

## Chapter 2

# Discrete Time Branching Process

We introduce branching processes. They are a recurring theme throughout the book. In this chapter we use them to model drug resistance and cancer risk.

### 1 The Model

In this chapter we are concerned with modeling population growth. Typically we start the population with a single individual. This individual has a random number of offspring. Each child has itself a number of offspring and so on. The first question is about survival. Can such a process survive forever? Can we compute the survival probability? In order to be able to do computations we need to make some assumptions. We will assume that different individuals have independent but identical offspring distributions. Hence, all individuals are put in the same condition (offspring wise) and we get a branching (or cascading) effect. We will concentrate on biological applications. This process turns out to be also a good model for a number of physical phenomena.

This stochastic process was introduced independently by Bienaymé (who got the mathematics right and was forgotten for many years) and by Galton and Watson (who got the mathematics wrong but got their names attached to this process) to model the survival of family names. An initial set of individuals which we call the zeroth generation have a number of offspring that are called the first generation; their offspring are called the second generation and so on. We denote the size of the  $n$ th generation by  $Z_n$ ,  $n \geq 0$ .

We now give the mathematical definition of the Bienaymé–Galton–Watson (BGW) process  $(Z_n)_{n \geq 0}$ . The state space  $\mathcal{S}$  of  $(Z_n)_{n \geq 0}$  is the set of positive (including zero) integers. We suppose that each individual gives birth to  $Y$  particles in the next generation where  $Y$  is a positive integer-valued random variable with distribution  $(p_k)_{k \geq 0}$ . In other words

$$P(Y = k) = p_k, \text{ for } k = 0, 1, \dots$$

Moreover we assume that the number of offspring of the various individuals in the various generations are chosen independently according to the distribution  $(p_k)_{k \geq 0}$ .

The process is governed by the so-called one-step transition probabilities

$$p(i, j) = P(Z_{n+1} = j | Z_n = i).$$

That is,  $p(i, j)$  is the conditional probability that  $Z_{n+1} = j$  given that  $Z_n = i$ . We also have

$$p(0, i) = 0 \text{ if } i \geq 1 \text{ and } p(0, 0) = 1.$$

That is, once the process is at 0 (or extinct) it stays there. State 0 (no individuals) is also said to be an absorbing state (or trap) for  $(Z_n)_{n \geq 0}$ .

Observe that

$$p(i, j) = P(Z_{n+1} = j | Z_n = i) = P\left(\sum_{k=1}^i Y_k = j\right) \text{ for } i \geq 1, j \geq 0,$$

where  $(Y_k)_{1 \leq k \leq i}$  is a sequence of independent identically distributed (i.i.d.) random variables with distribution  $(p_k)_{k \geq 0}$ . This shows that the distribution of  $Z_{n+1}$  can be computed using the distribution of  $Z_n$  only. That is, there is no need to know the complete history of the process given by  $Z_0, Z_1, \dots, Z_n$  in order to compute the distribution of  $Z_{n+1}$ . It is enough to know  $Z_n$ . This is called the Markov property.

A word on notation. For  $Z_n = i$  we should have written  $Z_{n+1}$  as

$$Z_{n+1} = \sum_{k=1}^i Y_{k,n} = j$$

where  $(Y_{k,n})_{1 \leq k \leq i}$  is an i.i.d. sequence. The subscript  $n$  is to indicate that we use a different independent sequence for every  $n$ . We omit the  $n$  in the notation to avoid a double index.

Let the mean offspring be

$$m = \sum_{k=0}^{\infty} k p_k,$$

where  $m$  is possibly  $+\infty$  if the series does not converge. Let  $q$  be the probability that the BGW process starting from a single individual eventually dies out. We also introduce the generating function of the offspring distribution

$$f(s) = \sum_{k=0}^{\infty} p_k s^k \text{ for } |s| \leq 1.$$

We now state the main result of this chapter. The process will be said to survive if there exists at least one individual for every generation  $n$ . Mathematically, surviving means

$$\{Z_n \geq 1, \text{ for all } n \geq 0\}.$$

**Theorem 1.1.** *Let  $(Z_n)_{n \geq 0}$  be a BGW process with offspring distribution  $(p_k)_{k \geq 0}$ . Assume that  $p_0 + p_1 < 1$ .*

*If  $m \leq 1$ , then  $P(Z_n \geq 1, \text{ for all } n \geq 0 | Z_0 = 1) = 0$ .*

*If  $m > 1$ , there exists  $q$  in  $[0, 1)$  such that  $P(Z_n \geq 1, \text{ for all } n \geq 0 | Z_0 = 1) = 1 - q > 0$ . Moreover  $q$ , the extinction probability, is the unique solution in  $[0, 1)$  of the equation  $f(s) = s$  when  $m > 1$ .*

The process BGW is said to be subcritical, critical, and supercritical according to whether  $m < 1$ ,  $m = 1$ , or  $m > 1$ . Observe that the BGW process may survive forever if and only if  $m > 1$ . So the only relevant parameter of the offspring distribution for survival is  $m$ . However, the probability  $1 - q$  of surviving forever depends on the whole distribution  $(p_k)_{k \geq 1}$  through its generating function.

The proof of Theorem 1.1 will be given in the last section of this chapter.

It is useful to have a graphical representation of the process  $(Z_n)_{n \geq 0}$ . See Fig. 2.1. Survival of the process corresponds to an infinite tree. Death of a process corresponds to a finite tree.

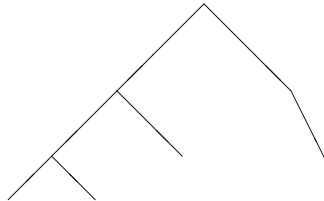
We now apply Theorem 1.1 to a few examples.

**Example 1.1.** Consider a BGW process with the offspring distribution  $P(Y = 0) = p_0 = 1/6$ ,  $P(Y = 1) = p_1 = 1/2$ , and  $P(Y = 2) = p_2 = 1/3$ . We first compute the average offspring per individual.

$$m = E(Y) = 7/6 > 1.$$

So the survival probability  $1 - q$  is strictly positive in this case. The generating function of  $Y$  is

$$f(s) = 1/6 + s/2 + s^2/3.$$



**Fig. 2.1** This is a graphical representation of a BGW. The process starts with one individual at the top of the tree. We see that  $Z_0 = 1$ ,  $Z_1 = 2$ ,  $Z_2 = 3$  and  $Z_3 = 4$

The extinction probability  $q$  is the only solution strictly less than 1 of  $f(s) = s$ . This equation can be written as

$$\frac{1}{6} - \frac{1}{2}s + \frac{1}{3}s^2 = \frac{1}{3}(s-1)(s-\frac{1}{2}) = 0.$$

There are two solutions  $s = 1$  and  $s = 1/2$ . We know that  $q$  is the solution in  $[0, 1)$ . Hence,  $q = 1/2$ . So starting with a single particle there is a probability  $1/2$  that the process will survive forever.

*Example 1.2.* Lotka (1939) has used a geometric distribution to fit the offspring of the American male population. He found that

$$p_0 = P(Y = 0) = 1/2 \text{ and } p_i = P(Y = i) = \left(\frac{3}{5}\right)^{i-1} \frac{1}{5} \text{ for } i \geq 1,$$

where  $Y$  represents the number of sons that a male has in his lifetime. Recall that

$$\sum_{n \geq 0} x^n = \frac{1}{1-x} \quad \sum_{n \geq 1} nx^{n-1} = \frac{1}{(1-x)^2} \text{ for } |x| < 1.$$

So

$$m = \sum_{n \geq 1} np_n = \sum_{n \geq 1} n \left(\frac{3}{5}\right)^{n-1} \frac{1}{5} = \frac{5}{4} > 1.$$

Hence, the extinction probability is strictly less than 1 and is a solution of  $f(s) = s$  where

$$f(s) = \frac{1}{2} + \sum_{n \geq 1} \left(\frac{3}{5}\right)^{n-1} \frac{1}{5} s^n = \frac{1}{2} + \frac{s}{5-3s}.$$

Solving the equation  $f(s) = s$  yields

$$\frac{3}{5}s^2 - \frac{11}{10}s + \frac{1}{2} = 0.$$

As always  $s = 1$  is a solution. The unique root strictly less than 1 is  $q = 5/6$ . So under this model a given male has a probability of  $1/6$  of generating a family that survives forever.

