2 Branching processes

2.1 Classification and extinction

Informally, a branching process\(^{10}\) is described as follows: let \(\{p_k\}_{k \geq 0}\) be a fixed probability mass function. A population starts with a single ancestor who forms generation number 0. This initial individual splits into \(k\) offspring with probability \(p_k\). These offspring constitute the first generation. Each of the offspring in the first generation splits independently into a random number of offspring according to the pmf \(p_k\). This process continues until extinction, which occurs when all the members of a generation fail to produce offspring.

This model has a number of applications in biology (e.g., it can be thought as a model of population growth), physics (chain reaction in nuclear fission), queueing theory etc. Originally it arose from a study of the likelihood of survival of family names (“how fertile must a family be to insure that in no future generation will the family name die out?”).

Formally, let \(\{Z_{n,k}\}, n \geq 1, k \geq 1\), be a family of i.i.d. random variables in \(\mathbb{Z}^+ \equiv \{0, 1, 2, \ldots\}\), each having a common distribution \(\{p_k\}_{k \geq 0}\). Then the branching process \((Z_n)_{n \geq 0}\) (generated by \(\{p_k\}_{k \geq 0}\)) is defined via \(Z_0 = 1\), and, for \(n \geq 1\),

\[
Z_n \stackrel{\text{def}}{=} Z_{n,1} + Z_{n,2} + \cdots + Z_{n,Z_{n-1}},
\]

where the empty sum is interpreted as zero. Notice that \(Z_n\) is a Markov chain in \(\mathbb{Z}^+\). We shall use \(P(\cdot) \equiv P_1(\cdot)\) and \(E(\cdot) \equiv E_1(\cdot)\) to denote the corresponding probability measure and the expectation operator.\(^{11}\)

If \(\varphi_n(s) \equiv \mathbb{E}s^{Z_n}\) is the generating function of \(Z_n\), a straightforward induction based on (2.1) and (1.3) implies

\[
\varphi_0(s) \equiv s, \quad \varphi(s) \equiv \varphi_1(s) \equiv \mathbb{E}s^1, \quad \varphi_k(s) \equiv \varphi_{k-1}(\varphi(s)), \quad k > 1. \tag{2.2}
\]

Usually explicit calculations are hard, but at least in principle, equations (2.2) determine the distribution of \(Z_n\) for any \(n \geq 0\).

Example 2.1. Let \(\varphi_1(s) \equiv \varphi(s) = q + ps\) for some \(0 < p = 1 - q < 1\). Then

\[
\varphi_n(s) \equiv q(1 + p + \cdots + p^{n-1}) + p^n s = 1 + p^n(s - 1).
\]

Notice that here we have \(\varphi_n(s) \to 1\) as \(n \to \infty\) for all \(s \in [0, 1]\), i.e., the distribution of \(Z_n\) converges to that of \(Z_\infty \equiv 0\).

The following result is a straightforward corollary of (1.3).

Exercise 2.2. In a branching process \((Z_n)_{n \geq 0}\) with \(Z_0 = 1\), let the offspring distribution have mean \(m\), variance \(\sigma^2\), and generating function \(\varphi(s)\). Write \(\varphi_n(s)\) for the generating function of the \(n\)th generation size \(Z_n\).

a) Using \(\varphi_n(s) \equiv \varphi_{n-1}(\varphi(s))\) or otherwise, show that \(\mathbb{E}Z_n = m^n;\)

\(^{10}\)sometimes called Galton-Watson-Bienaymé process

\(^{11}\)If \(Z_0 = k\), we shall explicitly write \(P_k(\cdot)\) and \(E_k(\cdot)\).
b) Using \( \varphi_n(s) \equiv \varphi(\varphi_{n-1}(s)) \) or otherwise, show that
\[
\text{Var}(Z_n) = \begin{cases} 
\sigma^2 m^{n-1} \frac{m^n - 1}{m - 1}, & m \neq 1, \\
\sigma^2 n, & m = 1.
\end{cases}
\]

c) Deduce that \( \mathbb{E}[(Z_n/m^n)^2] \) is uniformly bounded.

This result suggests that if \( m \equiv \mathbb{E}Z \neq 1 \), the branching process might explode (for \( m > 1 \)) or die out (for \( m < 1 \)). One classifies branching process into critical (if \( m = 1 \)), subcritical (\( m < 1 \)), and supercritical (\( m > 1 \)).

**Remark 2.2.1.** It is straightforward to describe the case \( m < 1 \). Indeed, the Markov inequality implies that
\[
P(Z_n > 0) = P(Z_n > 1) \leq \mathbb{E}(Z_n) = m^n,
\]
so that \( P(Z_n > 0) \to 0 \) as \( n \to \infty \) (ie., \( Z_n \to 0 \) in probability). Moreover, since \( \sum_{n \geq 0} P(Z_n > 0) < \infty \), the Borel-Cantelli lemma implies that \( P(Z_n \to 0) = 1 \) (ie., \( Z_n \to 0 \) almost surely). We also notice that the average total population in this case is finite, \( \mathbb{E}(\sum_{n \geq 0} Z_n) = \sum_{n \geq 0} m^n = (1 - m)^{-1} < \infty \).

