

**Revision:** a branching process consists of reproducing individuals.

- All individuals are independent.
- Start with a single individual at time 0:  $Z_0 = 1$ .
- Each individual lives a single unit of time, then has  $Y$  offspring and dies.
- Let  $Z_n$  be the size of generation  $n$ : the number of individuals born at time  $n$ .
- The branching process is  $\{Z_0 = 1, Z_1, Z_2, \dots\}$ .

### Branching Process Recursion Formula

This is the fundamental formula for branching processes. Let  $G_n(s) = \mathbb{E}(s^{Z_n})$  be the PGF of  $Z_n$ , the population size at time  $n$ . Let  $G(s) = G_1(s)$ , the PGF of the family size distribution  $Y$ , or equivalently, of  $Z_1$ . Then:

$$G_n(s) = G_{n-1}^{(1)}(G(s)) = \underbrace{G(G(G(\dots G(s) \dots)))}_{n \text{ times}} = G^{(1)}(G_{n-1}(s)).$$

## 9.1 Extinction Probability

One of the most interesting applications of branching processes is calculating the probability of eventual extinction. For example, what is the probability that a colony of cancerous cells becomes extinct before it overgrows the surrounding tissue? What is the probability that an infectious disease dies out before reaching an epidemic? What is the probability that a family line (e.g. for royal families) becomes extinct?

It is possible to find several results about the probability of eventual extinction.

### Extinction by generation $n$

The population is extinct by generation  $n$  if  $Z_n = 0$ .  
(no individuals at time  $n$ .)

If  $Z_n = 0$ , then the population is extinct forever:  
 $Z_t = 0$  for all  $t > n$ .



*Definition:* Define event  $E_n$  to be the event

$E_n = \{Z_n = 0\}$  : event of extinction by generation  $n$ .  
↖ (at or before.)

**Note:**  $E_0 \subseteq E_1 \subseteq E_2 \subseteq E_3 \subseteq \dots$

This is because event  $E_i$  forces  $E_j$  to be true for all  $j \geq i$ , so  $E_i$  is a 'part' or subset of  $E_j$  for  $j \geq i$ .

### Ultimate extinction

At the start of the branching process, we are interested in the probability of *ultimate extinction*: the probability that the population will be extinct ultimately — i.e. that there is some  $n$  at which extinction occurs.

We can express this probability in different ways:

$$\mathbb{P}(\text{ultimate extinction}) = \mathbb{P}\left(\bigcup_{n=0}^{\infty} E_n\right)$$

ie. extinct by 0  
or 1  
or 2  
or ...

**Or:**  $\mathbb{P}(\text{ultimate extinction}) = \mathbb{P}\left(\lim_{n \rightarrow \infty} E_n\right)$

ie. prob extinct by "generation  $\infty$ ".

$$P(E_0) \leq P(E_1) \leq P(E_2) \leq \dots$$

**Note:** By the Continuity Theorem (Chapter 2), and because  $E_0 \subseteq E_1 \subseteq E_2 \subseteq \dots$ , we have:

$$P(\text{ultimate extinction}) = \mathbb{P}\left(\lim_{n \rightarrow \infty} E_n\right) = \lim_{n \rightarrow \infty} P(E_n).$$

*limit of a sequence of sets* (pointing to  $\lim_{n \rightarrow \infty} E_n$ )  
*limit of a sequence of real numbers* (pointing to  $\lim_{n \rightarrow \infty} P(E_n)$ )  
*generates 1 real number* (pointing to  $\lim_{n \rightarrow \infty} E_n$ )  
*real number between 0 & 1* (pointing to  $\lim_{n \rightarrow \infty} P(E_n)$ )

Thus the probability of eventual extinction is the limit as  $n \rightarrow \infty$  of the probability of extinction by generation  $n$ .

We will use the Greek letter Gamma ( $\gamma$ ) for the probability of extinction: think of Gamma for 'all Gone'!

$$\gamma_n = P(E_n) = P(\text{extinction by generation } n) \leftarrow$$

$$\gamma = P(\text{ultimate extinction}) \leftarrow$$

By the Note above, we have established that we are looking for:



$$P(\text{ultimate extinction}) = \gamma = \lim_{n \rightarrow \infty} \gamma_n$$

$\gamma$

Extinction is Forever

**Theorem 9.1:** Let  $\gamma$  be the probability of ultimate extinction. Then

$\gamma$  is the smallest non-negative solution of the equation  $G(s) = s$ , where  $G(s)$  is the PGF of the family size distribution,  $Y$ .

To find the probability of ultimate extinction, we therefore:

- Find the PGF of family size,  $Y$  :  $G(s) = \mathbb{E}(s^Y)$
- Find values of  $s$  that satisfy  $G(s) = s$ .
- Find the smallest of these values that is  $\geq 0$ . This is the required value  $\gamma$ .

$$G(\gamma) = \gamma, \text{ and } \gamma \text{ is the smallest value } \geq 0 \text{ for which this holds.}$$

$\gamma$  is a hitting probability:  $P(\text{ever reach state 0 (extinction)} \mid \text{starting at State 1})$   
So Thm 5.11, p.101 applies. See Bonus Q10 for link between minimal non-neg soln in Thm 5.11 and min non-neg soln of  $G(s) = s$ .