**Proposition 1.1.** Assume that  $m$  (the mean offspring) is finite. We have

$$E(Z_n | Z_0 = 1) = m^n \text{ for } n \geq 0.$$

*Proof of Proposition 1.1.* We do a proof by induction. Note that

$$E(Z_1 | Z_0 = 1) = E(Y) = m = m^1.$$

Hence, the formula holds for  $n = 1$ . Assume now that it holds for  $n$ . By conditioning on  $Z_n$  we get

$$E(Z_{n+1}|Z_0 = 1) = \sum_{k \geq 1} E(Z_{n+1}|Z_n = k)P(Z_n = k|Z_0 = 1),$$

where by the Markov property we are using that

$$E(Z_{n+1}|Z_0 = 1, Z_n = k) = E(Z_{n+1}|Z_n = k).$$

For every  $k \geq 1$ ,

$$E(Z_{n+1}|Z_n = k) = E\left(\sum_{i=1}^k Y_i\right) = km.$$

Thus,

$$E(Z_{n+1}|Z_0 = 1) = \sum_{k \geq 1} kmP(Z_n = k|Z_0 = 1) = mE(Z_n|Z_0 = 1).$$

Since by the induction hypothesis we have  $E(Z_n|Z_0 = 1) = m^n$  we can conclude that

$$E(Z_{n+1}|Z_0 = 1) = mm^n = m^{n+1}.$$

This completes the proof of Proposition 1.1.

## Problems

1. Consider a BGW process with offspring distribution

$$P(Y = 0) = 1 - p \text{ and } P(Y = 2) = p$$

where  $p$  is a fixed parameter in  $(0, 1)$ .

- Compute the mean offspring  $m$ .
- For which  $p$  does the process have a positive probability of surviving?
- Sketch the extinction probability  $q$  as a function of  $p$ .

2. Consider a BGW with offspring distribution  $(p_k)_{k \geq 0}$  and extinction probability  $q$ .

- Show that  $q \geq p_0$ .
- Show that  $q = 0$  (survival is certain) if and only if  $p_0 = 0$ .

3. Let  $r$  be in  $(0, 1)$ . Consider a BGW with the offspring distribution

$$p_k = (1 - r)r^k \text{ for } k \geq 0.$$

Find the extinction probability  $q$  as a function of  $r$ .

4. Find the extinction probability  $q$  if the offspring distribution  $(p_k)_{k \geq 0}$  is given by

$$p_k = \binom{3}{k} (1/2)^3$$

for  $k = 0, 1, 2, 3$ .

5. Consider a BGW with mean offspring  $m$ .

(a) Show that

$$P(Z_n \geq 1) \leq E(Z_n).$$

(b) Assume that the mean offspring  $m = 1/2$ . Show that the probability that  $Z_n$  has survived ten generations is less than  $\frac{1}{2^{10}}$ .

6. Redo Example 1.2 with a truncated geometric distribution. More precisely, take  $p_0 = P(Y = 0) = 1/2$  and

$$p_i = P(Y = i) = \left(\frac{3}{5}\right)^{i-1} c \text{ for } 1 \leq i \leq 10.$$

(a) Find  $c$ .

(b) Compute (approximately) the extinction probability.

7. Find approximately the extinction probability in the case  $p_k = e^{-2} 2^k / k!$  for  $k \geq 0$ .

8. Discuss the behavior of  $Z_n$  in the case  $p_0 + p_1 = 1$ .

9. Consider an isolated island where the original stock of surnames is 100. Assume that each surname has an extinction probability  $q = 9/10$ .

(a) After many generations how many surnames do you expect in the island?

(b) Do you expect the total population of the island to be increasing or decreasing?

10. Consider a supercritical BGW for which the extinction probability  $q = 1/2$ . Start the process with five particles. What is the probability that the process will survive forever?

11. Consider a supercritical BGW with offspring distribution  $(p_k)_{k \geq 0}$ . Let the extinction probability be  $q$ .

- (a) Show that for every integer  $k \geq 0$  we have

$$P(\text{extinction} | Z_1 = k) = q^k.$$

- (b) Show that

$$P(Z_1 = k | Z_0 = 1) = p_k.$$

- (c) Use (a) and (b) to show that

$$P(\text{extinction} | Z_0 = 1) = \sum_{k=0}^{\infty} p_k q^k.$$

- (d) Use (c) to show that  $q$  is a solution of the equation  $f(s) = s$  where  $f$  is the generating function of the offspring distribution.

**12.** Let  $f$  be the generating function of the probability distribution  $(p_k)_{k \geq 0}$ .

- (a) Show that  $f$  is increasing on  $[0, 1]$ .
- (b) Show that  $f(0) \geq 0$  and  $f(1) = 1$ .
- (c) Show that  $f$  is concave up on  $[0, 1]$ .
- (d) Graph three functions with properties (a), (b), and (c) and  $f'(1) < 1$ ,  $f'(1) = 1$  and  $f'(1) > 1$ , respectively.
- (e) Using the graphs in (d) show that the equation  $f(s) = s$  has a solution  $0 \leq q < 1$  if and only if  $f'(1) > 1$ .
- (f) How does (e) relate to Theorem 1.1?

**13.** Consider a BGW process  $Z_n$  with offspring distribution

$$P(Y = 0) = 1 - p \text{ and } P(Y = 2) = p$$

where  $p$  is a fixed parameter in  $(0, 1)$ . A simulation of the offspring distribution  $Y$  with  $p = 3/4$  has yielded the following observations 2,2,0,2,2,0,2,2,2,0.

- (a) Use the simulation to graph the corresponding BGW tree.
- (b) Use the simulation to find  $Z_n$  for as many  $n$  as possible.

**14.** Consider the following algorithm. Let  $U$  be a continuous uniform random variable on  $(0, 1)$ . If  $U < p$  set  $Y = 2$ , if  $U > p$  set  $Y = 0$ .

Show that this algorithm generates a random variable  $Y$  with the distribution

$$P(Y = 0) = 1 - p, P(Y = 2) = p.$$

(Recall that  $P(U < x) = x$  for  $x$  in  $(0, 1)$ .)

**15.** Use Problem 13 to simulate a BGW process  $Z_n$  with offspring distribution

$$P(Y = 0) = 1 - p \text{ and } P(Y = 2) = p,$$

for  $p = 1/4$ ,  $p = 1/2$  and  $p = 3/4$ . Do multiple runs. Interpret the results.

## 2 The Probability of a Mutation in a Branching Process

Consider the following question. We have a population modeled by a BGW process. Each time an individual is born in the population it has a probability  $\mu$  of having a certain mutation. We want to compute the probability that this mutation will eventually appear in the population. We will actually do the computation for a particular BGW and we will need several steps.

### 2.1 An Equation for the Total Progeny Distribution

We will start by finding the probability distribution of

$$X = \sum_{n \geq 0} Z_n = 1 + \sum_{n \geq 1} Z_n.$$

The random variable  $X$  counts all the births that ever occur in the population (i.e., the total progeny) as well as the founding individual (as usual we are taking  $Z_0 = 1$ ). Our first goal is to find the distribution of  $X$ . The main step is to find the generating function  $g$  of  $X$ . Let

$$g(s) = \sum_{k \geq 1} s^k P(X = k) = E(s^X).$$

The sum above starts at 1 since  $X$  is always at least 1. We know that a generating function is always defined (i.e., the power series converges) for  $s$  in  $[0, 1]$ .