**Definition 2.3.** The extinction event \( \mathcal{E} \) is the event \( \mathcal{E} = \bigcup_{n=1}^{\infty} \{ Z_n = 0 \} \). Since \( \{ Z_n = 0 \} \subset \{ Z_{n+1} = 0 \} \) for all \( n \geq 0 \), the extinction probability \( \rho \) is
\[
\rho = P(\mathcal{E}) = \lim_{n \to \infty} P(Z_n = 0)
\]
with \( P(Z_n = 0) \equiv \varphi_n(0) \) being the extinction probability before \( n+1 \)st generation.

The following result helps to derive the extinction probability \( \rho \) without need to compute iterates \( \varphi_n(\cdot) \). To avoid trivialities we shall assume that \( p_0 = P(Z = 0) \) satisfies\(^{12} 0 < p_0 < 1 \); notice that under this assumption \( \varphi(s) \) is a strictly increasing function of \( s \in [0, 1] \).

**Theorem 2.4.** If \( 0 < p_0 < 1 \), then the extinction probability \( \rho \) is given by the smallest positive solution to the equation
\[
s = \varphi(s).
\] (2.3)

In particular, if \( m = \mathbb{E}Z \leq 1 \), then \( \rho = 1 \); otherwise, we have \( 0 < \rho < 1 \).

**Remark 2.4.1.** The relation \( \rho = \varphi(\rho) \) has a clear probabilistic sense. Indeed, if \( \rho = P_1(\mathcal{E}) \) is the extinction probability starting from a single individual, \( Z_0 = 1 \), then by independence we get \( P_k(\mathcal{E}) = P(\mathcal{E} \mid Z_0 = k) = \rho^k \), and thus the first step decomposition for the Markov chain \( Z_n \) gives
\[
\rho = P(\mathcal{E}) = \sum_{k \geq 0} P(\mathcal{E}, Z_1 = k) = \sum_{k \geq 0} P(\mathcal{E} \mid Z_1 = k)P(Z_1 = k)
\]
\[
= \sum_{k \geq 0} \rho^k P(Z_1 = k) \equiv E(\rho^{Z_1}) \equiv \varphi(\rho),
\]
in agreement with (2.3).

\(^{12}\) otherwise the model is degenerated: if \( p_0 = 0 \), then \( Z_n \geq 1 \) for all \( n \geq 0 \) so that \( \rho = 0 \); if \( p_0 = 1 \), then \( P(Z_1 = 0) = \rho = 1 \).
Proof of Theorem 2.4. [Draw the picture!]

Denote \( \rho_n = P(Z_n = 0) \equiv \varphi_n(0) \).
By continuity and strict monotonicity of the generating function \( \varphi(\cdot) \) we have (recall (2.2))
\[
0 < \rho_1 = \varphi(0) < \rho_2 = \varphi(\rho_1) < \cdots < 1,
\]
so that \( \rho_1 \neq \rho \in (0, 1] \) with \( \rho = \varphi(\rho) \).

Now if \( \rho \) is another fixed point of \( \varphi(\cdot) \) in \([0, 1]\), ie., \( \rho = \varphi(\rho) \), then
\[
0 < \rho_1 = \varphi(0) < \rho_2 < \cdots < \varphi(\rho) = \rho
\]
so that \( \rho = \lim_{n \to \infty} \rho_n \leq \rho \), ie., \( \rho \) is the smallest positive solution to (2.3).

We finally observe that in view of convexity of \( \varphi(\cdot) \), the condition \( m = \varphi'(1) < 1 \) implies \( \rho = 1 \) and the condition \( m = \varphi'(1) > 1 \) implies that \( \rho \) is the unique solution to the fixed point equation (2.3) in \((0, 1)\).

A similar argument gives the following result.

\[\textbf{Corollary 2.5.} \text{ If } s \in [0, 1), \text{ we have } \varphi_n(s) \equiv E[s^{Z_n}] \to \rho \in (0, 1] \text{ as } n \to \infty.\]

\[\textbf{Remark 2.5.1.} \text{ Consequently, the distribution of } Z_n \text{ converges to that of the limiting random variable } Z_{\infty}, \text{ where } P(Z_{\infty} = 0) = \rho \text{ and } P(Z_{\infty} = \infty) = 1 - \rho.\]

\[\textbf{Exercise 2.6.} \text{ For a branching process with generating function } \varphi(s) = a + bs + cs, \text{ where } a > 0, b > 0, c > 0, \varphi(1) = 1, \text{ compute the extinction probability } \rho \text{ and give the condition for sure extinction. Can you interpret your results?}\]

\[\textbf{Exercise 2.7.} \text{ Let } (Z_n)_{n \geq 0} \text{ be a branching process with generating function } \varphi(s) \equiv E_s Z_1^n \text{ satisfying } 0 < \varphi(0) < 1. \text{ Let } \tilde{\varphi}_n(u) \overset{\text{def}}{=} E u Z_n \text{ be the generating function of } \tilde{Z}_n = \sum_{k=0}^{n} Z_k, \text{ the total population size up to time } n. \text{ Show that } \tilde{\varphi}_{n+1}(u) = u \varphi(\tilde{\varphi}_n(u)) \text{ for all } n \geq 0 \text{ and } u \geq 0. \text{ Deduce that } E(u^{\tilde{Z}} \mathbb{1}_{\tilde{Z}_{\infty} < \infty}), \text{ } u \geq 0, \text{ where } \tilde{Z} = \sum_{k \geq 0} Z_k \text{ is the total population size of }(Z_n)_{n \geq 0}, \text{ is given by the smallest solution } s \geq 0 \text{ to the equation } s = u \varphi(s), \text{ when it exists.}\]

