14 Branching processes

In this chapter we will consider a random model for population growth in the absence of spatial or any other resource constraints. So, consider a population of individuals which evolves according to the following rule: in every generation $n = 0, 1, 2, \ldots$, each individual produces a random number of offspring in the next generation, independently of other individuals. The probability mass function for offspring is often called the *offspring distribution* and is given by

$$p_i = P(\text{number of offspring} = i),$$

for $i = 0, 1, 2, \ldots$ We will *assume* that $p_0 < 1$ and $p_1 < 1$ to eliminate the trivial cases. This model was introduced by F. Galton, in late 1800s, to study the disappearance of family names; in this case p_i is the probability that a man has i sons. We will start with a single individual in generation 0, and generate the resulting random *family tree*. This tree is either finite (when some generation produces no offspring at all) or infinite — in the former case, we say that the branching process *dies out*, and in the latter that it *survives*.

We can look at this process as a Markov chain, where X_n is the number of individuals at generation n. Let's start with the following observations.

- If X_n reaches 0, it stays there, so 0 is an absorbing state.
- Of $p_0 > 0$, $P(X_{n+1} = 0 | X_n = k) > 0$ for all k.
- Therefore, by Proposition 13.5, all states other than 0 are transient if $p_0 > 0$; the population must either die out or increase to infinity. If $p_0 = 0$, then the population cannot decrease, and increases each generation with probability at least $1 - p_1$, therefore must increase to infinity.

It is possible to write down the transition probabilities for this chain, but they have a rather complicated explicit form, as

$$P(X_{n+1} = i | X_n = k) \quad = \quad P(W_1 + W_2 + \ldots + W_k = i),$$

where $W_1, \ldots, W_k$ are independent random variables, each with the offspring distribution. This suggest the use of the moment generating functions, which we will indeed do. Recall that we are assuming that $X_0 = 1$.

Let

$$\delta_n = P(X_n = 0)$$

be the probability that the population is extinct by generation (which we also think of as time) n. The probability π_0 that the brandcing process dies out is then the limit of these probabilities:

$$\pi_0 = P(\text{the process dies out}) = P(X_n = 0 \text{ for some } n) = \lim_{n\to\infty} P(X_n = 0) = \lim_{n\to\infty} \delta_n.$$

Note that $\pi_0 = 0$ if $p_0 = 0$. Our main task will be to compute π_0 for general probabilities p_k. We start, however, with computing expectation and variance of the population at generation n.

Let μ and σ^2 be the expectation and variance of the offspring distribution, that is,

$$\mu = EX_1 = \sum_{k=0}^{\infty} kp_k,$$

and

$$\sigma^2 = \mathrm{Var}(X_1).$$

Let $m_n = E(X_n)$ and $v_n = \mathrm{Var}(X_n)$. Now, X_{n+1} is the sum of a random number, which equals X_n, of independent random variables, each with the offspring distribution. Thus we have, by Theorem 11.1,

$$m_{n+1} = m_n \mu,$$

and

$$v_{n+1} = m_n \sigma^2 + v_n \mu^2.$$

Together with initial conditions $m_0 = 1$, $v_0 = 0$, the two recursive equations determine m_n and v_n. We can very quickly solve the first recursion to get $m_n = \mu^n$, and so

$$v_{n+1} = \mu^n \sigma^2 + v_n \mu^2.$$

This recursion has a general solution $v_n = A\mu^n + B\mu^{2n}$. The constant A must satisfy

$$A\mu^{n+1} = \sigma^2 \mu^n + A\mu^{n+2},$$

so that, when $\mu \neq 1$,

$$A = \frac{\sigma^2}{\mu(1-\mu)}.$$

From $v_0 = 0$ we get $A + B = 0$ and the solution is given in the next theorem.

Theorem 14.1. *Expectation m_n and variance v_n of the n'th generation count.*

We have

$$m_n = \mu^n$$

and

$$v_n = \begin{cases} \frac{\sigma^2}{\mu(1-\mu)}\mu^n - \frac{\sigma^2}{\mu(1-\mu)}\mu^{2n} & \text{if } \mu \neq 1, \\ n\sigma^2 & \text{otherwise.} \end{cases}$$

We can immediately conclude that $\mu < 1$ implies $\pi_0 = 1$, as

$$P(X_n \neq 0) = P(X_n \geq 1) \leq EX_n = \mu^n \to 0.$$

That is, if the individuals have less than one offspring on the average, the branching process dies out.