We will find an equation for  $g$  by conditioning on the first generation  $Z_1$ . We have

$$g(s) = E(s^X) = \sum_{k \geq 0} E(s^X | Z_1 = k) P(Z_1 = k | Z_0 = 1) \quad (2.1)$$

Given that  $Z_1 = k$  we have  $k$  individuals at time 1 starting  $k$  independent BGW. Hence,  $X$  is the sum of the founding individual and all the individuals in these  $k$  BGW. Moreover, for each one of these  $k$  BGW the total number of births (plus the founding individual) has the same distribution as  $X$ . Therefore,

$$E(s^X | Z_1 = k) = E(s^{1+X_1+X_2+\dots+X_k})$$

where the  $X_i$ ,  $1 \leq i \leq k$ , are independent random variables with the same distribution as  $X$ . Thus, by independence

$$E(s^X | Z_1 = k) = s E(s^{X_1}) E(s^{X_2}) \dots E(s^{X_k}).$$



Since the  $X_i$ ,  $1 \leq i \leq k$ , have the same distribution as  $X$  we get

$$E(s^X | Z_1 = k) = sE(s^X)^k.$$

Using the last equality in (2.1) yields

$$g(s) = E(s^X) = \sum_{k \geq 0} sE(s^X)^k P(Z_1 = k | Z_0 = 1) = s \sum_{k \geq 0} g(s)^k p_k$$

since

$$P(Z_1 = k | Z_0 = 1) = P(Y = k) = p_k.$$

Hence,

$$g(s) = s \sum_{k \geq 0} p_k g(s)^k.$$

Observe now that the generating function  $f$  of the offspring distribution is

$$f(s) = \sum_{k \geq 0} p_k s^k.$$

Therefore,

$$g(s) = sf(g(s)) \tag{2.2}$$

This is an elegant equation for  $g$  but for most offspring distributions (represented by  $f$ ) it cannot be solved. Next we solve the equation in a particular case.

## 2.2 The Total Progeny Distribution in a Particular Case

Consider the following offspring distribution:  $p_0 = 1 - p$  and  $p_2 = p$  where  $p$  is a fixed parameter in  $[0, 1]$ . That is, for  $n \geq 0$ , every individual of the  $n$ th generation gives birth to two individuals with probability  $p$  or does not give birth at all with probability  $1 - p$ . This yields the  $(n + 1)$ th generation. Biologically, we are thinking of a population of bacteria or virus that either divide into two individuals (corresponding to births in the model) or die.

Recall that  $g$  is the generating function of  $X$ :

$$g(s) = \sum_{n=1}^{\infty} P(X = 2n - 1 | Z_0 = 1) s^{2n-1}.$$

We take into account only the odd values for  $X$  because the births that occur always occur by pairs and since we start with a single individual the random variable  $X$  is odd.

By (2.2)  $g$  is the solution of the equation

$$g(s) = sf(g(s)),$$

where  $f$  is the generating function of the offspring distribution. For this particular case

$$f(s) = 1 - p + ps^2.$$

Hence,

$$g(s) = s(1 - p + pg(s)^2),$$

and

$$psg(s)^2 - g(s) + s(1 - p) = 0.$$

For a fixed  $s$  let  $g(s) = u$ , we get

$$psu^2 - u + s(1 - p) = 0$$

and

$$u^2 - \frac{1}{ps}u + \frac{s(1 - p)}{ps} = 0.$$

This is a quadratic equation in  $u$ . We complete the square to get

$$\left(u - \frac{1}{2ps}\right)^2 - \frac{1}{4p^2s^2} + \frac{s(1 - p)}{ps} = 0.$$

Hence,

$$\left(u - \frac{1}{2ps}\right)^2 = \frac{1}{4p^2s^2} - \frac{s(1 - p)}{ps} = \frac{1 - 4ps^2(1 - p)}{4p^2s^2}.$$

Note that  $1 - 4ps^2(1 - p) > 0$  if and only if

$$s^2 < \frac{1}{4p(1 - p)}.$$

Since  $4p(1-p) \leq 1$  for  $p$  in  $(0, 1)$  (why?) we have that  $\frac{1}{4p(1-p)} \geq 1$ . Hence, for  $s$  in  $(0, 1)$  we have

$$s^2 < \frac{1}{4p(1-p)}.$$

Therefore, we have two solutions for  $u = g(s)$

$$u_1 = \frac{1}{2ps} - \sqrt{\frac{1 - 4ps^2(1-p)}{4p^2s^2}} \text{ or } u_2 = \frac{1}{2ps} + \sqrt{\frac{1 - 4ps^2(1-p)}{4p^2s^2}}.$$

We need to decide which expression is  $g(s)$ . Observe that

$$u_2 = \frac{1}{2ps} + \sqrt{\frac{1 - 4ps^2(1-p)}{4p^2s^2}} \sim \frac{1}{4ps}$$

where  $f \sim g$  means that

$$\lim_{s \rightarrow 0} \frac{f(s)}{g(s)} = 1.$$

Since  $\frac{1}{4ps}$  is not bounded near 0 and  $g$  is defined at 0 (in fact  $g(0) = 0$ )  $g(s)$  cannot be  $u_2$ . It must be  $u_1$ . Hence,

$$g(s) = \frac{1}{2ps} - \sqrt{\frac{1 - 4ps^2(1-p)}{4p^2s^2}} \quad (2.3)$$

Since

$$g(s) = \sum_{n=1}^{\infty} P(X = 2n - 1 | Z_0 = 1) s^{2n-1}$$

the expression in (2.3) is useful in computing the distribution of  $X$  only if we can find a power series expansion for that expression. This is our next task.

Using the binomial expansion from Calculus we have the following lemma.

**Lemma 2.1.** For  $|x| < 1$  we have

$$1 - \sqrt{1-x} = \sum_{n=1}^{\infty} c_n x^n$$

where

$$c_n = \frac{(2n-2)!}{2^{2n-1}n!(n-1)!} \text{ for } n \geq 1.$$

Lemma 2.1 will be proved in the exercises. Going back to (2.3) we have

$$g(s) = \frac{1}{2ps} - \sqrt{\frac{1 - 4ps^2(1-p)}{4p^2s^2}} = \frac{1}{2ps}(1 - \sqrt{1 - 4ps^2(1-p)}).$$

Let  $x = 4ps^2(1-p)$  in Lemma 2.1 to get

$$g(s) = \frac{1}{2ps} \sum_{n=1}^{\infty} c_n (4p(1-p)s^2)^n = \sum_{n=1}^{\infty} \frac{(2n-2)!}{n!(n-1)!} (1-p)^n p^{n-1} s^{2n-1}.$$

This power series expansion shows that  $g$  has no singularity at  $s = 0$ . In fact, it is infinitely differentiable at any  $s$  in  $(-1, 1)$ . We now have a power series expansion for  $g$  and this yields the distribution of  $X$ . We get for  $n \geq 1$  that

$$P(X = 2n - 1 | Z_0 = 1) = \frac{(2n-2)!}{n!(n-1)!} (1-p)^n p^{n-1}.$$

### 2.3 The Probability of a Mutation

We are finally ready to compute the probability that a certain mutation eventually appears in a population modeled by a BGW. We will do the computation for the particular BGW we have been considering so far. Recall that in this particular case the offspring distribution is  $p_0 = 1-p$  and  $p_2 = p$ . We also assume that at each birth in the BGW there is a probability  $\mu$  that the new individual has the mutation. Moreover, the new individual has the mutation (or not) independently of everything else. Let  $M$  be the event that the mutation eventually appears in the population. Let  $M^c$  be the complement of  $M$ , that is, the event that the mutation never appears in the population. Let

$$X = \sum_{n \geq 0} Z_n = 1 + \sum_{n \geq 1} Z_n,$$

where the BGW starts with one individual, that is,  $Z_0 = 1$ . Note that  $X$  can be infinite. In fact, the process  $(Z_n)_{n \geq 0}$  survives if and only if  $X = +\infty$  (why?). Intuitively it is clear that if there are infinitely many births and each birth has a fixed probability  $\mu$  of a mutation independently of all other births then the mutation will occur with probability 1. We will examine this issue more carefully in the last subsection of this section.