**Note:** Recall that, for any (non-defective) random variable  $Y$  with PGF  $G(s)$ ,  $G(1)=1$ :

$$G(1) = \mathbb{E}(1^Y) = \sum_y 1^y \mathbb{P}(Y=y) = \sum_y \mathbb{P}(Y=y) = 1.$$

So  $G(1) = 1$  always, and therefore there always exists a solution for  $G(s) = s$  in  $[0, 1]$ .

The required value  $\gamma$  is the smallest such solution  $\geq 0$ .

Before proving Theorem 9.1 we prove the following Lemma.

**Lemma:** Let  $\gamma_n = \mathbb{P}(Z_n = 0)$ . Then  $\gamma_n = G(\gamma_{n-1})$ .

**Proof:** If  $G_n(s)$  is the PGF of  $Z_n$ , then  $\mathbb{P}(Z_n = 0) = G_n(0)$  (Chapter 7)  $\uparrow$

So  $\gamma_n = G_n(0)$ . Likewise,  $\gamma_{n-1} = G_{n-1}(0)$ .

$$\text{Now } G_n(0) = \underbrace{G(G(\dots G(0)\dots))}_{n \text{ times}} = G(G_{n-1}(0))$$

$$\text{So } \gamma_n = G(\gamma_{n-1}). \quad \leftarrow \quad \square$$

**Proof of Theorem 9.1:** We need to prove:

(i)  $G(\gamma) = \gamma$ ;

(ii)  $\gamma$  is the smallest non-negative value for which  $G(\gamma) = \gamma$ .

That is, if  $s \geq 0$  and  $G(s) = s$ , then  $\gamma \leq s$ .

Intuitively,  
Lemma  $\Rightarrow \gamma_n = G(\gamma_{n-1})$   
As  $n \rightarrow \infty$ ,  $\gamma_n \rightarrow \gamma$ ,  $\gamma_{n-1} \rightarrow \gamma$   
So  $\gamma = G(\gamma)$ .

**Proof of (i):**

From overleaf,

$$\begin{aligned} \gamma &= \lim_{n \rightarrow \infty} \gamma_n \\ &= \lim_{n \rightarrow \infty} G(\gamma_{n-1}) \quad \text{by Lemma} \\ &= G\left(\lim_{n \rightarrow \infty} \gamma_{n-1}\right) \quad \text{allowed because } G \text{ is continuous} \\ &= G(\gamma). \quad \text{by definition.} \quad \square \end{aligned}$$

Proof of (ii): *don't need to learn.*

First note that  $G(s)$  is an increasing function on  $[0, 1]$ :

$$\begin{aligned} G(s) &= \mathbb{E}(s^Y) = \sum_{y=0}^{\infty} s^y \mathbb{P}(Y = y) \\ \Rightarrow G'(s) &= \sum_{y=0}^{\infty} y s^{y-1} \mathbb{P}(Y = y) \quad \leftarrow \text{only terms } \geq 0 \text{ in here.} \\ \Rightarrow \underline{G'(s) \geq 0} &\quad \text{for } 0 \leq s \leq 1, \quad \text{so } \underline{G \text{ is increasing on } [0, 1].} \end{aligned}$$

$G(s)$  is increasing on  $[0, 1]$  means that:

$$s_1 \leq s_2 \Rightarrow G(s_1) \leq G(s_2) \quad \text{for any } s_1, s_2 \in [0, 1].$$

The branching process begins with  $Z_0 = 1$ , so

$$\mathbb{P}(\text{extinct by generation 0}) = \gamma_0 = 0.$$

At any later generation,  $\gamma_n = G(\gamma_{n-1})$  by Lemma.

Now suppose that  $s \geq 0$  and  $G(s) = s$ . Then we have:

*"imposter solution"*

$$\underline{0 \leq s} \Rightarrow \underline{\gamma_0 \leq s} \quad (\text{because } \gamma_0 = 0)$$

$$\Rightarrow G(\gamma_0) \leq G(s) \quad (\text{by } \clubsuit)$$

$$\text{i.e. } \underline{\gamma_1 \leq s} \quad \text{by Lemma because } s = G(s) \text{ by hypothesis}$$

$$\Rightarrow G(\gamma_1) \leq G(s) \quad (\text{by } \clubsuit)$$

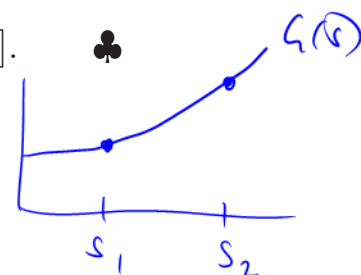
$$\text{i.e. } \underline{\gamma_2 \leq s}$$

$\vdots$

Thus  $\gamma_n \leq s$  for all  $n$ .

So if  $s \geq 0$  and  $G(s) = s$ , then  $\gamma = \lim_{n \rightarrow \infty} \gamma_n \leq s$ .

□



**Example 1:** Let  $\{Z_0 = 1, Z_1, Z_2, \dots\}$  be a branching process with family size distribution  $Y \sim \text{Binomial}(2, \frac{1}{4})$ . Find the probability that the process will eventually die out.