Let now ϕ be the moment generating function of the offspring distribution. It is more convenient to replace e^t in our original definition with s, so that

$$\phi(s) = \phi_{X_1}(s) = E(s^{X_1}) = \sum_{k=0}^{\infty} p_k s^k.$$

In combinatorics, this would be exactly the generating function of the sequence p_k. The moment generating function of X_n then is

$$\phi_{X_n}(s) = E[s^{X_n}] = \sum_{k=0}^{\infty} P(X_n = k)s^k.$$

We will assume that $0 \leq s \leq 1$, and observe that for such s this power series converges. Let's try to get a recursive equation for ϕ_{X_n}, by conditioning on the population count at generation $n-1$:

$$\begin{aligned}
\phi_{X_n}(s) &= E[s^{X_n}] \\
&= \sum_{k=0}^{\infty} E[s^{X_n} | X_{n-1} = k] P(X_{n-1} = k) \\
&= \sum_{k=0}^{\infty} E[s^{W_1+\ldots+W_k}] P(X_{n-1} = k) \\
&= \sum_{k=0}^{\infty} (E(s^{W_1}) E(s^{W_2}) \cdots E(s^{W_k}) P(X_{n-1} = k) \\
&= \sum_{k=0}^{\infty} \phi(s)^k P(X_{n-1} = k) \\
&= \phi_{X_{n-1}}(\phi(s)).
\end{aligned}$$

So ϕ_{X_n} is the n'th iterate of ϕ,

$$\phi_{X_2}(s) = \phi(\phi(s)),\ \phi_{X_3}(s) = \phi(\phi(\phi(s))),\ \ldots$$

and we can also write

$$\phi_{X_n}(s) = \phi(\phi_{X_{n-1}}(s)).$$

We will take a closer look at the properties of ϕ next. Clearly,

$$\phi(0) = p_0 > 0,$$

and

$$\phi(1) = \sum_{k=0}^{\infty} p_k = 1.$$

Moreover,

$$\phi'(s) = \sum_{k=0}^{\infty} k p_k s^{k-1} \geq 0,$$

so ϕ is increasing, with

$$\phi'(1) = \mu.$$

Finally,

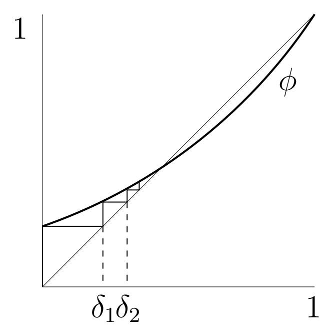
$$\phi''(s) = \sum_{k=1}^{\infty} k(k-1)p_k s^{k-2} \geq 0,$$

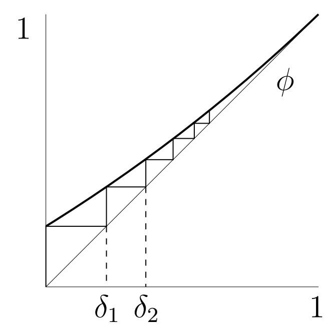
so ϕ is also convex. The crucial observation is that

$$\delta_n = \phi_{X_n}(0),$$

and so it is obtained by starting at 0 and computing the n'th iteration of ϕ. It is also clear that δ_n is a nondecreasing sequence (because $X_{n-1} = 0$ implies that $X_n = 0$). We now consider separately two cases:

- Assume that ϕ is always above the diagonal, $\phi(s) \geq s$ for all $s \in [0,1]$. This happens exactly when $\mu = \phi'(1) \leq 1$. In this case δ_n converges to 1, and so $\pi_0 = 1$. This is the case in the right graph of the figure below.
- Now assume that ϕ is not always above the diagonal, which happens when $\mu > 1$. In this case there exists exactly one $s < 1$ which solves $s = \phi(s)$. As δ_n converges to this solution, we conclude that $\pi_0 < 1$ is now the smallest solution to $s = \phi(s)$. This is the case in the left graph of the figure below.





Next theorem is the summary of our findings.

Theorem 14.2. *Probability that the branching process dies out.*

$$\boxed{\text{If } \mu \leq 1,\ \pi_0 = 1. \text{ If } \mu > 1, \text{ then } \pi_0 \text{ is the smallest solution on } [0,1] \text{ to } s = \phi(s).}$$

Example 14.1. Assume that a branching process is started with $X_0 = k$ instead of $X_0 = 1$. How does this change the survival probability? The k individuals all evolve independent family trees, so that the probability of eventual death is π_0^k. It follows that also

$$P(\text{the process ever dies out} | X_n = k) = \pi_0^k$$

for every n.

If μ is barely larger than 1, the probability π_0 of extinction is quite close to 1. In the context of family names, this means that the ones with already large number of representatives in the population are at a distinct advantage, as the probability that they die out by chance is much lower than that of those with only a few representatives. Thus common family names become ever more common, especially in societies that have used family names for a long time. The most famous example of this phenomenon is Korea, where three family names (Kim, Lee, and Park in common English transcriptions) account for about 45% of the population.