The probability that the mutation never appears in the population is therefore

$$P(M^c) = P(M^c; X < +\infty).$$

By the rule of averages we have

$$P(M^c) = \sum_{k=1}^{\infty} P(M^c | X = k) P(X = k) = \sum_{k=1}^{\infty} (1 - \mu)^{k-1} P(X = k).$$

Recall the total progeny distribution

$$P(X = 2n - 1 | Z_0 = 1) = \frac{(2n - 2)!}{n!(n - 1)!} (1 - p)^n p^{n-1}.$$

Hence,

$$\begin{aligned} P(M^c) &= \sum_{n=1}^{\infty} (1 - \mu)^{2n-2} P(X = 2n - 1 | Z_0 = 1) \\ &= \sum_{n=1}^{\infty} (1 - \mu)^{2n-2} \frac{(2n - 2)!}{n!(n - 1)!} (1 - p)^n p^{n-1}. \end{aligned}$$

To sum this series we will use Lemma 2.1. We first rearrange the general term of the series to get

$$P(M^c) = \frac{1}{2p(1 - \mu)^2} \sum_{n=1}^{\infty} \frac{(2n - 2)!}{2^{2n-1} n!(n - 1)!} \left( 4(1 - \mu)^2 (1 - p)p \right)^n.$$

Therefore,

$$P(M^c) = \frac{1}{2p(1 - \mu)^2} \sum_{n=1}^{\infty} c_n \left( 4(1 - \mu)^2 (1 - p)p \right)^n$$

where the sequence  $c_n$  is defined in Lemma 2.1. By that lemma we have

$$P(M^c) = \frac{1}{2p(1 - \mu)^2} \left( 1 - \sqrt{1 - 4p(1 - p)(1 - \mu)^2} \right) \quad (2.4)$$

This is the probability no mutation ever appears in the population. Of course, the probability that a mutation does eventually appear is  $P(M) = 1 - P(M^c)$ .

## 2.4 Application: The Probability of Drug Resistance

Drug resistance is a constant threat to the health of individuals who are being treated for a variety of ailments: HIV, tuberculosis, cancer to cite a few. It is also a threat to the population as a whole since there is a risk that a treatable disease may be replaced by a non-treatable one. This is the case, for instance, for tuberculosis.

We are interested in evaluating the risk of a treatment induced drug resistance. In the presence of a drug the drug sensitive strain is weakened (how much it is weakened depends on the efficacy of the drug) and this gives an edge to the drug resistant strain if it appears before the drug is able to eradicate all pathogens. Therefore, what determines the treatment outcome is whether total eradication takes place before the appearance of a drug resistance mutation. We propose a model to compute the probability of pathogen eradication before drug resistance appears.

We now recall the model we have been studying in the previous subsections. We assume that at every unit time a given pathogen may die with probability  $1 - p$  or divide in two with probability  $p$ . Thus, the mean offspring per pathogen is  $2p$ . We assume that  $p$  is strictly between 0 and 1. If  $2p > 1$ , then there is a positive probability for the family tree of a single drug sensitive pathogen to survive forever. If  $2p \leq 1$ , then eradication is certain for drug sensitive pathogens. The parameter  $p$  is a measure of efficacy of the drug. The smaller the  $p$  the more efficient the drug is and the more likely eradication of the drug sensitive pathogen is.

As always for branching processes, we assume that the number of pathogens each pathogen gives birth to is independent of the number of pathogens any other pathogen gives birth to at the same time. We also assume that for each birth of pathogen there is a probability  $\mu$  that the new pathogen is drug resistant. We denote by  $N$  the number of pathogens at the beginning of treatment.

Recall from (2.4) above that the probability of no mutation starting with a single pathogen is

$$P(M^c | Z_0 = 1) = \frac{1}{2p(1-\mu)^2} \left( 1 - \sqrt{1 - 4p(1-p)(1-\mu)^2} \right) \quad (2.5)$$

Usually the treatment will start when the patient has some symptoms. These symptoms start when the number of pathogens is high enough. Therefore we are interested in the model for  $Z_0 = N$  where  $N$  is a rather large number. Define the function  $f$  as

$$f(N, \mu, p) = P(M^c | Z_0 = N).$$

That is,  $f$  is the probability that the drug resistance mutation never appears given that the treatment starts with  $N$  pathogens. In our model each of the  $N$  pathogens starts its own independent BGW. Hence, the probability that there is no mutation in the population is the probability that none of the  $N$  independent BGW generate a mutation. Therefore,

$$f(N, \mu, p) = f(1, \mu, p)^N = \left( \frac{1}{2p(1-\mu)^2} \left( 1 - \sqrt{1 - 4p(1-p)(1-\mu)^2} \right) \right)^N.$$

We are now going to see that the function  $f$  behavior changes drastically depending whether  $p < 1/2$  or  $p > 1/2$ .

- Subcritical case:  $p < 1/2$ . In order to obtain a friendlier expression for  $f$  we compute a linear approximation in  $\mu$  as  $\mu$  approaches 0. Note that the linear approximation for  $(1 - \mu)^{-2}$  is  $1 + 2\mu$ . A little algebra shows

$$\sqrt{1 - 4p(1 - p)(1 - \mu)^2} = |1 - 2p| \sqrt{1 + \frac{8p(1 - p)}{(1 - 2p)^2} \mu - \frac{4p(1 - p)}{(1 - 2p)^2} \mu^2}.$$

By the binomial expansion

$$\sqrt{1 + x} \sim 1 + \frac{1}{2}x$$

where  $f \sim g$  means that

$$\lim_{x \rightarrow 0} \frac{f(x)}{g(x)} = 1.$$

Hence,

$$\sqrt{1 - 4p(1 - p)(1 - \mu)^2} \sim |1 - 2p| \left(1 + \frac{4p(1 - p)}{(1 - 2p)^2} \mu\right)$$

as  $\mu \rightarrow 0$ . Thus, for  $p < 1/2$  we have the linear approximation

$$f(1, \mu, p) \sim 1 - \frac{2p}{1 - 2p} \mu.$$

Since  $f(N, \mu, p) = f(1, \mu, p)^N$  we have for  $p < 1/2$

$$f(N, \mu, p) \sim \left(1 - \frac{2p}{1 - 2p} \mu\right)^N \sim \exp\left(-\frac{2p}{1 - 2p} N \mu\right) \quad (2.6)$$

where we are using that

$$(1 - x)^N \sim \exp(-Nx)$$

as  $x$  approaches 0. Formula (2.6) tells us that the critical parameter for a successful drug treatment is  $N\mu$ . The smaller  $N\mu$  the larger  $f(N, \mu, p)$  and therefore the larger the probability of no drug resistance. The model confirms what has been found by experience. For HIV for instance better start the treatment early (smaller  $N$ ) than late (larger  $N$ ). It also has been found that it is better to use simultaneously three drugs rather than one. The probability  $\mu$  of the appearance of a mutation which is resistant to all three drugs is much smaller than the probability of the appearance of a mutation which is resistant to a single drug. The model also suggests that it is not necessary to have both  $N$  and  $\mu$  small. It is enough to have  $N\mu$  small.

- Supercritical case:  $p > 1/2$ . In this case too we may do an approximation similar to what was done in the subcritical case. But more importantly we have for  $p > 1/2$

$$f(N, \mu, p) \leq \left(\frac{1-p}{p}\right)^N \quad (2.7)$$

We will prove (2.7) below. First note that  $\frac{1-p}{p} < 1$  (for  $p > 1/2$ ) and  $N$  is very large. Therefore,  $f(N, \mu, p)$  will be very small. That is, drug resistance will almost certainly appear. In this case, the model suggests that treatment is futile at best. The drug will make appear something worse (a drug resistant strain) than what it is supposed to cure.

We now prove (2.7). Starting with one pathogen the probability that the BGW will go extinct is  $\frac{1-p}{p}$ . See the exercises. Starting with  $N$  pathogens in order for the drug resistant mutation not to appear it is necessary (but not sufficient) that all  $N$  independent BGW to go extinct. For if one the BGW survives forever the mutation is certain to appear (in a BGW that survives forever there are infinitely many births and each one has the constant probability  $\mu > 0$  of being resistant). Hence, the event “Drug resistance does not occur” is included in “All  $N$  BGW go extinct.” Therefore, the probability of the first event (i.e.  $f(N, \mu, p)$ ) is less than the probability of the second event. Now the probability that  $N$  independent BGW go extinct is  $\left(\frac{1-p}{p}\right)^N$ . This proves (2.7).

## 2.5 Application: Cancer Risk

Cancer has long been thought to appear after several successive mutations. We assume here that a cancerous cell appears after two successive mutations. We consider a tissue in the human body (a tissue is an ensemble of similar cells that together carry out a specific function). The cells of this tissue undergo a fixed number  $D$  of divisions over the lifetime of the tissue. We also assume that there is a probability  $\mu_1$  per division of producing a cell with a type 1 mutation. A cell carrying a type 1 mutation is a pre-cancerous cell. If a type 1 cell appears, it starts a BGW process. At each unit time each cell in this BGW may die with probability  $1 - p_1$  or divide in two type 1 cells with probability  $p_1$ . At each division of a type 1 cell there is a probability  $\mu_2$  for each daughter cell that it be a type 2 cell. A type 2 cell is a cancerous cell. We are interested in computing the probability that a cancerous cell appear over the lifetime of the tissue.