**Solution:** Let  $G(s) = \mathbb{E}(s^Y)$ . We need  $\gamma$ , the smallest soln  $\geq 0$  to  $G(s) = s$ .

For  $Y \sim \text{Bin}(n, p)$ , the PGF is  $G(s) = (ps + q)^n$  (§7.2)  
So if  $Y \sim \text{Bin}(2, \frac{1}{4})$ , then  $G(s) = (\frac{1}{4}s + \frac{3}{4})^2$ .

We need to solve  $G(s) = s$ :

$$G(s) = \left(\frac{1}{4}s + \frac{3}{4}\right)^2 = s$$

$$(s+3)^2 = 16s$$

$$s^2 + 6s + 9 - 16s = 0$$

$$s^2 - 10s + 9 = 0$$

Trick! We know that  $G(1) = 1$ , so  $s=1$  must be a solution. Use this for a quick factorisation.

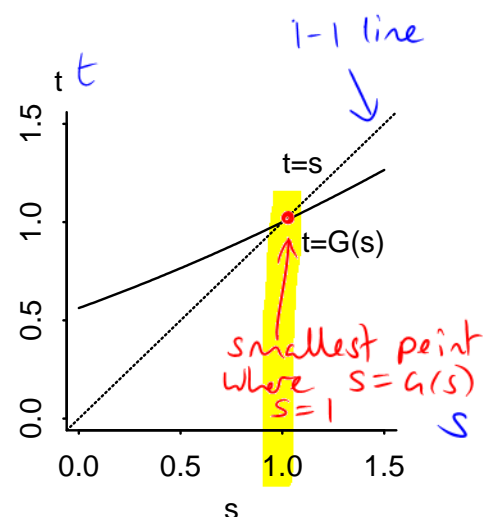
$$(s-1)(s-9) = 0$$

$$\text{So } s=1, \text{ or } s=9.$$

The smallest solution  $\geq 0$  is  $s=1$ .

$$\text{So } \mathbb{P}(\text{ultimate extinction}) = \gamma = 1.$$

Extinction is definite when the family size distribution is  $Y \sim \text{Bin}(2, \frac{1}{4})$ .



$$\mathbb{E} Y = \frac{q}{p} = \frac{3/4}{1/4} = 3$$

$\Rightarrow$  if  $\gamma = 1$  or not.

**Example 2:** Let  $\{Z_0 = 1, Z_1, Z_2, \dots\}$  be a branching process with family size distribution  $Y \sim \text{Geometric}(\frac{1}{4})$ . Find the probability that the process will eventually die out.

**Solution:** Let  $G(s) = \mathbb{E}(s^Y)$ . Then

$\gamma = \mathbb{P}(\text{ultimate extinction}) = (\text{smallest solution } \geq 0 \text{ to } G(s) = s)$ .

For  $Y \sim \text{Geometric}(p)$ , the PGF is  $G(s) = \frac{p}{1 - qs}$  (Chapter 7).

So if  $Y \sim \text{Geometric}(\frac{1}{4})$ , then  $G(s) = \frac{1/4}{1 - (3/4)s} = \frac{1}{4 - 3s}$ .

We need to solve  $G(s) = s$ :

$$G(s) = \frac{1}{4 - 3s} = s$$

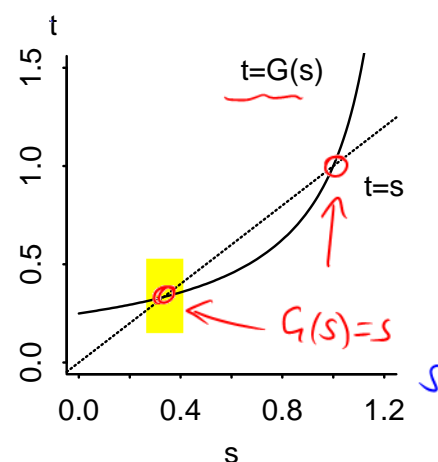
$$\Rightarrow 4s - 3s^2 = 1$$

$$3s^2 - 4s + 1 = 0$$

$$(s - 1)(3s - 1) = 0$$

because we know  $(s - 1)$  must be a factor

$\Rightarrow$  easy factorization.



$$\Rightarrow s = 1 \text{ or } s = \frac{1}{3}$$

\* The smallest solution  $\geq 0$  is  $s = \frac{1}{3}$ .

$$\text{So } \mathbb{P}(\text{ultimate extinction}) = \gamma = \frac{1}{3}.$$

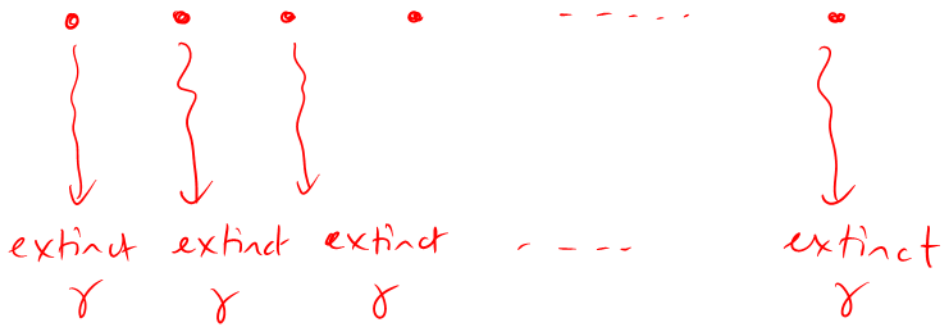
Extinction is possible, but not definite, when the family size is  $Y \sim \text{Geometric}(\frac{1}{4})$ .