We now turn to classification of states for the Markov chain \( Z_n \) in \( \mathbb{Z}^+ \). Of course, since 0 is an absorbing state, it is recurrent.

\[\textbf{Lemma 2.8.} \text{ If } p_1 = P(Z = 1) \neq 1, \text{ then every } k \in \mathbb{N} \text{ is transient. As a result, } \]
\[
P(Z_n \to \infty) = 1 - P(Z_n \to 0) = 1 - \rho.
\]

\[\textbf{Proof.} \text{ We first show that every } k \in \mathbb{N} \text{ is transient:}\]

If \( p_0 = 0 \), then \( Z_n \) is a non-decreasing Markov chain (ie., \( Z_{n+1} \geq Z_n \)), so that for every \( k \in \mathbb{N} \) the first passage probability \( f_k \) satisfies
\[
f_{kk} = P(Z_{n+1} = k \mid Z_n = k) = P_k(Z_1 = k) = (p_1)^k < 1.
\]

If \( p_0 \in (0, 1] \), we have
\[
f_{kk} \leq P(Z_{n+1} \neq 0 \mid Z_n = k) = P_k(Z_1 \neq 0) = 1 - P_k(Z_1 = 0) = 1 - (p_0)^k < 1.
\]

We now turn to the explosion event \( \{Z_n \to \infty\} \). Fix \( K > 0 \); since the states 1, 2, \ldots, \( K \) are transient, we eventually have \( \{Z_n \to 0\} \cup \{Z_n > K\} \) and therefore
\[
P(Z_n \to 0 \text{ or } Z_n \to \infty) = 1.
\]
As the LHS above equals \( P(Z_n \to 0) + P(Z_n \to \infty) \), the result follows from the observation \( P(Z_n \to 0) \equiv P(E) = \rho \).
Remark 2.8.1. In the case \( p_0 \in (0,1] \) the transience result above is well known in the general theory of Markov chains: since state 0 is absorbing and for every \( k \in \mathbb{N} \) the event

\[
\{ k \text{ leads to } 0 \} \equiv \{ Z_n = 0 \text{ for some } n \text{ given } Z_0 = k \}
\]

has probability at least \( p_0^k > 0 \), state \( k \) is transient.

Remark 2.8.2. A formal justification of the final step of the argument above can be derived as follows. Let \( A = \bigcup_{n \geq 0} \{ Z_n = 0 \} \equiv \{ Z_n \to 0 \} \) be the absorption event. With \( C_m = \{1,2,\ldots,m\} \), by the above,

\[
\mathbb{P}(Z_n \in C_m \text{ f.o.}) = \mathbb{P}(Z_n \notin C_m \text{ eventually}) = 1
\]

for all \( m \geq 1 \). Further, denote \( A_m^c = \{ Z_n > m \text{ eventually} \} \). Since

\[
A_m^c \setminus A_m^{c-1} \subseteq \{ Z_n = m \text{ i.o.} \} \subseteq \{ Z_n \in C_m \text{ i.o.} \}
\]

has probability 0, we deduce that

\[
\mathbb{P}(A_m^c) = \mathbb{P}(A_m^c) = \mathbb{P}(A^c) = 1 - \mathbb{P}(A)
\]

for all \( m \geq 1 \). By monotonicity of \( A_m \) in \( m \), the event \( \{ Z_n \to \infty \} = \bigcap_{n \geq 1} A^c_m \) has probability \( \mathbb{P}(A^c) \), i.e.,

\[
\mathbb{P}(Z_n \to 0) + \mathbb{P}(Z_n \to \infty) = 1.
\]

Exercise 2.9. Let \((Z_n)_{n \geq 0}\) be a supercritical branching process with offspring distribution \( \text{Poi}(\lambda) \), \( \lambda > 1 \). Let \( T_0 = \min\{n \geq 0 : Z_n = 0\} \) be its extinction time, and let \( \rho = \mathbb{P}(T_0 < \infty) > 0 \) be its extinction probability. Define \((\bar{Z}_n)_{n \geq 0}\) as \( Z_n \) conditioned on extinction, i.e., \( \bar{Z}_n = (Z_n \mid T_0 < \infty) \).

a) Show that the transition probabilities \( \bar{p}_{xy} \) of \((\bar{Z}_n)_{n \geq 0}\) and the transition probabilities \( p_{xy} \) of the original process \((Z_n)_{n \geq 0}\) are related via \( \bar{p}_{xy} = p_{xy}\rho^{-x} \), \( x, y \geq 0 \);

b) Deduce that the generating functions \( \hat{\varphi}(s) \equiv \mathbb{E}_1[s^{\bar{Z}_1}] \) and \( \varphi(s) \equiv \mathbb{E}_1[s^{Z_1}] \) are related via \( \hat{\varphi}(s) = \frac{1}{\rho}\varphi(\rho s) \), \( 0 \leq s \leq 1 \);

c) Use the fixed point equation \( \rho = e^{\lambda(\rho-1)} \) to show that \( \hat{\varphi}(s) = e^{\lambda \rho(s-1)} \), i.e., that the offspring distribution for \((\bar{Z}_n)_{n \geq 0}\) is just \( \text{Poi}(\lambda \rho) \).

