$0 \ < \ \frac{q - f(s)}{q - s} \ \leq \ f'(q)$$

for each $s \in [0,q)$. Hence, we infer (1.12) from

$$0 \ < \ \frac{q - f_n(s)}{q - s} \ = \prod_{k=0}^{n-1} \frac{q - f(f_k(s))}{q - f_k(s)} \ \leq \ f'(q)^n$$

for all $n \geq 1$ and $s \in [0,q)$, where $f_0(s) := s$. Cleary, $f'(q) = \mathsf{m} < 1$ in the subcritical case. But if $\mathsf{m} > 1$ and thus $0 \leq q < 1$, we have

$$f'(q) \;<\; \frac{f(1)-f(q)}{1-q} \;=\; 1$$

by the strict convexity of f. $\square$

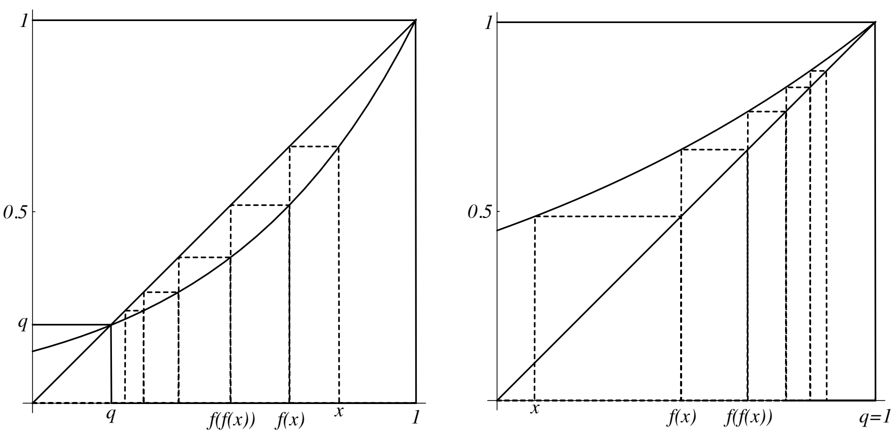

Fig. 1.3 Approximation of q via the iteration sequence $f_n(x)$, $n \in \mathbb{N}_0$.

The last result of this section shows a fundamental dichotomy for GWP's, which is actually typical for any branching process with independent reproduction and known as the *extinction-explosion principle*.

Theorem 1.9. [Extinction-explosion principle] *Every GWP $(Z_n)_{n\geq 0}$ satisfying $\mathbb{P}(Z_1 = 1) \neq 1$ is bound to either extinction or explosion, i.e.*

$$\mathbb{P}(Z_n = 0 \text{ eventually}) \;+\; \mathbb{P}(Z_n \to \infty) \;=\; 1.$$

Proof. An equivalent formulation of the extinction-explosion principle in the framework of Markov chains is that 0 is an absorbing state while all other states $k \in \mathbb{N}$ are transient. As before, let $(p_n)_{n\geq 0}$ be the offspring distribution of $(Z_n)_{n\geq 0}$, thus $p_1 \neq 1$. For any $k \in \mathbb{N}$, we then infer

$$\mathbb{P}_k(Z_n \neq k \text{ for all } n \geq 1) \;\geq\; \left\{\begin{array}{rl} p_0^k, & \text{if } p_0 > 0 \\ 1-p_1^k, & \text{if } p_0 = 0 \end{array}\right\} \;>\; 0$$

which proves the asserted transience of k. $\square$

The extinction-explosion principle naturally calls for a more detailed description of the asymptotic behavior of Z_n on the two disjoint events of extinction and explosion. As it turns out, this leads to quite different results for the three cases $\mathsf{m} < 1$, $\mathsf{m} = 1$ and $\mathsf{m} > 1$. Due to this fact, $(Z_n)_{n\ge 0}$ is called

- *subcritical*, if $\mathsf{m} < 1$,
- *critical*, if $\mathsf{m} = 1$,
- *supercritical*, if $\mathsf{m} > 1$.

For each of the three cases, Figure 1.4 shows 100 simulated trajectories of a GWP with Poissonian offspring distribution and a glance suffices to see that they exhibit a quite different behavior.