If  $\gamma = P(\text{ultimate extinction})$

Suppose  $Z_{10} = 15$  (i.e. there are 15 indivs alive at generation 10)

What is  $P(\text{ultimate extinction} \mid Z_{10} = 15)$  ?



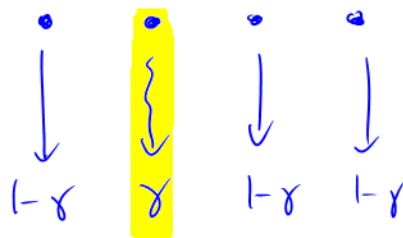
15 chances to stay alive

$$P(\text{ultimate extinction} \mid Z_{10} = 15) = \gamma^{15}$$

General:  $P(\text{extinction} \mid Z_n = k) = \gamma^k$

Suppose  $Z_2 = 4$  (n trials). What's probability that 3 of the 4 indivs alive at generation 2 have surviving descendants for ever? (x successes  $P(\text{success}) = 1 - \gamma$ )

Answer: Binomial probability.



$$\binom{4}{3} \gamma (1 - \gamma)^3$$



## 9.2 Conditions <sup>that guarantee</sup> for ultimate extinction : is $\gamma=1$ , or not?

It turns out that the probability of extinction depends crucially on the value of  $\mu$ , the mean of the family size distribution,  $Y$ .

Some values of  $\mu$  guarantee that the branching process will die out with probability 1. Other values guarantee that the probability of extinction will be strictly less than 1. We will see below that the threshold value is  $\mu=1$ .

$\gamma < 1$  If the mean number of offspring per individual  $\mu$  is more than 1 (so on average, individuals replace themselves plus a bit extra), then the branching process is not guaranteed to die out — although it might do. However, if the mean number of offspring per individual  $\mu$  is 1 or less, the process is guaranteed to become extinct (unless  $Y=1$  with probability 1). The result is not too surprising for  $\mu > 1$  or  $\mu < 1$ , but it is a little surprising that extinction is generally guaranteed if  $\mu = 1$ .

**Theorem 9.2:** Let  $\{Z_0 = 1, Z_1, Z_2, \dots\}$  be a branching process with family size distribution  $Y$ . Let  $\mu = \mathbb{E}(Y)$  be the mean family size distribution, and let  $\gamma$  be the probability of ultimate extinction. Then

- (i) If  $\mu > 1$ , then  $\gamma < 1$  : extinction is NOT certain if  $\mu > 1$ .
- (ii) If  $\mu < 1$ , then  $\gamma = 1$  : extinction IS certain if  $\mu < 1$ .
- (iii) If  $\mu = 1$ , then  $\gamma = 1$ , unless the family size is always constant at  $Y=1$ .  
(pretty boring branching process!!)

**Lemma:** Let  $G(s)$  be the PGF of family size  $Y$ . Then  $G(s)$  and  $G'(s)$  are strictly increasing for  $0 < s < 1$ , as long as  $Y$  can take values  $\geq 2$ .

**Proof:**  $G(s) = \mathbb{E}(s^Y) = \sum_{y=0}^{\infty} s^y \mathbb{P}(Y = y)$ .

So  $G'(s) = \sum_{y=1}^{\infty} y s^{y-1} \mathbb{P}(Y = y) > 0$  for  $0 < s < 1$ ,

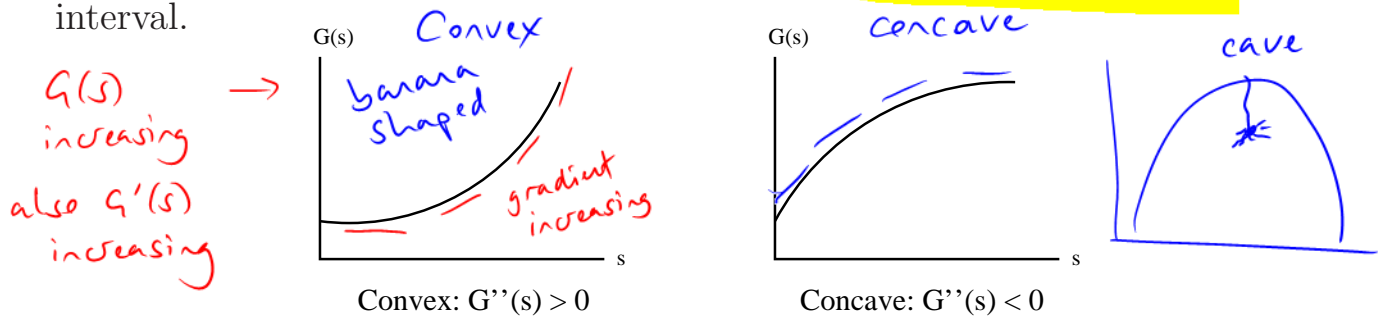
because all terms are  $\geq 0$  and at least 1 term is  $> 0$  (if  $\mathbb{P}(Y \geq 2) > 0$ ).