Exercise 2.10. Let \((Z_n)_{n \geq 0}\) be a supercritical branching process with offspring distribution \( \{p_k\}_{k \geq 0} \), offspring generating function \( \varphi(s) \) and extinction probability \( \rho \in (0,1) \).

a) If \( Z_0 = 1 \), let \( \bar{p}_k \) be the probability that conditioned on survival the first generation has exactly \( k \) individuals with an infinite line of descent. Show that

\[
\bar{p}_k = \frac{1}{1-\rho} \sum_{n=k}^{\infty} p_n \binom{n}{k} (1-\rho)^k \rho^{n-k}.
\]

\[\text{geometrically, the graph of } \hat{\varphi}(\cdot) \text{ is a rescaled version of that of } \varphi(\cdot);\]
b) Let \((\tilde{Z}_n)_{n \geq 0}\) count only those individuals in \((Z_n)_{n \geq 0}\), who conditioned on survival have an infinite line of descent. Show that \((\tilde{Z}_n)_{n \geq 0}\) is a branching process with offspring generating function\(^{13}\)

\[ \tilde{\varphi}(s) = \frac{1}{1-\rho} \left( \varphi((1-\rho)s + \rho) - \rho \right). \]

**Exercise 2.11.** Let \((Z_n)_{n \geq 0}\) be a subcritical branching process whose generating function \(\varphi(s) = E(s^{Z_1})\) is finite for some \(s > 1\), i.e., the offspring distribution has finite exponential moments in a neighbourhood of the origin.

a) Using the result of Exercise 2.7 or otherwise, show that the total population size \(Z = \sum_{k \geq 0} Z_k\) satisfies \(E(u^Z) < \infty\) for some \(u > 1\).

b) Suppose that for each \(1 \leq i \leq \tilde{Z}\), individual \(i\) produces wealth of size \(W_i\), where \(W_i\) are independent random variables with common distribution satisfying \(E(s^{W}) < \infty\) for some \(s > 1\). Show that for some \(u > 1\) we have \(E(u^{\tilde{W}}) < \infty\), where \(\tilde{W} = W_1 + \cdots + W_{\tilde{Z}}\) is the total wealth generated by \((Z_n)_{n \geq 0}\).

**Exercise 2.12.** Let \((Z_n)_{n \geq 0}\) be a supercritical branching process whose generating function \(\varphi(s) = E(s^{Z_1})\) is finite for some \(s > 1\).

a) Using the result of Exercise 2.7 or otherwise, show that the total population size \(Z = \sum_{k \geq 0} Z_k\) satisfies \(E(u^Z 1_{\tilde{Z} < \infty}) < \infty\) for some \(u > 1\).

b) Define \((\hat{Z}_n)_{n \geq 0}\) as \(Z_n\) conditioned on extinction, i.e., \(\hat{Z}_n = (Z_n \mid \tilde{Z} < \infty)\). Deduce that \((\hat{Z}_n)_{n \geq 0}\) is a subcritical branching process such that \(E(u^{\hat{Z}_1}) < \infty\) for some \(u > 1\).

**Remark 2.12.1.** The fact that a probability generating function is finite at \(u = 1\) (or has a finite (left) derivative there) does not, in general, imply any regularity beyond the unit disk:

Let \(X\) be a random variable such that

\[ P(X = k) = \frac{1}{k(k+1)} \quad \text{for all } k \geq 1, \]

and let \(\varphi_X(u)\) be its generating function. It is easy to check that \(\varphi_X(1) = 1\) while \(E(X) = \varphi_X'(1-) = \infty\), and thus \(|\varphi_X(u)| \leq \varphi_X(1) = 1\) if \(|u| \leq 1\) but \(\varphi_X(u) = \infty\) for all \(|u| > 1\).

Now suppose that

\[ P(X = k) = \frac{4}{k(k+1)(k+2)} \quad \text{for } k \geq 1. \]

It is easy to see that the generating function \(\varphi_X(u)\) of \(X\) satisfies \(\varphi_X(1) = 1\), \(\varphi_X'(1-) = 2 < \infty\), but \(\varphi_X''(1-) = \infty\). Notice that in this case \(\varphi_X'(u) \leq 2\) uniformly in \(|u| < 1\) while still \(\varphi_X(u) = \infty\) for all \(|u| > 1\).

Following a similar pattern, for every \(m \in \mathbb{N}\), one can construct a function \(\varphi(u)\), which is continuous and bounded on the closed unit disk \(|u| \leq 1\) together with derivatives up to order \(m\), while \(\varphi(u) = \infty\) for all \(|u| > 1\).
2.2 Critical case \( m = 1 \)

The following example is among very few, for which the computation in the critical case \( m = EZ = 1 \) can be done explicitly.