Problems

The notation of the previous sections is kept. In particular, $(Z_n)_{n\ge 0}$ always denotes a GWP with one ancestor, offspring distribution $(p_n)_{n\ge 0}$ having gf f and mean m, and extinction probability q.

Problem 1.10 (Cell splitting). Find the extinction probability of a *cell splitting process*, described by a GWP $(Z_n)_{n\ge 0}$ with offspring distribution of the simple form

$$p_0, p_2 > 0, \quad p_1 \in [0,1) \quad \text{and} \quad p_n = 0 \text{ otherwise.}$$

In the case $p_1 = 0$, one may think of a cell population whose members either split or die at the end of their life cycles.

Problem 1.11. Show that

$$\frac{p_0}{1-p_1} \;\le\; q \;\le\; \frac{p_0}{1-p_0-p_1} \tag{1.13}$$

holds true under the convention that $\frac{0}{0} := 1$.

Problem 1.12. [This is an example of a linear fractional distribution to be discussed in full generality in the next section] Find the extinction probability of a supercritical GWP with geometric offspring distribution, that is

$$p_n \;=\; \frac{1}{\mathsf{m}+1}\left(\frac{\mathsf{m}}{\mathsf{m}+1}\right)^n \quad \text{for } n \in \mathbb{N}_0.$$

Compare your result (as a function of m) graphically with the bounds

$$h_1(\mathsf{m}) \;=\; \frac{\mathsf{m}+1}{(\mathsf{m}+1)^2-\mathsf{m}} \quad \text{and} \quad h_2(\mathsf{m}) \;=\; \frac{\mathsf{m}+1}{\mathsf{m}^2}$$

derived from (1.13).