In order for a cancerous cell to appear we first need a type 1 mutation and then a type 2 mutation appearing in the BGW started by the type 1 cell. Assume that at each of the  $D$  divisions we have the same probability  $p$  that the two successive mutations appear. Assuming also independence of these  $D$  events we get

$$P(\text{no cancer}) = (1 - p)^D.$$



We now compute  $p$ . At a given cell division let  $A_1$  be the event of a first mutation and let  $A_2$  be the event that a second mutation eventually occurs. The probability of  $A_1 \cap A_2$  is exactly  $p$ . Therefore,

$$p = P(A_1 \cap A_2) = P(A_1)P(A_2|A_1).$$

We know that  $P(A_1) = \mu_1$  and  $P(A_2|A_1)$  is the probability that a mutation occurs in a BGW starting with a single individual with mutation probability  $\mu_2$  and division probability  $p_1$ . Hence,

$$P(A_2|A_1) = 1 - f(1, \mu_2, p_1)$$

where

$$f(1, \mu_2, p_1) = \frac{1}{2p_1(1 - \mu_2)^2} \left( 1 - \sqrt{1 - 4p_1(1 - p_1)(1 - \mu_2)^2} \right),$$

has been computed in Sect. 2.4. Recall that  $f(1, \mu, p)$  is the probability that no mutation occurs in a BGW with mutation probability  $\mu$  and division probability  $p$ . So  $1 - f(1, \mu_2, p_1)$  is the probability that a mutation does occur. Therefore,

$$p = \mu_1(1 - f(1, \mu_2, p_1)).$$

Let

$$S(p_1, \mu_2) = 1 - f(1, \mu_2, p_1).$$

Hence,

$$P(\text{no cancer}) = (1 - p)^D \sim \exp(-pD) = \exp(-\ell S(p_1, \mu_2))$$

where  $\ell = \mu_1 D$  and the approximation holds for  $p$  approaching 0. The formula above is interesting in several ways. It shows that  $\mu_1$  and  $D$  are important only through their product  $\ell$ . Moreover, the parameter  $\ell$  determines whether  $p_1$  and  $\mu_2$  are important. We now see why.

- Small  $\ell$ . Note that  $S(p_1, \mu_2)$  is a probability and is therefore in  $[0, 1]$  for all  $p_1$  and  $\mu_2$ . Hence,  $\ell S(p_1, \mu_2) \leq \ell$  and

$$P(\text{no cancer}) \geq \exp(-\ell) \sim 1 - \ell$$

where the approximation holds for  $\ell$  approaching 0. That is, the risk of cancer is of order  $\ell$ . The parameters  $p_1$  and  $\mu_2$  are almost irrelevant. This is so because if  $\ell$  is small then the first mutation is unlikely during the lifetime of the tissue. If the first mutation is unlikely, then so is the second mutation since the second mutation can only happen on top of the first.

- Large  $\ell$ . If  $\ell$  is large, then it is quite likely that a first mutation will occur during the lifetime of the tissue. Whether the second mutation occurs depends on  $p_1$  and  $\mu_2$ . It turns out that what determines whether the second mutation occurs is  $p_1$ . The parameter  $\mu_2$  is not really relevant. See the problems.

## 2.6 The Total Progeny May Be Infinite

We now revisit the distribution of the total progeny computed in Sect. 2.2. Recall that the total progeny  $X$  is defined by

$$X = \sum_{n \geq 0} Z_n = 1 + \sum_{n \geq 1} Z_n,$$

where the BGW  $(Z_n)_{n \geq 0}$  starts with one individual, that is,  $Z_0 = 1$ . If the distribution of  $X$  is correct, then  $\sum_{n=1}^{\infty} P(X = 2n - 1 | Z_0 = 1)$  should be 1. Or should it? We now do the computation.

$$\sum_{n=1}^{\infty} P(X = 2n - 1 | Z_0 = 1) = \sum_{n=1}^{\infty} \frac{(2n - 2)!}{n!(n - 1)!} (1 - p)^n p^{n-1}.$$

We will use Lemma 2.1 to compute this infinite series. In order to do so we rearrange the general term of the series. We have

$$\begin{aligned} \frac{(2n - 2)!}{n!(n - 1)!} (1 - p)^n p^{n-1} &= \frac{(2n - 2)!}{2^{2n-1} n!(n - 1)!} 2^{2n-1} (1 - p)^n p^n p^{-1} \\ &= \frac{(2n - 2)!}{2^{2n-1} n!(n - 1)!} (4p(1 - p))^n (2p)^{-1}. \end{aligned}$$

Hence,

$$\begin{aligned} \sum_{n=1}^{\infty} P(X = 2n - 1 | Z_0 = 1) &= \frac{1}{2p} \sum_{n=1}^{\infty} \frac{(2n - 2)!}{2^{2n-1} n!(n - 1)!} (4p(1 - p))^n \\ &= \frac{1}{2p} \sum_{n=1}^{\infty} c_n (4p(1 - p))^n \end{aligned}$$

where the sequence  $(c_n)_{n \geq 1}$  is defined in Lemma 2.1. We now let  $x = 4p(1 - p)$  in Lemma 2.1 to get

$$\sum_{n=1}^{\infty} P(X = 2n - 1 | Z_0 = 1) = \frac{1}{2p} (1 - \sqrt{1 - 4p(1 - p)}).$$

Note that  $1 - 4p(1 - p) = (1 - 2p)^2$ . Hence,

$$\sqrt{1 - 4p(1 - p)} = |1 - 2p|.$$

Therefore,

$$\sum_{n=1}^{\infty} P(X = 2n - 1 | Z_0 = 1) = \frac{1}{2p}(1 - |1 - 2p|).$$

There are two cases to consider. If  $p \leq 1/2$  then  $|1 - 2p| = 1 - 2p$  and

$$\sum_{n=1}^{\infty} P(X = 2n - 1 | Z_0 = 1) = 1$$

as expected. However, if  $p > 1/2$ , then  $|1 - 2p| = -1 + 2p$  and

$$\sum_{n=1}^{\infty} P(X = 2n - 1 | Z_0 = 1) = \frac{1}{2p}(2 - 2p) = \frac{1 - p}{p}$$

which is not 1 (except when  $p = 1/2$ )! What is going on? Recall our definition of  $X$ .

$$X = \sum_{n \geq 0} Z_n$$

where  $(Z_n)_{n \geq 0}$  is a BGW. Now note that the  $Z_n$  are positive or 0 integers. So  $X$  is finite if and only if  $Z_n = 0$  for all  $n$  larger than some fixed integer (why?). This is the same as saying that  $X$  is finite if and only if the BGW  $(Z_n)_{n \geq 0}$  dies out. It is easy to check (see the exercises) that for this particular BGW extinction occurs if and only if  $p \leq 1/2$ . If  $p > 1/2$ , then there is a positive probability that the BGW does not die out. That is, there is a positive probability that  $X$  is infinite. Observe also that  $\sum_{n=1}^{\infty} P(X = 2n - 1 | Z_0 = 1)$  is the probability that  $X$  takes a finite value. This series does not include the possibility that  $X$  is infinite. This is why when  $p > 1/2$  the series is strictly less than 1. We have

$$\sum_{n=1}^{\infty} P(X = 2n - 1 | Z_0 = 1) = P(X < +\infty | Z_0 = 1) = \frac{1 - p}{p}.$$

We will check in the exercises that if  $p > 1/2$  the probability that  $(Z_n)_{n \geq 0}$  dies out is indeed  $\frac{1-p}{p}$ .