Similarly,  $G''(s) = \sum_{y=2}^{\infty} y(y-1) s^{y-2} \mathbb{P}(Y = y) > 0$  for  $0 < s < 1$ .

So  $G(s)$  and  $G'(s)$  are strictly increasing for  $0 < s < 1$ .  $\square$

$\gamma = 0, 1$   
otherwise

**Note:** When  $G''(s) > 0$  for  $0 < s < 1$ , the function  $G$  is said to be **convex** on that interval.

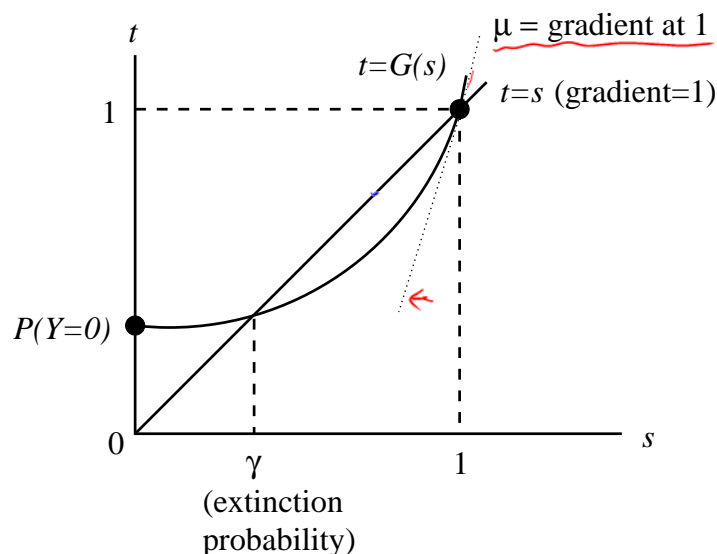


$G''(s) > 0$  means that the gradient of  $G$  is constantly increasing for  $0 < s < 1$ .

**Proof of Theorem 9.2:** This is usually done graphically. *appears often on past exams: not in 2012...*

The graph of  $G(s)$  satisfies the following conditions:

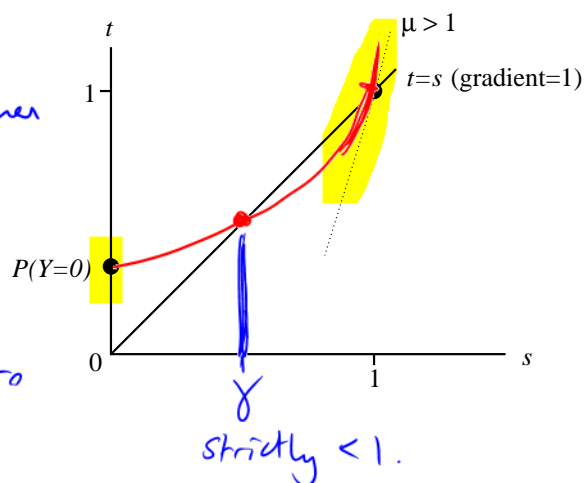
- $G(s)$  is **increasing** and **strictly convex** *(as long as  $Y$  can be  $\geq 2$ )*
- $G(0) = P(Y=0) \geq 0$  *(true for any PGF)*
- $G(1) = 1$  *(any non-defective r.v.)*
- $G'(1) = E(Y) = \mu$  so the slope of  $G(s)$  at  $s=1$  gives the value  $\mu$ .
- The extinction probability,  $\gamma$ , is the smallest value  $\geq 0$  for which  $G(s) = s$  : i.e. the smallest intersection point between  $G(s)$  and the straight line  $t=s$ .



Case (i):  $\mu > 1$  When  $\mu > 1$ , the curve  $G(s)$  is forced beneath the line  $t=s$  at  $s=1$  (gradient of  $G(s)$  is steeper than gradient of the line  $t=s$ ).

The curve  $G(s)$  has to cross the line  $t=s$  again to meet the  $t$ -axis at  $G(0)=P(Y=0)$ .

Thus there must be a solution  $\gamma < 1$  to the equation  $G(s) = s$ .



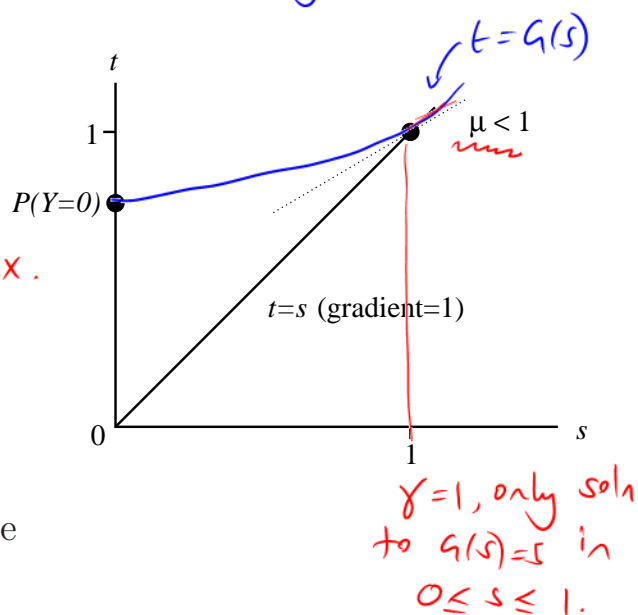
Case (ii):  $\mu < 1$

When  $\mu < 1$ , the curve  $G(s)$  is forced above the line  $t=s$  for  $s < 1$ .