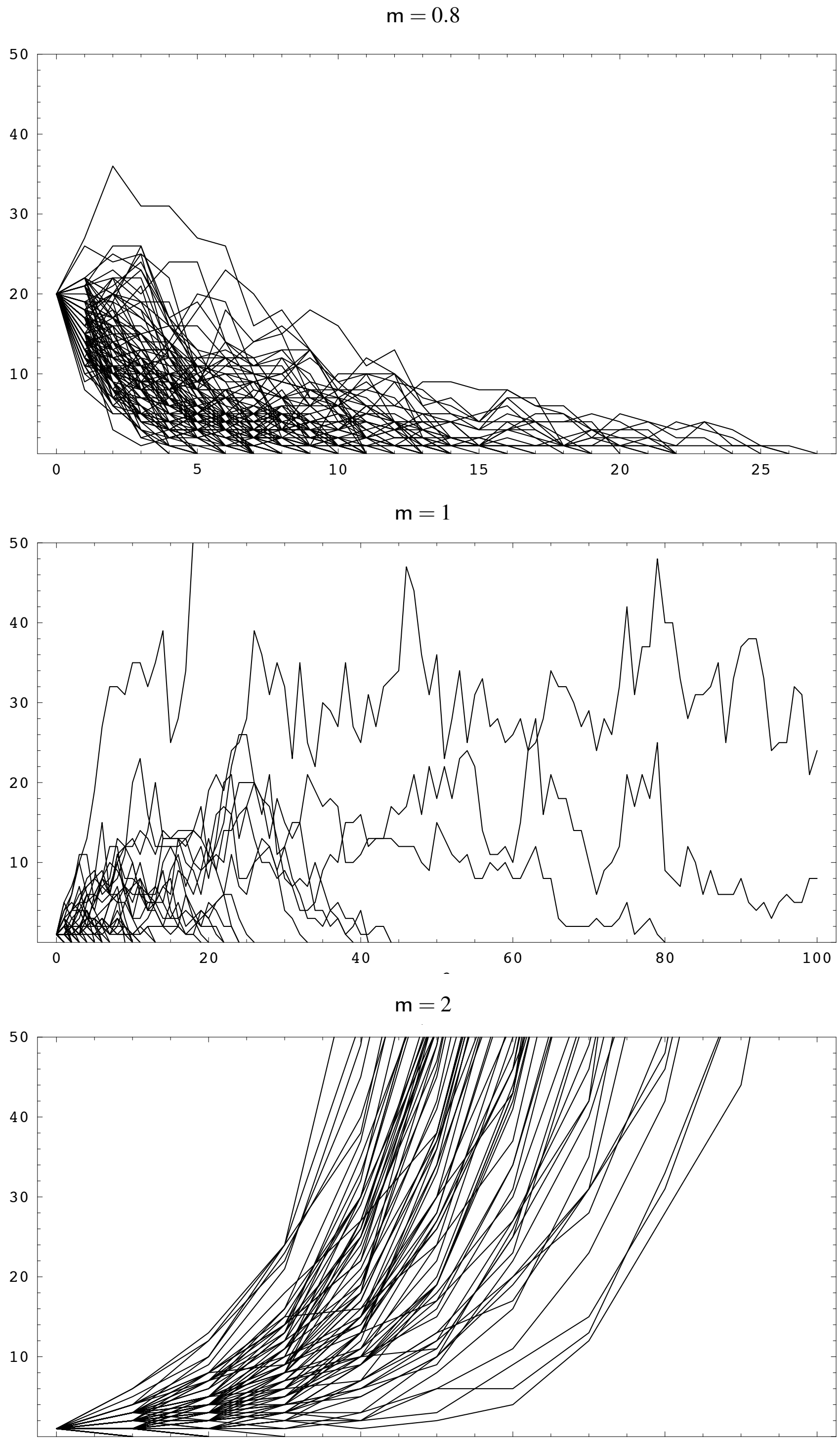

Fig. 1.4 100 simulations of a GWP with Poissonian offspring distribution having mean $\mathsf{m} = 0.8$ (top), $\mathsf{m} = 1$ (middle), and $\mathsf{m} = 2$ (bottom).

Problem 1.13. Let $(p_n)_{n\ge 0}$ be a Poisson distribution with mean $\mathsf{m} > 1$, i.e., $p_n = e^{-\mathsf{m}}\mathsf{m}^n/n!$ for $n \in \mathbb{N}_0$.

(a) Show that $q < \mathsf{m}^{-1}$ and $f'(q) \le f'(\mathsf{m}^{-1}) < 1$. By Corollary 1.8, this provides us with the error bound $f'(\mathsf{m}^{-1})^n$ when approximating q by $f_n(0)$.

(b) Compare the bound $h_3(\mathsf{m}) = \mathsf{m}^{-1}$ for q (as a function of m) graphically with the bounds

$$h_1(\mathsf{m}) \;=\; \frac{e^{-\mathsf{m}}}{1-\mathsf{m}e^{-\mathsf{m}}} \quad \text{and} \quad h_2(\mathsf{m}) \;=\; \frac{e^{-\mathsf{m}}}{1-(1+\mathsf{m})e^{-\mathsf{m}}}$$

derived from (1.13).

(c) Use iteration to compute q up to an error of at most 10^{-4} for the cases $\mathsf{m} = 1.1, 1.2, ..., 3.0$ and show your results graphically together with $h_i(\mathsf{m})$ for $i = 1,2,3$.

Problem 1.14. Let $(p_n)_{n\ge 0}$ be a binomial distribution with parameters N and m/N for some $N \in \mathbb{N}$ and $1 < \mathsf{m} \le N$, thus $p_n = \binom{N}{n}(\mathsf{m}/N)^n(1-(\mathsf{m}/N))^{N-n}$ for $n = 0,...,N$ and $p_n = 0$ otherwise.

(a) Show that, as in the Poisson case, $q < \mathsf{m}^{-1}$ but $f'(q) \le f'(\mathsf{m}^{-1}) > 1$ for sufficiently small $\mathsf{m} > 1$.

(b) Use iteration to compute q up to an error of at most 10^{-4} for $N = 3$ and $\mathsf{m} = 1.1, 1.2, ..., 2.0$. Show your results graphically together with the bounds

$$h_1(\mathsf{m}) \;=\; \frac{(1-(\mathsf{m}/N))^N}{1-\mathsf{m}(1-(\mathsf{m}/N))^{N-1}} \quad \text{and}$$

$$h_2(\mathsf{m}) \;=\; \frac{(1-(\mathsf{m}/N))^N}{1-(1+\mathsf{m}(1-1/N))(1-(\mathsf{m}/N))^{N-1}}$$

derived from (1.13) and $h_3(\mathsf{m}) = \mathsf{m}^{-1}$ (range $1 \le \mathsf{m} \le 3$).