Example 2.13. Consider the so-called linear-fractional case, where the offspring distribution is given by

\[
p_j = \frac{1}{2j+1}, \quad j \geq 0.
\]

We then have \( \varphi(s) = \sum_{j \geq 0} s^j/2^{j+1} = (2-s)^{-1} \) and a straightforward induction gives (check this!)

\[
\varphi_k(s) = \frac{k-(k-1)s}{(k+1) - ks} = \frac{k}{k+1} + \frac{1}{k(k+1)} \sum_{m \geq 1} \left( \frac{ks}{k+1} \right)^m,
\]

so that

\[
P(Z_k = 0) = \varphi_k(0) = \frac{k}{k+1}, \quad P(Z_k > 0) = \frac{1}{k+1},
\]

and

\[
P(Z_k = m \mid Z_k > 0) = \frac{1}{k+1} \left( \frac{k}{k+1} \right)^{m-1},
\]

ie., \((Z_k \mid Z_k > 0)\) has geometric distribution with success probability \(1/(k+1)\).

Remark 2.13.1. By the partition theorem,

\[
1 = E(Z_k) = E(Z_k \mid Z_k > 0) P(Z_k > 0) + E(Z_k \mid Z_k = 0) P(Z_k = 0),
\]

so that in the previous example we have

\[
E(Z_k \mid Z_k > 0) = \frac{1}{P(Z_k > 0)} = k+1,
\]

ie., conditional on survival, the average generation size grows linearly with time.

The following example is known as the general linear-fractional case:

Exercise 2.14. For fixed \( b > 0 \) and \( p \in (0,1) \), consider a branching process with the offspring distribution

\[
p_j = bp^{j-1}, \quad j \geq 1, \quad p_0 = 1 - \sum_{j \geq 1} p_j.
\]

a) Show that for \( b \in (0,1-p) \) the distribution above is well defined; find the corresponding \( p_0 \), and show that

\[
\varphi(s) = \frac{1-b-p}{1-p} + \frac{bs}{1-ps};
\]

b) Find \( b \) for which the branching process is critical and show that then

\[
\varphi_k(s) = E(s^{Z_k}) = \frac{kp - (kp+p-1)s}{(1-p+kp) - kps};
\]

c) Deduce that \((Z_k \mid Z_k > 0)\) is geometrically distributed with parameter \(1/(kp+1-p)\).
Straightforward computer experiments (see the R script on the webpage!) show that a similar linear growth of \( \mathbb{E}(Z_k \mid Z_k > 0) \) takes place for other critical offspring distributions, e.g., the one with \( \varphi(s) = (1 + s^2)/2 \).

**Theorem 2.15.** If the offspring distribution of the branching process \((Z_k)_{k \geq 0}\) has mean \( m = 1 \) and finite variance \( \sigma^2 > 0 \), then

\[
k \mathbb{P}(Z_k > 0) \to \frac{2}{\sigma^2} \quad \text{as } k \to \infty;
\]
equivalently,

\[
\frac{1}{k} \mathbb{E}(Z_k \mid Z_k > 0) \to \frac{\sigma^2}{2} \quad \text{as } k \to \infty. \tag{2.4}
\]

**Remark 2.15.1.** This general result suggests that, conditional on survival, a general critical branching process exhibits linear intermittent behaviour;\(^{14}\) namely, with small probability (of order \( 2/(k\sigma^2) \)) the values of \( Z_k \) are of order \( k \).

Our argument is based on the following general fact:\(^{15}\)

**Lemma 2.16.** Let \((y_n)_{n \geq 0}\) be a real-valued sequence. If for some constant \( a \) we have \( y_{n+1} - y_n \to a \) as \( n \to \infty \), then \( n^{-1}y_n \to a \) as \( n \to \infty \).

**Proof.** By changing the variables \( y_n \mapsto y'_n = y_n - na \) if necessary, we can and shall assume that \( a = 0 \). Fix arbitrary \( \delta > 0 \) and find \( K > 0 \) such that for \( n \geq K \) we have \( |y_{n+1} - y_n| \leq \delta \). Decomposing, for \( n > K \),

\[ y_n - y_K = \sum_{j=K}^{n-1} (y_{j+1} - y_j) \]

we deduce that \( |y_n - y_K| \leq \delta(n - K) \) so that the claim follows from the estimate

\[
\left| \frac{y_n}{n} \right| \leq \left| \frac{y_n - y_K}{n} \right| + \left| \frac{y_K}{n} \right| \leq \delta + \left| \frac{y_K}{n} \right| \leq 2\delta,
\]

provided \( n \) is chosen sufficiently large. \( \square \)

**Proof of Theorem 2.15.** We only derive the second claim of the theorem, Eq. (2.4). By assumptions, the offspring generating function is \( \varphi(s) = s + f(s - 1) \), where \( f(\cdot) \) satisfies

\[ f(0) = f'(0) = 0, \quad f''(0) = \sigma^2; \]
in other words, \( \varphi(s) \) is well approximated by \( s + \sigma^2(s - 1)^2/2 \) in a small enough neighbourhood of \( s = 1 \). Denote \( x_k = \mathbb{P}(Z_k = 0) = \varphi_k(0) \); recall that by criticality \( x_k \to 1 \) as \( k \to \infty \). We now have

\[
\frac{1}{1 - x_{k+1}} = \frac{1}{1 - \varphi(x_k)} = \frac{1}{1 - x_k} \left( 1 - \frac{f(x_k) - 1}{1 - x_k} \right)^{-1}
\]

\(^{14}\) Intermittency follows from the criticality condition, \( 1 = \mathbb{E}(Z_k \mid Z_k > 0)\mathbb{P}(Z_k > 0) \); it is the linearity which is surprising here!