There is no possibility for the curve  $G(s)$  to cross the line  $t=s$  again before meeting the  $t$ -axis. *while being convex.*

Thus there can be no solution  $< 1$  to the equation  $G(s) = s$ , so  $\gamma = 1$ .

The exception is where  $Y$  can take only values 0 and 1, so  $G(s)$  is not strictly convex (see Lemma). However, in that case  $G(s) = p_0 + p_1 s$  is a straight line, giving the same result  $\gamma = 1$ .



Case (iii):  $\mu = 1$

When  $\mu = 1$ , the situation is the same as for  $\mu < 1$ .

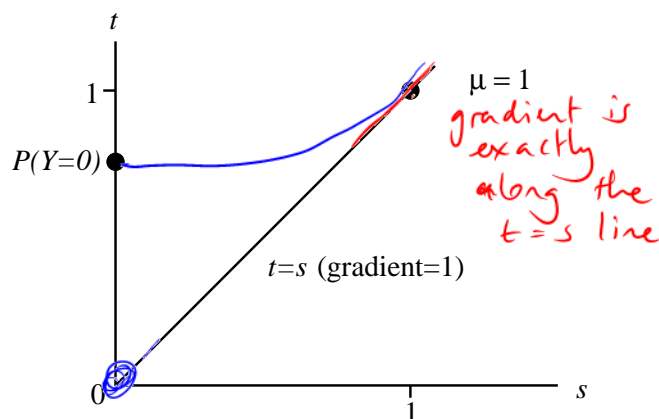
The exception is where  $Y$  takes only the value 1. Then  $G(s) = s$  for all  $0 \leq s \leq 1$ , so the smallest solution  $\geq 0$  is  $\gamma = 0$ .

Thus extinction is guaranteed for  $\mu = 1$ , **unless**  $Y = 1$  with probability 1.

The PGF of  $Y$  is  $\mathbb{E}(s^Y) = G(s)$

so  $G(s) = \mathbb{E}(s^Y) = s$  if  $Y=1$  with certainty.

(but the process  $\{Z_t : t \geq 0\}$  is just constant:  $Z_t = 1$  for all  $t$ )



**Example 1:** Let  $\{Z_0 = 1, Z_1, Z_2, \dots\}$  be a branching process with family size distribution  $Y \sim \text{Binomial}(2, \frac{1}{4})$ , as in Section 9.1. Find the probability of eventual extinction.

**Solution:**  $\mathbb{E}(Y) = \mu = 2 * \frac{1}{4}$  for  $Y \sim \text{Bin}(2, \frac{1}{4})$

$$\therefore \mu = \frac{1}{2}$$

$$\mu < 1.$$

So  $\gamma = \mathbb{P}(\text{extinction}) = 1.$

(The longer calculation in § 9.1 was not necessary).

**Example 2:** Let  $\{Z_0 = 1, Z_1, Z_2, \dots\}$  be a branching process with family size distribution  $Y \sim \text{Geometric}(\frac{1}{4})$ , as in Section 9.1. Find the probability of eventual extinction.

**Solution:**  $\mathbb{E}(Y) = \mu = \frac{q}{p}$  when  $Y \sim \text{Geometric}(p)$

$$\therefore \mu = \frac{3/4}{1/4} = 3 > 1.$$

So we know  $\gamma < 1$  but we still have to do the solution in § 9.1 to find out what  $\gamma$  is (ie. solve  $\gamma(s) = s$ ).

Answer:  $\gamma = \frac{1}{3}.$

**Note:** The mean  $\mu$  of the offspring distribution  $Y$  is known as the

**criticality parameter.**

- If  $\mu < 1$ , extinction is definite ( $\gamma = 1$ ). The process is called **subcritical**.  
Note that  $\mathbb{E}(Z_n) = \mu^n \rightarrow 0$  as  $n \rightarrow \infty$ .
- If  $\mu = 1$ , extinction is definite unless  $Y \equiv 1$ . The process is called **critical**.  
Note that  $\mathbb{E}(Z_n) = \mu^n = 1 \forall n$ , even though extinction is definite.
- If  $\mu > 1$ , extinction is not definite ( $\gamma < 1$ ). The process is called **supercritical**.  
Note that  $\mathbb{E}(Z_n) = \mu^n \rightarrow \infty$  as  $n \rightarrow \infty$ .





But how long have you got...?

### 9.3 Time to Extinction

Suppose the population is doomed to extinction — or maybe it isn't. Either way, it is useful to know how long it will take for the population to become extinct. This is the distribution of  $T$ , the number of generations before extinction. For example, how long do we expect a disease epidemic like SARS to continue? How long have we got to organize ourselves to save the kakapo or the tuatara before they become extinct before our very eyes?