(c) Do the same as in (b), but for $N = 2$. [This is the cell splitting case and thus q explicitly known from Problem 1.10].

Problem 1.15. Let $p_n = 1/[(n+1)(n+2)]$ for $n \in \mathbb{N}_0$. Find the gf of this offspring distribution with infinite mean and, via iteration, an approximation of the pertinent extinction probability.

Problem 1.16. Prove or give a counterexample for the assertion that $q \le \mathsf{m}^{-1}$ for any supercritical GWP $(Z_n)_{n\ge 0}$.

Problem 1.17. Prove or give a counterexample for the assertion that any two supercritical GWP's with one ancestor, the same offspring mean and extinction probability are already stochastically equivalent, that is, have the same offspring distribution. [Note that this is obviously false for two critical GWP's].

Problem 1.18. Let $\{(p_n(N))_{n\ge 0} : N \ge 1\}$ be a family of offspring distributions such that

$$\lim_{N\to\infty} p_n(N) \;=\; p_n \quad \text{for all } n \ge 0.$$

Show that, if $q(N)$ denotes the extinction probability associated with $(p_n(N))_{n\ge0}$, then $\lim_{N\to\infty} q(N) = q$.

1.5 The linear fractional case

The following offspring distribution, called *linear fractional* due to the form (1.14) of its gf f, constitutes one of the few nontrivial examples for which all iterations f_n can be computed in closed form. The subsequent derivation is taken from [4, I.4].

Given parameters $b,p \in (0,1)$, $b+p \le 1$, let

$$p_n \;=\; bp^{n-1} \quad \text{for } n \in \mathbb{N} \quad \text{and} \quad p_0 \;=\; 1 - \sum_{n\ge1} p_n \;=\; \frac{1-b-p}{1-p}.$$

If $b = p(1-p)$, this is the familiar geometric distribution with parameter $1-p$. The gf f of $(p_n)_{n\ge0}$ is easily seen to be of the form

$$f(s) \;=\; 1 - \frac{b}{1-p} + \frac{bs}{1-ps} \tag{1.14}$$

with mean value

$$\mathsf{m} \;=\; f'(1) \;=\; \frac{b}{(1-p)^2}. \tag{1.15}$$

Note that, as a power series, f has radius of convergence $p^{-1} > 1$. Since (1.14) is not yet the appropriate form for a computation of f_n, we will proceed with a number of transformations. For arbitrary $u,v < p^{-1}$, we have

$$\frac{f(s)-f(u)}{f(s)-f(v)} \;=\; \left(\frac{s-u}{s-v}\right)\left(\frac{1-pv}{1-pu}\right), \tag{1.16}$$

and the task is now to choose u and v in a useful manner. Towards this end, let 1 and $\hat{q}$ denote the two fixed points of f which are possible equal. Indeed, $\hat{q} = 1$ if $\mathsf{m} = 1$, while $\hat{q} = q < 1$ if $\mathsf{m} > 1$, and $q = 1 < \hat{q}$ if $\mathsf{m} < 1$.

Suppose first $\mathsf{m} \ne 1$. Then, choosing $u = \hat{q}$ and $v = 1$ in (1.16), we obtain

$$\frac{1-p}{1-p\hat{q}} \;=\; \lim_{s\uparrow1}\left(\frac{f(s)-\hat{q}}{s-\hat{q}}\right)\left(\frac{f(s)-1}{s-1}\right)^{-1} \;=\; \frac{1}{f'(1)} \;=\; \frac{1}{\mathsf{m}} \tag{1.17}$$

and thus in (1.16)

$$\frac{f(s)-\hat{q}}{f(s)-1} \;=\; \frac{s-\hat{q}}{\mathsf{m}(s-1)}.$$

This relation may now easily be iterated to yield

$$\frac{f_n(s)-\hat{q}}{f_n(s)-1} = \frac{s-\hat{q}}{\mathsf{m}^n(s-1)}$$

for each $n \geq 1$. Finally solving for $f_n(s)$, we arrive at the following result.