\(^{15}\) Compare the result to Cesàro limits of real sequences: if \((a_k)_{k \geq 1}\) is a real-valued sequence, and \( s_n = a_1 + \cdots + a_n \) is its \( n \)th partial sum, then \( \frac{1}{n} s_n \) are called the Cesàro averages for the sequence \((a_k)_{k \geq 1}\). Lemma 2.16 claims that if \( a_k \to a \) as \( k \to \infty \), then the sequence of its Cesàro averages also converges to \( a \). The converse is, of course, false. (Find a counterexample!)
so that
\[ \frac{1}{1 - x_{k+1}} - \frac{1}{1 - x_k} = \frac{f(x_k - 1)}{(1 - x_k)^2} \left( 1 - \frac{f(x_k - 1)}{1 - x_k} \right) ^{-1} \rightarrow \frac{\sigma^2}{2} \]
(2.5)
as \( k \rightarrow \infty \). Noticing that the LHS above is just
\[ \mathbb{E}(Z_{k+1} \mid Z_{k+1} > 0) - \mathbb{E}(Z_k \mid Z_k > 0), \]
we deduce the claim of the theorem from Lemma 2.16. \( \square \)

**Remark 2.16.1.** With a bit of extra work\(^\text{16}\) one can generalize (2.5) to
\[ \lim_{n \rightarrow \infty} n^{-1} \left( \frac{1}{1 - \varphi_n(s)} - \frac{1}{1 - s} \right) = \frac{\sigma^2}{2} \]
and use this relation to derive the convergence in distribution:

**Theorem 2.17.** If \( \mathbb{E}Z_1 = 1 \) and \( \text{Var}(Z_1) = \sigma^2 \in (0, \infty) \), then for every \( z \geq 0 \) we have
\[ \lim_{k \rightarrow \infty} \mathbb{P} \left( \frac{Z_k}{k} > z \mid Z_k > 0 \right) = \exp \left\{ \frac{-2z}{\sigma^2} \right\}, \]
IE., the distribution of \( (k^{-1}Z_k \mid Z_k > 0) \) is approximately exponential with parameter \( 2/\sigma^2 \).

**Remark 2.17.1.** In the setup of Example 2.13, we have
\[ \mathbb{P}(Z_k > m \mid Z_k > 0) = \left( \frac{k}{k+1} \right)^m = \left( 1 - \frac{1}{k+1} \right)^m, \]
so that \( \mathbb{P}(Z_k > kz \mid Z_k > 0) \rightarrow e^{-z} \) as \( k \rightarrow \infty \); in other words, for large \( k \) the distribution of \( (k^{-1}Z_k \mid Z_k > 0) \) is approximately \( \text{Exp}(1) \).

**Exercise 2.18.** Check carefully that the result of the last theorem holds for the critical branching process from Exercise 2.14.

### 2.3 Non-homogeneous case

If the offspring distribution changes with time, the previous approach must be modified. Let \( \psi_n(u) \) be the generating function of the offspring distribution of a single ancestor in the \( (n-1) \)st generation,
\[ \psi_n(u) = \mathbb{E}(u^{Z_n} \mid Z_{n-1} = 1); \]
then the generating function \( \varphi_n(u) = \mathbb{E}(u^{Z_n} \mid Z_0 = 1) \) of the population size at time \( n \) given a single ancestor at time 0, can be defined recursively as follows:
\[ \varphi_0(u) \equiv u, \quad \varphi_k(u) = \varphi_{k-1}(\psi_n(u)), \quad \forall k \geq 1. \]
If \( \mu_n = \mathbb{E}(Z_n \mid Z_{n-1} = 1) = \psi_n(1) \) denotes the average offspring size in the \( n \)th generation given a single ancestor in the previous generation, then
\[ m_n \equiv \mathbb{E}(Z_n \mid Z_0 = 1) = \mu_1\mu_2 \cdots \mu_{n-1}\mu_n. \]
It is natural to call the process \( (Z_n)_{n \geq 0} \) supercritical if \( m_n \rightarrow \infty \) and subcritical if \( m_n \rightarrow 0 \) as \( n \rightarrow \infty \).