We now tie another loose end. What is the probability that a mutation appears when  $X$  is infinite? We will show below that this probability is 1. Let  $M$  be the probability that a mutation occurs at some point in the process  $(Z_n)_{n \geq 0}$ . Let  $k$  be a positive integer. We have

$$P(M^c; X \geq k) = P(M^c | X \geq k)P(X \geq k) \leq (1 - \mu)^{k-1}P(X \geq k) \quad (2.8)$$

This is so because given  $X \geq k$  there are at least  $k - 1$  births in the population and each birth has independently failed to carry the mutation. Observe also that the sequence  $\{X \geq k\}$  for  $k \geq 1$  is decreasing. That is, for  $k \geq 1$

$$\{X \geq k + 1\} \subset \{X \geq k\}.$$

Hence, by Proposition 1.1 in the appendix

$$\lim_{k \rightarrow \infty} P(X \geq k) = P\left(\bigcap_{k \geq 1} \{X \geq k\}\right).$$

Note that if  $X \geq k$  for every  $k \geq 1$  then  $X = +\infty$ . Thus,

$$\lim_{k \rightarrow \infty} P(X \geq k) = P(X = +\infty).$$

With a similar argument we show that

$$\lim_{k \rightarrow \infty} P(M^c; X \geq k) = P(M^c; X = +\infty).$$

Letting  $k$  go to infinity in (2.8) yields

$$P(M^c; X = +\infty) \leq P(X = +\infty) \lim_{k \rightarrow \infty} (1 - \mu)^{k-1} = 0$$

since  $0 < 1 - \mu < 1$ . Hence, a mutation occurs with probability 1.

## Problems

1. Let

$$X = \sum_{n \geq 0} Z_n = 1 + \sum_{n \geq 1} Z_n,$$

where the BGW  $(Z_n)_{n \geq 0}$  starts with one individual, that is,  $Z_0 = 1$ . Show that the process  $(Z_n)_{n \geq 0}$  survives if and only if  $X = +\infty$ .

2. Consider a BGW with the following offspring distribution:  $p_0 = 1 - p$  and  $p_2 = p$  where  $p$  is a fixed parameter in  $[0, 1]$ .

- (a) Show that the BGW may survive if and only if  $p > 1/2$ .
- (b) Show that the moment generating function of the offspring distribution is

$$f(s) = 1 - p + ps^2.$$

- (c) Show that the extinction probability is  $\frac{1-p}{p}$  when  $p > 1/2$ .

**3.** Consider a BGW with the offspring distribution

$$p_k = (1 - r)r^k \text{ for } k \geq 0,$$

where  $r$  is in  $(0, 1)$ .

- (a) Show that the generating function of the offspring distribution is for  $s \leq 1$

$$f(s) = \frac{1 - r}{1 - rs}.$$

- (b) Let  $g$  be the generating function of the total progeny for this BGW starting with a single individual. Use Eq. (2.2) to show that

$$rg(s)^2 - g(s) + s(1 - r) = 0.$$

- (c) The quadratic equation in (b) has two solutions. Explain why  $g(s)$  is in fact

$$g(s) = \frac{1}{2r}(1 - \sqrt{1 - 4s(1 - r)r}).$$

- (d) Use Lemma 2.1 to show that

$$g(s) = \frac{1}{2r} \sum_{n \geq 1} c_n 4^n (1 - r)^n r^n s^n$$

where  $c_n$  is defined in Lemma 2.1.

- (e) Let  $X$  be the total progeny of this BGW starting with a single individual. Show that

$$P(X = n | Z_0 = 1) = \frac{1}{2r} c_n 4^n (1 - r)^n r^n.$$

- (f) Use (e) and Lemma 2.1 to show that

$$\sum_{n \geq 1} P(X = n | Z_0 = 1) = \frac{1}{2r}(1 - |1 - 2r|).$$

- (g) Show that  $\sum_{n \geq 1} P(X = n) = 1$  for  $r \leq 1/2$  and  $\sum_{n \geq 1} P(X = n) < 1$  for  $r > 1/2$ . Could you have known that without computing the distribution of  $X$ ?  
 (h) Set  $r = 1/4$ . Compute  $P(X = n)$  for  $n = 1, 2, \dots, 10$ .  
 (i) Set  $r = 3/4$ . Compute  $P(X = n)$  for  $n = 1, 2, \dots, 10$ .

**4.** In this problem we compute the probability of a given mutation for the BGW studied in Problem 3. Let  $M^c$  be the event that no mutation ever appears in the BGW that started with a single individual. Assume that for each birth in the BGW there is a probability  $\mu$  that the new individual has the mutation.

(a) Show that

$$P(M^c | Z_0 = 1) = \sum_{n \geq 1} P(X = n | Z_0 = 1)(1 - \mu)^{n-1}$$

where  $X$  is the total progeny of the BGW.

(b) Use Problem 2 (e) and Lemma 2.1 in (a) to show that

$$P(M^c | Z_0 = 1) = \frac{1}{2r(1 - \mu)}(1 - \sqrt{1 - 4(1 - r)r(1 - \mu)}).$$

(c) Explain why for every integer  $N \geq 1$

$$P(M^c | Z_0 = N) = P(M^c | Z_0 = 1)^N.$$

(d) Set  $r = 1/10$ . Compute  $P(M^c | Z_0 = N)$  for several values of  $N$  and  $\mu$  for which  $N\mu = 1$ . What do these computations suggest?

(e) Set  $r = 6/10$  and  $N = 10$ . Compute  $P(M^c | Z_0 = N)$  for  $\mu = 10^{-4}$ ,  $\mu = 10^{-5}$ ,  $\mu = 10^{-6}$ . What do these computations suggest?

5. Recall from Calculus the binomial expansion. Let  $\alpha$  be a real number. Then, for all  $x$  in  $(-1, 1)$ ,

$$(1 + x)^\alpha = 1 + \sum_{k=1}^{\infty} a_k x^k$$

where

$$a_k = \frac{\alpha(\alpha - 1) \dots (\alpha - k + 1)}{k!}$$

for  $k \geq 1$ . Use the binomial expansion (in the case  $\alpha = 1/2$ ) to prove Lemma 2.1. Do a proof by induction.

6. For  $p < 1/2$  we have approximated

$$f(N, \mu, p) = f(1, \mu, p)^N = \left( \frac{1}{2p(1 - \mu)^2} \left( 1 - \sqrt{1 - 4p(1 - p)(1 - \mu)^2} \right) \right)^N$$

by using

$$h(N, \mu, p) = \exp\left(-\frac{2p}{1 - 2p} N\mu\right)$$

as  $\mu \rightarrow 0$ .

How good is the approximation? Compute  $f$  and  $h$  for  $p$  in  $[0.1, 0.4]$ ,  $\mu$  in  $[10^{-8}, 10^{-4}]$  and  $N$  in  $[10, 10^6]$ . Find out the maximal error for this approximation.

7. The cells of a tissue undergo a fixed number  $D$  of divisions over the lifetime of the tissue. Assume that there is a probability  $\mu_1$  per division of producing a cell with a type 1 mutation.

- (a) Show that the probability of having at least one type 1 mutation over the lifetime of the tissue is

$$s = 1 - (1 - \mu_1)^D.$$

- (b) Show that

$$s \sim 1 - \exp(-\mu_1 D)$$

as  $\mu_1$  approaches 0.

- (c) Let  $m = \mu_1 D$ . Sketch the graph of  $s$  as a function of  $m$ .

8. In the cancer risk model of Sect. 2.6 we have shown that the risk  $r$  of cancer for a certain tissue is

$$r \sim 1 - \exp(-\ell S(p_1, \mu_2)).$$

Let  $\mu_2 = 10^{-6}$ . Sketch the graphs of  $r$  as a function of  $p_1$  for  $\ell = 0.01$ ,  $\ell = 0.1$ , and  $\ell = 1$ . Interpret these graphs.

### 3 Proof of Theorem 1.1

This proof involves mostly analysis arguments and is not important for the sequel. We include it for the sake of completeness.

Before proving Theorem 1.1 we will need a few properties of generating functions. Recall that the generating function of the probability distribution  $(p_k)_{k \geq 0}$  is

$$f(s) = \sum_{k \geq 0} p_k s^k.$$

We have seen already that a generating function is defined on  $(-1, 1)$ . Since  $f$  is also defined at 1 and we are only interested in positive numbers we will take the domain of  $f$  to be  $[0, 1]$ .

An useful Analysis lemma is the following.