Lemma 1.19. *Let f be the gf of a linear fractional distribution with mean $\mathsf{m} \neq 1$ and fixed points $1, \hat{q}$. Then*

$$f_n(s) = 1 - \mathsf{m}^n \left(\frac{1-\hat{q}}{\mathsf{m}^n-\hat{q}}\right) + \frac{\mathsf{m}^n \left(\frac{1-\hat{q}}{\mathsf{m}^n-\hat{q}}\right)^2 s}{1-\left(\frac{\mathsf{m}^n-1}{\mathsf{m}^n-\hat{q}}\right)s} \tag{1.18}$$

for each $n \geq 1$.

Moreover, we infer from (1.15) and (1.17) in the case $\mathsf{m} > 1$ that

$$q = \hat{q} = \frac{1-b-p}{(1-p)p} = \frac{p_0}{p}. \tag{1.19}$$

There is a one-to-one correspondence between the pairs $(\mathsf{m}, q) \in (1,\infty) \times [0,1)$ and $(p,b) \in (0,1)^2,\ p+b<1$, i.e., for each reproduction mean m and extinction probability $q<1$, there exists exactly one linear fractional distribution with mean m and extinction probability q. Namely, by (1.15) and (1.19),

$$p = \frac{\mathsf{m}-1}{\mathsf{m}-q} \quad \text{and} \quad b = \frac{\mathsf{m}(1-q)^2}{(\mathsf{m}-q)^2}.$$

Turning to the critical case $\mathsf{m}=1$, we have $b=(1-p)^2$ by (1.15) and thus (☞ (1.14))

$$f(s) = \frac{p-(2p-1)s}{1-ps}. \tag{1.20}$$

Unlike the noncritical case, this relation may directly be iterated to give the following result.

Lemma 1.20. *Let f be the gf of a linear fractional distribution with mean $\mathsf{m}=1$. Then*

$$f_n(s) = \frac{np-((n+1)p-1)s}{(n-1)p+1-nps} \tag{1.21}$$

for each $n \geq 1$.

Based on the census in the United States in 1920, LOTKA [19, 20, 21] studied the extinction probability of white American families when restricting to male members. He found a good fit of empirical reproduction probabilities by a linear fractional distribution with parameters $b = 0.2126$ and $p = 0.5892$, which entails $p_0 = 0.4825$. These values may be found in [21] and differ slightly from those of his earlier contributions. Using a GWP as the underlying model, the theoretical extinction probability equals $q = 0.819$.

Problems

Problem 1.21. This extension of the linear fractional case is taken from [13, p. 10].

(a) Let f be a gf, f_n its n^{th} iteration and h an injective function on $[0,1]$ such that $g := h^{-1} \circ f \circ h$ constitutes a gf as well. Show that the n^{th} iteration of g is given by $g_n = h^{-1} \circ f_n \circ h$ for each $n \geq 1$.

(b) Consider the special case that $f(s) = (\mathsf{m} - (\mathsf{m}-1)s)^{-1}s$ for some $\mathsf{m} > 1$ and $h(s) = s^k$ for some $k \in \mathbb{N}$. Show that f and $g = h^{-1} \circ f \circ h$ are both gf and that

$$g_n(s) \;=\; \frac{s}{(\mathsf{m}^n - (\mathsf{m}^n - 1)s^k)^{1/k}}$$

for each $n \geq 1$.

Problem 1.22. Let $f(s) = 1 - p(1-s)^{\alpha}$ for arbitrary $p, \alpha \in (0,1)$. Show that f is a gf with iterations

$$f_n(s) \;=\; 1 - p^{1+\alpha+\ldots+\alpha^{n-1}}(1-s)^{\alpha^n}$$

for each $n \geq 1$.

1.6 A martingale and first limit theorems

Having provided a classification of simple GWP's as to their chance of survival, which turned out to be positive only in the supercritical case, the natural question to be addressed next is about the growth behavior of a GWP if it survives. We will give an answer at this time only under the restriction of finite reproduction variance because the general situation is more complicated and therefore requiring a deeper analysis that must wait until Section 2.1

We keep the notation from before and let $(Z_n)_{n\geq 0}$ denote a GWP in a standard model with offspring distribution $(p_n)_{n\geq 0}$ and mean $\mathsf{m} = \sum_{n\geq 1} np_n \in \mathbb{R}_>$. Recall from Section 1.1 that $\mathscr{F}_0 = \sigma(Z_0)$ and $\mathscr{F}_n = \sigma(Z_0, \{X_{j,k} : 1 \leq j \leq n, k \geq 1\})$ for $n \geq 1$. Put also $\mathscr{F}_\infty = \sigma(\mathscr{F}_n, n \geq 0)$.