#### 1. Extinction by time $n$

The branching process is extinct by time  $n$  if  $Z_n = 0$ .

Thus the probability that the process has become extinct by time  $n$  is:

$$\mathbb{P}(Z_n = 0) = q_n(0) = \gamma_n$$

**Note:** Recall that  $G_n(s) = \mathbb{E}(s^{Z_n}) = \underbrace{G(G(G(\dots G(s)\dots)))}_{n \text{ times}}$ .

There is no guarantee that the PGF  $G_n(s)$  or the value  $G_n(0)$  can be calculated easily. However, we can build up  $G_n(0)$  in steps:

$$\text{e.g. } q_2(0) = q(q(0))$$

$$\text{then } q_3(0) = q(q_2(0))$$

$$\text{or even } q_4(0) = q_2(q_2(0)).$$

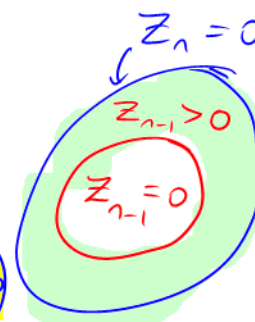
## 2. Extinction at time $n$

Let  $T$  be the exact time of extinction. That is,  $T = n$  if generation  $n$  is the first generation with no individuals:

$$T = n \iff Z_n = 0 \text{ and } Z_{n-1} > 0.$$

Now by the Partition Rule,

$$\mathbb{P}(Z_n = 0 \cap Z_{n-1} > 0) = \mathbb{P}(Z_n = 0) - \mathbb{P}(Z_n = 0 \cap Z_{n-1} = 0)$$



But the event  $\{Z_n = 0 \cap Z_{n-1} = 0\}$  is the event that the process is extinct by generation  $n - 1$  AND it is extinct by generation  $n$ . However, we know it will always be extinct by generation  $n$  if it is extinct by generation  $n - 1$ , so the  $Z_n = 0$  part is redundant. So

$$\begin{aligned} \mathbb{P}(Z_n = 0 \cap Z_{n-1} > 0) &= \mathbb{P}(Z_n = 0) - \mathbb{P}(Z_{n-1} = 0) \\ &= G_n(0) - G_{n-1}(0) \\ &= \gamma_n - \gamma_{n-1} \end{aligned}$$

Similarly,

So (\*) gives.

$$\mathbb{P}(T = n) = G_n(0) - G_{n-1}(0) = \gamma_n - \gamma_{n-1}$$

This gives the distribution of  $T$ , the exact time at which extinction occurs.

**Example: Binary splitting.** Suppose that the family size distribution is

$$Y = \begin{cases} 0 & \text{with probability } q = 1 - p, \\ 1 & \text{with probability } p. \end{cases}$$

Find the distribution of the time to extinction.

***Solution:***

Consider

$$G(s) = \mathbb{E}(s^Y) = qs^0 + ps^1 = q + ps.$$

$$G_2(s) = G(G(s)) = q + p(q + ps) = q(1 + p) + p^2s.$$

$$G_3(s) = G(G_2(s)) = q + p(q + pq + p^2s) = q(1 + p + p^2) + p^3s.$$

$\vdots$

$$G_n(s) = q(1 + p + p^2 + \dots + p^{n-1}) + p^n s.$$

Thus time to extinction,  $T$ , satisfies

$$\begin{aligned} \mathbb{P}(T = n) &= G_n(0) - G_{n-1}(0) \\ &= q(1 + p + p^2 + \dots + p^{n-1}) - q(1 + p + p^2 + \dots + p^{n-2}) \\ &= qp^{n-1} \quad \text{for } n = 1, 2, \dots \end{aligned}$$

Thus

$$T - 1 \sim \text{Geometric}(q).$$

It follows that  $\mathbb{E}(T - 1) = \frac{p}{q}$ , so

$$\mathbb{E}(T) = 1 + \frac{p}{q} = \frac{1 - p + p}{q} = \frac{1}{q}.$$

***Note:*** The expected time to extinction,  $\mathbb{E}(T)$ , is:

- finite if  $\mu < 1$ ;
- infinite if  $\mu = 1$  (despite extinction being definite), if  $\sigma^2$  is finite;
- infinite if  $\mu > 1$  (because with positive probability, extinction never happens).

(Results not proved here.)

Have a quick look – but Assignment is better preparation.

## 9.4 Case Study: Geometric Branching Processes

Recall that  $G_n(s) = \mathbb{E}(s^{Z_n}) = \underbrace{G\left(G\left(G\left(\dots G(s)\dots\right)\right)\right)}_{n \text{ times}}.$

In general, it is not possible to find a closed-form expression for  $G_n(s)$ . We achieved a closed-form  $G_n(s)$  in the Binary Splitting example (page 190), but binary splitting only allows family size  $Y$  to be 0 or 1, which is a very restrictive model.

The only non-trivial family size distribution that allows us to find a closed-form expression for  $G_n(s)$  is the **Geometric distribution**.