**Lemma 3.1.** *Let  $(b_n)_{n \geq 0}$  be a positive sequence and let*

$$g(t) = \sum_{n \geq 0} b_n t^n.$$

Assume that  $g$  is defined on  $[0, 1)$ . Then

$$\lim_{t \rightarrow 1^-} g(t) = \sum_{n \geq 0} b_n$$

where both sides are possibly  $+\infty$ .

For a proof see Proposition A1.9 in the Appendix of “Theoretical Probability for applications” by S.C. Port.

Applying Lemma 3.1 to the generating function  $f$  we see that

$$\lim_{s \rightarrow 1^-} f(s) = \sum_{n \geq 0} p_n = 1.$$

Since  $f(1) = 1$ ,  $f$  is left continuous at 1. On the other hand, a power series is infinitely differentiable (and hence continuous) on any open interval where it is defined. Therefore  $f$  is continuous on  $[0, 1]$ .

We will need another application of Lemma 3.1. As noted above the function  $f$  is differentiable on  $(0, 1)$  and since a power series can be differentiated term by term

$$f'(s) = \sum_{n \geq 1} n p_n s^{n-1}.$$

By Lemma 3.1

$$\lim_{s \rightarrow 1^-} f'(s) = \sum_{n \geq 1} n p_n = m$$

where  $\lim_{s \rightarrow 1^-} f'(s)$  and  $m$  may be both infinite.

We now go back to the BGW process and compute the generating function of  $Z_n$  for  $n \geq 1$ .

**Proposition 3.1.** *Let  $f_1 = f$  and  $f_{n+1} = f \circ f_n$  for  $n \geq 1$ . For  $n \geq 1$ , the generating function of  $Z_n$  conditioned on  $Z_0 = 1$  is  $f_n$ .*

*Proof of Proposition 3.1.* We prove this by induction. Let  $g_n$  be the generating function of  $Z_n$  given that  $Z_0 = 1$ . We have

$$g_1(s) = E(s^{Z_1} | Z_0 = 1) = E(s^{Y_1}) = f(s) = f_1(s),$$

so the property holds for  $n = 1$ . Assume that  $g_n = f_n$ . Given  $Z_n = k$ , the distribution of  $Z_{n+1}$  is the same as the distribution of  $\sum_{i=1}^k Y_i$  where the  $Y_i$  are i.i.d. with distribution  $(p_k)_{k \geq 0}$ . Hence,

$$E(s^{Z_{n+1}} | Z_n = k) = E(s^{\sum_{i=1}^k Y_i}) = E(s^{Y_1}) E(s^{Y_2}) \dots E(s^{Y_k}) = (E(s^{Y_1}))^k = f(s)^k.$$



By the Markov property

$$\begin{aligned}
 g_{n+1}(s) &= E(s^{Z_{n+1}} | Z_0 = 1) \\
 &= \sum_{k=0}^{\infty} E(s^{Z_{n+1}} | Z_n = k) P(Z_n = k | Z_0 = 1) \\
 &= \sum_{k=0}^{\infty} P(Z_n = k | Z_0 = 1) f(s)^k \\
 &= g_n(f(s))
 \end{aligned}$$

and by our induction hypothesis we get  $g_{n+1} = g_n \circ f = f_n \circ f = f_{n+1}$ . This completes the proof of Proposition 3.1.

We now prove Theorem 1.1. We start by dealing with the easiest case:  $m < 1$ . For any positive integer valued random variable  $X$

$$E(X) = \sum_{k \geq 0} k P(X = k) \geq \sum_{k \geq 1} P(X = k) = P(X \geq 1).$$

Hence,

$$P(X \geq 1) \leq E(X).$$

We use the preceding inequality and Proposition 1.1 to get

$$P(Z_n \geq 1 | Z_0 = 1) \leq E(Z_n | Z_0 = 1) = m^n.$$

Since  $m < 1$

$$\lim_{n \rightarrow \infty} P(Z_n \geq 1 | Z_0 = 1) = 0,$$

and the convergence occurs exponentially fast. Observe that since 0 is a trap for  $(Z_n)_{n \geq 0}$  the sequence of events  $\{Z_n \geq 1\}$  is decreasing. That is,

$$\{Z_{n+1} \geq 1\} \subset \{Z_n \geq 1\}.$$

In words, if  $Z_{n+1} \geq 1$  then we must have  $Z_n \geq 1$  (why?).

By Proposition 1.1 in the Appendix

$$\lim_{n \rightarrow \infty} P(Z_n \geq 1 | Z_0 = 1) = P\left(\bigcap_{n \geq 0} \{Z_n \geq 1\} | Z_0 = 1\right) =$$

$$P(Z_n \geq 1 \text{ for all } n \geq 0 | Z_0 = 1) = 0.$$

This proves Theorem 1.1 in the case  $m < 1$ .

For the cases  $m = 1$  and  $m > 1$  we will need the following observations. For any positive integer valued random variable  $X$  we can define the generating function  $g_X$  on  $[0, 1]$  by

$$g_X(s) = \sum_{k=0}^{\infty} P(X = k)s^k.$$

If we let  $s = 0$ , we get  $g_X(0) = P(X = 0)$ .

Since  $f_n$  is the moment generating function of  $Z_n$  conditioned on  $\{Z_0 = 1\}$  we get

$$P(Z_n = 0 | Z_0 = 1) = f_n(0),$$

and since the sequence of events  $\{Z_n = 0\}$  is increasing (why?) we have by Proposition 1.1 in the Appendix

$$\lim_{n \rightarrow \infty} f_n(0) = P\left(\bigcup_{n \geq 1} \{Z_n = 0\} | Z_0 = 1\right) \quad (3.1)$$

Let  $q$  to be the probability of extinction. Define

$$q = P(Z_n = 0 \text{ for some } n \geq 1 | Z_0 = 1).$$

Observe that

$$\{Z_n = 0 \text{ for some } n \geq 1\} = \bigcup_{n \geq 1} \{Z_n = 0\}.$$

Hence, by (3.1)

$$q = P(Z_n = 0 \text{ for some } n \geq 1 | Z_0 = 1) = P\left(\bigcup_{n \geq 1} \{Z_n = 0\} | Z_0 = 1\right) = \lim_{n \rightarrow \infty} f_n(0).$$

Therefore,

$$q = \lim_{n \rightarrow \infty} f_n(0) \quad (3.2)$$

Now

$$f_{n+1}(0) = f(f_n(0)) \quad (3.3)$$

Since  $f_{n+1}(0)$  and  $f_n(0)$  both converge to  $q$  and  $f$  is continuous on  $[0, 1]$  we get  $f(q) = q$  by letting  $n$  go to infinity in (3.3). That is,  $q$  is a fixed point of  $f$ .

Our task now will be to show that depending on  $m$  we will have  $q = 1$  (extinction is certain) or  $q < 1$  (survival has positive probability).

We first consider  $m = 1$ .

$$f'(s) = \sum_{k \geq 1} k p_k s^{k-1} < \sum_{k \geq 1} k p_k = m = 1 \text{ for } s < 1.$$

Therefore, for any  $s < 1$ , by the Mean Value Theorem there is a  $c$  in  $(s, 1)$  such that

$$f(1) - f(s) = f'(c)(1 - s) < 1 - s,$$

and so for any  $s < 1$

$$f(s) > s.$$

Therefore there is no solution to the equation  $f(s) = s$  in the interval  $[0, 1]$  other than  $s = 1$ . Hence, we must have  $q = 1$ . This completes the proof of Theorem 1.1 for the case  $m = 1$ .

Consider now  $m > 1$ . We have for  $s$  in  $[0, 1)$  that

$$f'(s) = \sum_{k=1}^{\infty} k p_k s^{k-1}.$$

Moreover, by Lemma 3.1 we have that

$$\lim_{s \rightarrow 1} f'(s) = \sum_{k=1}^{\infty} k p_k = m.$$

Hence, there exists an  $\eta > 0$  such that if  $s \geq 1 - \eta$  then  $1 < f'(s) < m$ . By the Mean Value Theorem, for any  $s$  in  $(1 - \eta, 1)$  there is a  $c$  in  $(s, 1)$  such that

$$f(1) - f(s) = (1 - s)f'(c).$$

Since  $f(1) = 1$  and  $f'(c) > 1$  (since  $c > s > 1 - \eta$ ) we have  $1 - f(s) > 1 - s$ . Hence, there is an  $\eta > 0$  such that

$$f(s) < s \text{ for } s \text{ in } (1 - \eta, 1) \quad (3.4)$$

Let  $g(x) = x - f(x)$ . This is a continuous function on  $[0, 1]$ . By (3.4)  $g(s) > 0$  for  $s > 1 - \eta$ . Moreover,  $f(0) \geq 0$  and therefore  $g(0) \leq 0$ . By the Intermediate Value Theorem we have at least one solution in  $[0, 1 - \eta) \subset [0, 1)$  to the equation  $g(s) = 0$  or equivalently to the equation  $f(s) = s$ . Denote this solution by  $s_1$ .