In view of $\mathbb{E}Z_n = \mathsf{m}^n$ for each $n \geq 0$ it is natural to study the normalized process

$$W_n = \frac{Z_n}{\mathsf{m}^n}, \quad n \geq 0,$$

we have already encountered in the proof of Prop. 1.4 and which is next shown to be a nonnegative and thus a.s. convergent martingale.

Proposition 1.23. *The normalized sequence $(W_n, \mathscr{F}_n)_{n\geq 0}$ forms a nonnegative martingale under each $\mathbb{P}_i$, $i \in \mathbb{N}_0$, and thus converges a.s. to a nonnegative random variable W satisfying $\mathbb{E}_i W \leq i$. Furthermore,*

$$\mathbb{P}_i(W \in \cdot) \;=\; \mathbb{P}(W \in \cdot)^{*i} \tag{1.22}$$

and particularly $\mathbb{P}_i(W = 0) = \mathbb{P}(W = 0)^i$.

Proof. By (1.1), $Z_n = \sum_{k=1}^{Z_{n-1}} X_{n,k}$, where Z_{n-1} is $\mathscr{F}_{n-1}$-measurable and independent of the iid $X_{n,k}$, $k \geq 1$, with common distribution $(p_n)_{n\geq 0}$. As $\mathsf{m} \in \mathbb{R}_>$, it follows that

$$\mathbb{E}(Z_n|\mathscr{F}_{n-1}) \;=\; \mathbb{E}(Z_n|Z_{n-1}) \;=\; Z_{n-1}\mathbb{E}Z_1 \;=\; \mathsf{m}\, Z_{n-1} \quad \mathbb{P}_i\text{-a.s.}$$

and thus the martingale property of $(W_n, \mathscr{F}_n)_{n\geq 0}$ under each $\mathbb{P}_i$. Almost sure convergence is ensured by the martingale convergence theorem, while Fatou's lemma ensures that

$$\mathbb{E}_i W \;=\; \mathbb{E}_i \lim_{n\to\infty} W_n \;\leq\; \lim_{n\to\infty} \mathbb{E}_i W_n \;=\; \mathbb{E}_i W_0 \;=\; i$$

for each $i \in \mathbb{N}_0$. A proof of the remaining assertions is left as an exercise [☞ Problem 1.27]. ☐

It must be noted that this result is of relatively limited use because it provides satisfactory information on the growth behavior of $(Z_n)_{n\geq 0}$ only on $\{W > 0\}$, which has probability zero if $\mathsf{m} \leq 1$ and $p_1 \neq 1$. In the supercritical case $\mathsf{m} > 1$, we would like to infer that $\{W > 0\} = \{Z_n \to \infty\}$ a.s., for this would mean that m^n is always the "right" normalization of Z_n on the event of survival, but in general only the obvious inclusion $\{W > 0\} \subset \{Z_n \to \infty\}$ holds true. Indeed, the next lemma shows that, if the inclusion is strict, we already have $W = 0$ a.s.

Lemma 1.24. *The following dichotomy holds true for the martingale limit W in Theorem 1.23. Either $W = 0$ $\mathbb{P}_i$-a.s. for each $i \in \mathbb{N}$, or*

$$\mathbb{P}_i(W > 0) \;=\; \mathbb{P}_i(Z_n \to \infty) \quad \textit{for each } i \in \mathbb{N}. \tag{1.23}$$

If $\mathsf{m} \leq 1$ and $p_1 \neq 1$, the first alternative occurs.

Proof. In view of the valid inclusion $\{W > 0\} \subset \{Z_n \to \infty\}$, it suffices to prove that $\rho_i := \mathbb{P}_i(W = 0)$ either equals 1 or $\mathbb{P}_i(Z_n = 0 \text{ eventually}) = q^i$ for each i. It further suffices to consider $i = 1$, for $\rho_i = \rho_1^i$ for each $i \in \mathbb{N}$ by Theorem 1.23.