\(^{16}\)using the fact that every \( s \in (0, 1) \) satisfies \( 0 < s < \varphi_k(0) < 1 \) for some \( k \geq 1 \);
Exercise 2.19. A strain of phototrophic bacteria uses light as the main source of energy. As a result individual organisms reproduce with probability mass function
\[ p_0 = 1/4, \quad p_1 = 1/4 \quad \text{and} \quad p_2 = 1/2 \] per unit of time in light environment, and with probability mass function \( p_0 = 1 - p, \quad p_1 = p \) (with some \( p > 0 \)) per unit of time in dark environment. A colony of such bacteria is grown in a laboratory, with alternating light and dark unit time intervals.

a) Model this experiment as a time non-homogeneous branching process \((Z_n)_{n \geq 0}\) and describe the generating function of the population size at the end of the \( n \)th interval.

[Hint: Notice that the behaviour of your system differs for odd and even time intervals!]

b) Characterise all values of \( p \) for which the branching process \( Z_n \) is subcritical and for which it is supercritical.

c) Let \((D_k)_{k \geq 0}\) be the original process observed at the end of each even interval, \( D_k \overset{\text{def}}{=} Z_{2k} \). Find the generating function of \((D_k)_{k \geq 0}\) and derive the condition for sure extinction. Compare your result with that of part b).

2.4 Two-type branching processes

Consider a process in which a single type I individual gives birth to \( \xi^1 \) individuals of type I and \( \eta^1 \) individuals of type II, while a single type II individual gives birth to \( \xi^2 \) individuals of type I and \( \eta^2 \) individuals of type II. Let
\[ P(\xi^1 = k, \eta^1 = l) = p_1(k,l), \quad P(\xi^2 = k, \eta^2 = l) = p_2(k,l), \quad k, l \geq 0, \]
be the offspring distributions for type I and type II ancestors.

Assuming, as before, that each individual reproduces independently of all other individuals, and writing \( U_n \) and \( V_n \) for the respective numbers of type I and type II individuals, we get
\[ U_{n+1} = \sum_{j=1}^{U_n} \xi_j^1 + \sum_{j=1}^{V_n} \xi_j^2, \quad V_{n+1} = \sum_{j=1}^{U_n} \eta_j^1 + \sum_{j=1}^{V_n} \eta_j^2, \]
where \( (\xi_j^i, \eta_j^i)_{j \geq 1} \) are i.i.d. random vectors with probability mass functions \( p_i(k,l) \) as above. The simplest situation arises for initial conditions
\[ (U_0, V_0) = (1,0) \quad \text{or} \quad (U_0, V_0) = (0,1). \]

Consider the single step generating functions (starting from a single type \( i \) individual at time \( n = 0 \))
\[ \varphi^{(i)}(s,t) = E[ \sum_{k,l \geq 0} p_i(k,l)s^kt^l ], \quad i \in \{1,2\}, \]
and their multiple step generalisations
\[ \varphi^{(1)}(s,t) = \sum_{k,l \geq 0} P(U_n = k, V_n = l \mid U_0 = 1, V_0 = 0)s^kt^l, \]
\[ \varphi^{(2)}(s,t) = \sum_{k,l \geq 0} P(U_n = k, V_n = l \mid U_0 = 0, V_0 = 1)s^kt^l \]

http://maths.dur.ac.uk/stats/courses/StochProc34/
with \( n \geq 0 \). Clearly,
\[
\varphi_0^{(1)}(s,t) \equiv s, \quad \varphi_0^{(2)}(s,t) \equiv t,
\]
\[
\varphi_1^{(1)}(s,t) \equiv \varphi^{(1)}(s,t), \quad \varphi_1^{(2)}(s,t) \equiv \varphi^{(2)}(s,t),
\]
and the key observation is that
\[
\varphi_{n+m}^{(i)}(s,t) = \varphi_m^{(i)}(\varphi_n^{(1)}(s,t), \varphi_n^{(2)}(s,t))
\]
for \( i \in \{1, 2\} \) and \( n, m \geq 0 \). Indeed, it follows by induction from
\[
\varphi_{k+1}^{(i)}(s,t) \equiv \mathbb{E}^i(s^{U_{k+1}}t^{V_{k+1}}) = \mathbb{E}^i\left(\mathbb{E}^i(s^{U_{k+1}}t^{V_{k+1}} \mid U_k, V_k)\right)
\]
\[
= \mathbb{E}^i\left((\varphi^{(1)}(s,t))^{U_k}(\varphi^{(2)}(s,t))^{V_k}\right) = \varphi_k^{(i)}(\varphi^{(1)}(s,t), \varphi^{(2)}(s,t)).
\]

Write \( X_n = (U_n, V_n) \) for the population composition of the \( n \)th generation and consider the expectations
\[
m_{11} = \mathbb{E}(U_1 \mid X_0 = (1, 0)) = \mathbb{E}\xi^1 = \frac{\partial}{\partial s}\varphi^{(1)}(s,t) \bigg|_{(1,1)},
\]
\[
m_{12} = \mathbb{E}(V_1 \mid X_0 = (1, 0)) = \mathbb{E}\eta^1 = \frac{\partial}{\partial t}\varphi^{(1)}(s,t) \bigg|_{(1,1)},
\]
\[
m_{21} = \mathbb{E}(U_1 \mid X_0 = (0, 1)) = \mathbb{E}\xi^2 = \frac{\partial}{\partial s}\varphi^{(2)}(s,t) \bigg|_{(1,1)},
\]
\[
m_{22} = \mathbb{E}(V_1 \mid X_0 = (0, 1)) = \mathbb{E}\eta^2 = \frac{\partial}{\partial t}\varphi^{(2)}(s,t) \bigg|_{(1,1)},
\]

ie., \( m_{i1i2} \) is the expected number of type \( i_2 \) individuals starting from a single type \( i_1 \) ancestor. Write
\[
M = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}.
\]