We now show that there is a unique solution to the equation  $f(s) = s$  in  $[0, 1)$ . By contradiction assume there is another solution  $t_1$  in  $[0, 1)$ . Assume without loss

of generality that  $s_1 < t_1$ . Since  $f(1) = 1$  we have at least three solutions to the equation  $g(s) = 0$  on  $[0, 1]$ . We apply Rolle's Theorem on  $[s_1, t_1]$  and on  $[t_1, 1]$  to get the existence of  $\xi_1$  in  $(s_1, t_1)$  and  $\xi_2$  in  $(t_1, 1)$  such that  $g'(\xi_1) = g'(\xi_2) = 0$ . Hence,  $f'(\xi_1) = f'(\xi_2) = 1$ . Observe that

$$f''(s) = \sum_{k \geq 2} k(k-1)p_k s^{k-2}.$$

Since  $p_0 + p_1 < 1$  we must have  $p_k > 0$  for at least one  $k \geq 2$  (why?). Therefore  $f''(s) > 0$  for  $s$  in  $(0, 1)$  and  $f'$  is strictly increasing on  $(0, 1)$ . Therefore, we cannot have  $\xi_1 < \xi_2$  and  $f'(\xi_1) = f'(\xi_2) = 1$ . We have reached a contradiction. Hence, there is a unique solution to the equation  $f(s) = s$  in  $[0, 1)$ .

At this point we know that  $q = s_1$  or  $q = 1$  since these are the two only solutions of  $f(s) = s$  in  $[0, 1]$ . By contradiction assume that  $q = 1$ . By (3.2),  $\lim_{n \rightarrow \infty} f_n(0) = q = 1$ . Hence, for  $n$  large enough,  $f_n(0) > 1 - \eta$ . By (3.4) (let  $s = f_n(0)$  there) this implies that  $f(f_n(0)) < f_n(0)$ . That is,  $f_{n+1}(0) < f_n(0)$ . But this contradicts the fact that  $(f_n(0))_{n \geq 1}$  is an increasing sequence. Hence  $q$  cannot be 1. It must be the unique solution of  $f(s) = s$  which is strictly less than 1. This completes the proof of Theorem 1.1.

## Problems

1. Show that for every  $n \geq 1$  we have

$$\{Z_{n+1} \geq 1\} \subset \{Z_n \geq 1\}.$$

2. Show that for every  $n \geq 1$  we have

$$\{Z_n = 0\} \subset \{Z_{n+1} = 0\}.$$

3. Consider a probability distribution  $(p_k)_{k \geq 0}$ . Show that if  $p_0 + p_1 < 1$  we must have  $p_k > 0$  for at least one  $k \geq 2$ .

4. Let  $f$  be the generating function of the probability distribution  $(p_k)_{k \geq 0}$ .

- (a) Show that if  $p_0 < 1$  then  $f$  is strictly increasing on  $(0, 1)$ .
- (b) Show that if  $p_0 + p_1 < 1$  then  $f'$  is strictly increasing on  $(0, 1)$ .

5. Consider the generating function  $f$  of the offspring distribution  $(p_k)_{k \geq 0}$ . We assume that  $p_0 + p_1 < 1$  and therefore  $f$  is strictly increasing. Assume also that the mean offspring distribution  $m > 1$ .

- (a) We have shown in (3.4) that  $f(s) < s$  for all  $s$  in  $(1 - \eta, 1)$  where  $\eta$  is some positive real number. Show that in fact

$$f(s) < s \text{ for all } s \text{ in } (q, 1).$$

(Do a proof by contradiction. Show that if  $f(s_0) \geq s_0$  for some  $s_0$  in  $(q, 1)$  then the equation  $f(s) = s$  would have at least two solutions in  $[0, 1)$ .)

- (b) Recall that  $f_n$  is the  $n$ th iterate of the generating function  $f$  (see Proposition 3.1). For a fixed  $s$  in  $(q, 1)$  define the sequence  $a_n = f_n(s)$  for  $n \geq 1$ . Show that for every  $n \geq 1$

$$a_n > q.$$

- (c) Show that the sequence  $a_n$  is decreasing.  
 (d) Show that  $a_n$  converges to a limit  $\ell$  which is in  $[q, 1)$ . Show also that  $f(\ell) = \ell$ .  
 (e) Show that  $\ell$  is in fact  $q$ . That is, we have shown that for any  $s$  in  $(q, 1)$ ,  $f_n(s)$  converges to  $q$ .  
 (f) Do steps similar to (a) through (e) to show that  $f_n(s)$  converges to  $q$  for any  $s$  in  $[0, q)$ .
6. By Proposition 3.1  $f_n$  is the generating function of  $Z_n$ . That is, for  $n \geq 1$

$$f_n(s) = \sum_{k \geq 0} P(Z_n = k | Z_0 = 1) s^k.$$

- (a) In Problem 5 we proved that  $f_n(s)$  converges to  $q$  for any fixed  $s$  in  $[0, 1)$ . Show that

$$\lim_{n \rightarrow \infty} P(Z_n = 0 | Z_0 = 1) = q.$$

- (b) Show that for any fixed  $k \geq 1$

$$\lim_{n \rightarrow \infty} P(Z_n = k | Z_0 = 1) = 0.$$

- (c) If  $Z_n$  does not get extinct, where does it go to?

## Notes

For a history of the Bienaymé–Galton–Watson process see Kendall (1975). There are entire books on branching processes. My favorite is Harris (1989). The drug resistance application is based on Schinazi (2006a,b), see also Iwasa et al. (2004) for another stochastic model. See Nowak and May (2000) for differential equation models for drug resistance. The cancer risk application is based on Schinazi (2006a,b). Durrett et al. (2009) and Schweinsberg (2008) have also studied stochastic models for multi-stage carcinogenesis but their models lead to rather involved mathematics. Schweinsberg (2008) gives a nice account of the history of the multi-stage carcinogenesis hypothesis.

## References

- Durrett, R., Schmidt, D., Schweinsberg, J.: A waiting time problem arising from the study of multi-stage carcinogenesis. *Ann. Appl. Probab.* **19**, 676–718 (2009)
- Harris, T.E.: *The Theory of Branching Processes*. Dover, New York (1989)
- Iwasa, Y., Michor, F., Nowak, M.A.: Evolutionary dynamics of invasion and escape. *J. Theor. Biol.* **226**, 205–214 (2004)
- Kendall, D.G.: The genealogy of genealogy branching processes before (and after) 1873. *Bull. Lond. Math. Soc.* **7**(3), 225–253 (1975)
- Lotka, A.J.: *Theorie analytique des associations biologiques*. Actualites scientifiques et industrielles, Hermann, Paris **780** (1939)
- Nowak, M.A., May, R.M.: *Virus Dynamics*. Oxford University Press, New York (2000)
- Port, S.C.: *Theoretical Probability for Applications*. Wiley, New York (1994)
- Schinazi, R.B.: The probability of treatment induced drug resistance. *Acta Biotheor.* **54**, 13–19 (2006a)
- Schinazi, R.B.: A stochastic model for cancer risk. *Genetics* **174**, 545–547 (2006b)
- Schweinsberg, J.: Waiting for  $m$  mutations. *Electron. J. Probab.* **13**, 1442–1478 (2008)

Classical and Spatial Stochastic Processes

With Applications to Biology

Schinazi, R.B.

2014, XII, 268 p. 334 illus., 206 illus. in color.,

Hardcover

ISBN: 978-1-4939-1868-3

A product of Birkhäuser Basel