Towards this end, put $\rho = \rho_1$ and note first that

$$Z_{n+1} = \sum_{j=1}^{Z_1} Z_n(j), \quad n \geq 0, \tag{1.24}$$

where $(Z_n(j))_{n\geq 0}$ denotes the GWP which pertains to the subpopulation originating from the j^{th} child of the ancestor $\varnothing$ [☞ also Problem 1.28]. Clearly, Z_1 is independent of the family $\{(Z_n(j))_{n\geq 0} : j \geq 1\}$. After normalization and letting n tend to infinity, we obtain from (1.24) the identity

$$W = \frac{1}{\mathsf{m}} \sum_{j=1}^{Z_1} W(j) \quad \text{a.s.} \tag{1.25}$$

where $W(1), W(2), ...$ are independent copies of W and also independent of Z_1. Consequently,

$$\begin{aligned} \rho &= \mathbb{P}(W(j) = 0 \text{ for } j = 1, ..., Z_1) \\ &= \sum_{n\geq 0} \mathbb{P}(Z_1 = n)\,\mathbb{P}(W(1) = ... = W(n) = 0) = \sum_{n\geq 0} p_n \rho^n = f(\rho) \end{aligned}$$

and so $\rho \in \{q, 1\}$. This proves the asserted dichotomy and finishes the proof because the last assertion is trivial. □

Remark 1.25. Let us point out that (1.25) provides an example of a so-called *distributional equation* of the general form

$$X \stackrel{d}{=} \Phi(X_1, X_2, ...) \tag{1.26}$$

for some *random* function Φ independent of the iid $X_1, X_2, ...$, which in turn are copies of X. Here $\stackrel{d}{=}$ means equality in distribution. Any distribution Q such that (1.26) holds true with $X \stackrel{d}{=} Q$ is called a *solution* to this equation, although in slight abuse of language we sometimes call the random variable X a solution as well. Distributional equations will be met frequently throughout this text because many limiting distributions in connection with branching processes are characterized by them, owing to the recursive structure of these processes that is reflected in these equations. Typically, these limiting distributions cannot be identified explicitly and all information about them must therefore be drawn from the pertinent distributional equation they solve. A first example for this procedure has been provided by the proof of the previous lemma. A more sophisticated one can be found in Section 2.1 where it will shown that the distribution of W if nontrivial is absolutely continuous with continuous density.

As for (1.25), we finally note that the random function Φ is given by $\Phi(x_1, x_2, ...) = \mathsf{m}^{-1} \sum_{k=1}^{Z_1} x_k$. Moreover, the distributional equality even appears in the stronger form of a ”physical” equality of random variables.

Lemma 1.24 calls for a further investigation of the question under which condition (1.23) holds true in the supercritical case. The bad news is that $1 < \mathsf{m} < \infty$ does not suffice as was first note by LEVINSON [18]. The highly nontrivial task of providing an answer in form of a necessary and sufficient condition will be accomplished in Section 2.1. Here we content ourselves with a sufficient condition that allows the desired conclusion by resorting to a standard result from martingale theory. By (1.22), it suffices to consider the case of one ancestor.

Theorem 1.26. *Suppose that $(Z_n)_{n\ge0}$ is a supercritical GWP with one ancestor and finite reproduction variance σ^2. Then*

$$\lim_{n\to\infty}\mathbb{E}(W_n-W)^2\ =\ 0, \tag{1.27}$$

$$\mathbb{E}W\ =\ 1\quad\textit{and}\quad \mathbb{V}\mathrm{ar}\,W\ =\ \frac{\sigma^2}{\mathsf{m}(\mathsf{m}-1)}, \tag{1.28}$$

$$\mathbb{P}(W=0)\ =\ q. \tag{1.29}$$

Proof. Since $(Z_n)_{n\ge0}$ has finite reproduction variance σ^2, we infer from (1.9) in the proof of Prop. 1.4 that [☞ Problem 1.5]

$$\mathbb{V}\mathrm{ar}\,W_n\ =\ \frac{\sigma^2(1-\mathsf{m}^{-n})}{\mathsf{m}(\mathsf{m}-1)}\quad\text{for each } n\ge1.$$