**Lemma 2.20.** For \( n, k \geq 0 \),
\[
\mathbb{E}(X_{n+k} \mid X_n) = X_n \cdot M^k.
\]

**Proof.** This follows easily by induction in \( k \) using
\[
F_n(s,t) \equiv \mathbb{E}(s^{U_{n+1}}t^{V_{n+1}} \mid U_n, V_n) = (\varphi^{(1)}(s,t))^{U_n}(\varphi^{(2)}(s,t))^{V_n}
\]
so that
\[
\mathbb{E}(U_{n+1} \mid U_n, V_n) = \frac{\partial}{\partial s} F_n(s,t) \bigg|_{(1,1)} = U_n \cdot \mathbb{E}\xi^1 + V_n \cdot \mathbb{E}\xi^2 = (X_n \cdot M)_1.
\]
Similarly, \( \mathbb{E}(V_{n+1} \mid U_n, V_n) = (X_n \cdot M)_2. \)
As in the classical case of branching processes with individuals of single type, write \( E_2 \) for the extinction event \( \{ X_n = (0,0) \text{ for some } n \} \) and introduce the extinction probabilities
\[
\rho^1 = P( E_2 \mid X_0 = (1,0)) , \quad \rho^2 = P( E_2 \mid X_0 = (0,1)) .
\]

Then the following analogue of Theorem 2.4 holds.

**Theorem 2.21.** Let the offspring generating functions \( \varphi^{(i)}(s,t) \), \( i \in \{1,2\} \), be monotone in each of the variables \( s,t \) and let all expectations \( m_{ij} \) be positive. Then the extinction probabilities \( (\rho^1,\rho^2) \) are given by the smallest non-negative solution of the fixed point system
\[
\rho^1 = \varphi^{(1)}(\rho^1,\rho^2), \quad \rho^2 = \varphi^{(2)}(\rho^1,\rho^2) .
\]

Further, let \( r \geq 0 \) be the maximal eigenvalue of the matrix \( M \) of averages \( m_{i1,i2} \). Then for \( r \leq 1 \) we have \( \rho^1 = \rho^2 = 1 \), while for \( r > 1 \) we have \( \rho^1, \rho^2 < 1 \).

**Remark 2.21.1.** The proof of this result can be obtained in full analogy with the one-dimensional case, by studying the sequence \( q^i_n = \varphi^{(i)}(0,0) \), \( n \geq 0 \), of approximates. By monotonicity of \( \varphi^{(i)}(s,t) \), \( i \in \{1,2\} \), each \( (q^i_n)_{n \geq 0} \), is a monotone sequence of real numbers in \([0,1]\), and hence converges to a limit. By continuity, the limit must solve the system (2.6), and a suitable generalisation of the single type argument shows that \( (\rho^1,\rho^2) \) is given by its smallest non-negative solution.

Notice that the strict monotonicity of \( \varphi^{(i)}(s,t) \), \( i \in \{1,2\} \), implies that a fixed point of (2.6) with \( (\rho^1,\rho^2) \neq (1,1) \) must satisfy \( \rho^1, \rho^2 < 1 \). The explicit characterisation of these cases in terms of the maximal eigenvalue of the matrix \( M \) can be achieved using the famous Perron-Frobenius theorem describing the limiting behaviour of powers of matrices with non-negative entries.\(^{17}\)

**Remark 2.21.2.** If all entries \( m_{ij} \) are positive, it is easy to see that the maximal eigenvalue \( r \) is positive. Indeed, then the quadratic polynomial
\[
f(\lambda) = \det (M - \lambda I) = (m_{11} - \lambda)(m_{22} - \lambda) - m_{12}m_{21}
\]
satisfies \( f(m_{11}) = f(m_{22}) = -m_{12}m_{21} < 0 \) and thus must have a positive root.

Alternatively, notice that the trace of \( M \) is positive, \( m_{11} + m_{22} > 0 \); since it equals the sum of the eigenvalues, at least one of them must be positive.

**Exercise 2.22.** Consider a branching process with two particle types, type I and type II. Let \( \varphi^{(i)}(s,t) = E(s^{U_1} t^{V_1}) \) be the generating function of the offspring distribution of a single individual of type \( i \). Suppose that
\[
\varphi^{(1)}(s,t) = \frac{1}{2} + \frac{1}{8} s + \frac{1}{8} t + \frac{1}{8} st + \frac{1}{8} st^2 , \\
\varphi^{(2)}(s,t) = \frac{1}{4} + \frac{1}{4} t + \frac{1}{8} st + \frac{1}{8} st^2 + \frac{1}{4} s^2 t .
\]

Find the maximal eigenvalue of the matrix \( M \) of averages and determine whether this process become extinct with probability one or not.

\(^{17}\)we will not do this here; get in touch if interested!