Example 14.2. Assume that

$$p_k = p^k(1-p),\ k = 0, 1, 2, \ldots$$

This means that the offspring distribution is Geometric($\frac{1}{1-p}$), minus 1. Thus

$$\mu = \frac{1}{1-p} - 1 = \frac{p}{1-p},$$

and if $p \le \frac{1}{2}$, $\pi_0 = 1$. Suppose now that $p > \frac{1}{2}$. Then we have to compute

$$\begin{aligned} \phi(s) &= \sum_{k=0}^{\infty} s^k p^k (1-p) \\ &= \frac{1-p}{1-ps}. \end{aligned}$$

The equation $\phi(s) = s$ has two solutions, $s = 1$ and $s = \frac{1-p}{p}$. Thus, when $p > \frac{1}{2}$.

$$\pi_0 = \frac{1-p}{p}.$$

Example 14.3. Assume that the offspring distribution is Binomial($3, \frac{1}{2}$). Compute π_0.

As $\mu = \frac{3}{2} > 1$, π_0 is given by

$$\phi(s) = \frac{1}{8} + \frac{3}{8}s + \frac{3}{8}s^2 + \frac{1}{8}s^3 = s,$$

with solutions $s = 1$, $-\sqrt{5} - 2$, and $\sqrt{5} - 2$. The one that lies in $(0, 1)$, $\sqrt{5} - 2 \approx 0.2361$, is the probability π_0.

Problems

1. For a branching process with offspring distribution given by $p_0 = \frac{1}{6}, p_1 = \frac{1}{2}, p_3 = \frac{1}{3}$, determine (a) expectation and variance of X_9, the population at generation 9, (b) probability that the branching process dies by generation 3, but not by generation 2, and (c) the probability that

the process ever dies out. Then assume you start 5 independent copies of this branching process at the same time (equivalently, change X_0 to 5), and (d) compute the probability the that the process ever dies out.

2. Assume that the offspring distribution of a branching process is Poisson with parameter λ. (a) Determine the expected *combined* population through generation 10. (b) Determine, with the aid of computer if necessary, the probability that the process ever dies out for $\lambda = \frac{1}{2}$, $\lambda = 1$ and $\lambda = 2$.

3. Assume that the offspring distribution of a branching process is given by $p_1 = p_2 = p_3 = \frac{1}{3}$. Note that $p_0 = 0$. Solve the following problem for $a = 1, 2, 3$. In a generation n, choose a random individual and let Y_n be the proportion of individuals, among X_n, from families of size a. (A family consists of individuals that are offspring of the same parent from previous generation.) Compute the limit of Y_n as $n \to \infty$.

Solutions to problems

1. For (a), compute $\mu = \frac{3}{2}$, $\sigma^2 = \frac{7}{2} - \frac{9}{4} = \frac{5}{4}$, and plug into the formula. Then compute

$$\phi(s) = \frac{1}{6} + \frac{1}{2}s + \frac{1}{3}s^3.$$

For (b),

$$P(X_3 = 0) - P(X_2 = 0) = \phi(\phi(\phi(0))) - \phi(\phi(0)) \approx 0.0462.$$

For (c), we solve $\phi(s) = s$, $0 = 2s^3 - 3s + 1 = (s-1)(2s^2 + 2s - 1)$, and so $\pi_0 = \sqrt{3} - 12 \approx 0.3660$. For (d), the answer is π_0^5.

2. For (a),

$$E(X_0 + X_1 + \ldots + X_{10}) = EX_0 + EX_1 + \ldots + EX_{10} = 1 + \mu + \cdots + \mu^{10} = \frac{\mu^{11} - 1}{\mu - 1},$$

if $\mu \neq 1$, and 11 if $\mu = 1$. For (b), if $\lambda \leq 1$ then $\pi_0 = 1$, while if $\lambda > 1$, then π_0 is the solution in $s \in (0, 1)$ to

$$e^{\lambda(s-1)} = s.$$

This equation cannot be solved analytically, but we can numerically obtained the solution for $\lambda = 2$, $\pi_0 = 0.2032$.

3. Assuming $X_{n-1} = k$, this is the number of families at time n. Each of these has, independently, a members with probability p_a. If k is large, which it will be for large n, as the branching process can't die out, then with overwhelming probability the number of children in such families is about $ap_a k$, while X_n is about μk. The proportion Y_n then is about $\frac{ap_a}{\mu}$, which works out to be $\frac{1}{6}$, $\frac{1}{3}$, and $\frac{1}{2}$, for $a = 1$, 2, and 3.