Hence, $(W_n)_{n\ge0}$ is an L^2-bounded martingale and thus convergent in L^2 to W as claimed [☞ [12, Cor. 2.2]]. Furthermore, $\mathbb{E}W=1$ and $\mathbb{V}\mathrm{ar}\,W=\lim_{n\to\infty}\mathbb{V}\mathrm{ar}\,W_n=(\mathsf{m}(\mathsf{m}-1))^{-1}\sigma^2$, which shows (1.28). Finally, as $\mathbb{P}(W>0)$ is positive it must equal $1-q$ by Lemma 1.24. □

Problems

Problem 1.27. Let $(Z_n)_{n\ge0}$ be a GWP in a standard model and $(k_n)_{n\ge0}$ a normalizing sequence of positive reals such that $W_n^*=k_n^{-1}Z_n$ converges $\mathbb{P}_i$-a.s. for each $i\in\mathbb{N}_0$ to a random variable W^* taking values in $[0,\infty]$. Show that

(a) (1.22) holds true for W^*;
(b) If $k_n^{-1}k_{n+1}\to\mathsf{m}<\infty$, then W^* satisfies the distributional equation (1.25) and thus Lemma 1.24 as well. Moreover, $\kappa:=\mathbb{P}(W^*<\infty)$ is a fixed point of f and thus either equals q or 1.

Problem 1.28. Consider the genealogical model described in Section 1.2 and let, for any $\mathsf{u}\in\mathbb{V}$, $\boldsymbol{GW}(\mathsf{u})$ be the Galton-Watson random tree emanating from u as the root and based on the $\{X_{\mathsf{uv}}:\mathsf{v}\in\mathbb{V}\}$. Then define the associated GWP $(Z_n(\mathsf{u}))_{n\ge0}$ with ancestor u by

$$Z_n(\mathsf{u}) \;=\; |\{\mathsf{uv} \in \boldsymbol{GW}(\mathsf{u}) : |\mathsf{v}| = n\} \;=\; \sum_{\mathsf{v}\in\mathbb{V}:|\mathsf{v}|=n-1 \text{ and } \mathsf{uv}\in\boldsymbol{GW}(\mathsf{u})} X_{\mathsf{uv}}$$

for $n \ge 1$. Finally, put $Z_n = Z_n(\varnothing)$, $W_n^* = k_n^{-1} Z_n$ and $W_n^*(\mathsf{u}) = k_n^{-1} Z_n(\mathsf{u})$ for a normalizing sequence $(k_n)_{n\ge 0}$ of positive reals. Show that:

(a) $W_{j+n}^* = (k_j/k_{j+n}) \sum_{\mathsf{u}\in\boldsymbol{GW}:|\mathsf{u}|=n} W_j^*(\mathsf{u})$ for each $j \ge 1$;

(b) If W_n^* converges a.s. to a random variable W^* taking values in $[0,\infty]$ and $k_n^{-1} k_{n+1} \to \mathsf{m} < \infty$, then W^* satisfies the distributional equation

$$W^* \;=\; \frac{1}{\mathsf{m}^n} \sum_{\mathsf{u}\in\boldsymbol{GW}:|\mathsf{u}|=n} W^*(\mathsf{u}) \quad \text{a.s.} \tag{1.30}$$

for each $n \ge 1$, where the $W^*(\mathsf{u})$ have the obvious meaning and are independent of Z_n. Notice that this particularly implies

$$W^* \;\stackrel{d}{=}\; \frac{1}{\mathsf{m}^n} \sum_{k=1}^{Z_n} W_k^* \tag{1.31}$$

for each $n \ge 1$ if $W_1^*, W_2^*, \ldots$ denote iid copies of W^* that are independent of $(Z_n)_{n\ge 0}$.

Problem 1.29. Show that, under the general conditions stated at the beginning of this section,

$$\mathbb{P}(W \in \cdot|\mathscr{F}_n) \;=\; \mathbb{P}(W \in \cdot|Z_n) \quad \mathbb{P}_i\text{-a.s.} \tag{1.32}$$

for all $i, n \in \mathbb{N}_0$.