When family size  $Y \sim \text{Geometric}(p)$ , we can do the following:

- Derive a closed-form expression for  $G_n(s)$ , the PGF of  $Z_n$ .
- Find the probability distribution of the **exact time of extinction,  $T$** : not just the probability that extinction will occur at some unspecified time ( $\gamma$ ).
- Find the **full probability distribution of  $Z_n$** : probabilities  $\mathbb{P}(Z_n = 0)$ ,  $\mathbb{P}(Z_n = 1)$ ,  $\mathbb{P}(Z_n = 2)$ ,  $\dots$ .

With  $Y \sim \text{Geometric}(p)$ , we can therefore calculate just about every quantity we might be interested in for the branching process.

### 1. Closed form expression for $G_n(s)$

**Theorem 9.4:** Let  $\{Z_0 = 1, Z_1, Z_2, \dots\}$  be a branching process with family size distribution  $Y \sim \text{Geometric}(p)$ . The PGF of  $Z_n$  is given by:

$$G_n(s) = \mathbb{E}(s^{Z_n}) = \begin{cases} \frac{n - (n-1)s}{n+1 - ns} & \text{if } p = q = 0.5, \\ \frac{(\mu^n - 1) - \mu(\mu^{n-1} - 1)s}{(\mu^{n+1} - 1) - \mu(\mu^n - 1)s} & \text{if } p \neq q, \text{ where } \mu = \frac{q}{p}. \end{cases}$$



### Proof (sketch):

The proof for both  $p = q$  and  $p \neq q$  proceed by mathematical induction. We will give a sketch of the proof when  $p = q = 0.5$ . The proof for  $p \neq q$  works in the same way but is trickier.

Consider  $p = q = \frac{1}{2}$ . Then

$$G(s) = \frac{p}{1 - qs} = \frac{\frac{1}{2}}{1 - \frac{s}{2}} = \frac{1}{2 - s}.$$

Using the Branching Process Recursion Formula (Chapter 8),

$$G_2(s) = G(G(s)) = \frac{1}{2 - G(s)} = \frac{1}{2 - \frac{1}{2-s}} = \frac{2 - s}{2(2 - s) - 1} = \frac{2 - s}{3 - 2s}.$$

The inductive hypothesis is that  $G_n(s) = \frac{n - (n - 1)s}{n + 1 - ns}$ , and it holds for  $n = 1$  and  $n = 2$ . Suppose it holds for  $n$ . Then

$$\begin{aligned} G_{n+1}(s) &= G_n(G(s)) = \frac{n - (n - 1)G(s)}{n + 1 - nG(s)} = \frac{n - (n - 1)\left(\frac{1}{2-s}\right)}{n + 1 - n\left(\frac{1}{2-s}\right)} \\ &= \frac{(2 - s)n - (n - 1)}{(2 - s)(n + 1) - n} \\ &= \frac{n + 1 - ns}{n + 2 - (n + 1)s}. \end{aligned}$$

Therefore, if the hypothesis holds for  $n$ , it also holds for  $n + 1$ . Thus the hypothesis is proved for all  $n$ .  $\square$

---

## 2. Exact time of extinction, $T$

Let  $Y \sim \text{Geometric}(p)$ , and let  $T$  be the exact generation of extinction.

From Section 9.3,

$$\mathbb{P}(T = n) = \mathbb{P}(Z_n = 0) - \mathbb{P}(Z_{n-1} = 0) = G_n(0) - G_{n-1}(0).$$

By using the closed-form expressions overleaf for  $G_n(0)$  and  $G_{n-1}(0)$ , we can find  $\mathbb{P}(T = n)$  for any  $n$ .

### 3. Whole distribution of $Z_n$

From Chapter 7,  $\mathbb{P}(Z_n = r) = \frac{1}{r!} G_n^{(r)}(0)$ .

Now our closed-form expression for  $G_n(s)$  has the same format regardless of whether  $\mu = 1$  ( $p = 0.5$ ), or  $\mu \neq 1$  ( $p \neq 0.5$ ):

$$G_n(s) = \frac{A - Bs}{C - Ds}.$$

(For example, when  $\mu = 1$ , we have  $A = D = n$ ,  $B = n - 1$ ,  $C = n + 1$ .) Thus:

$$\mathbb{P}(Z_n = 0) = G_n(0) = \frac{A}{C}$$

$$G'_n(s) = \frac{(C - Ds)(-B) + (A - Bs)D}{(C - Ds)^2} = \frac{AD - BC}{(C - Ds)^2}$$

$$\Rightarrow \mathbb{P}(Z_n = 1) = \frac{1}{1!} G'_n(0) = \frac{AD - BC}{C^2}$$

$$G''_n(s) = \frac{(-2)(-D)(AD - BC)}{(C - Ds)^3} = \frac{2D(AD - BC)}{(C - Ds)^3}$$

$$\Rightarrow \mathbb{P}(Z_n = 2) = \frac{1}{2!} G''_n(0) = \left( \frac{AD - BC}{CD} \right) \left( \frac{D}{C} \right)^2$$

$\vdots$

$$\Rightarrow \mathbb{P}(Z_n = r) = \frac{1}{r!} G_n^{(r)}(0) = \left( \frac{AD - BC}{CD} \right) \left( \frac{D}{C} \right)^r \quad \text{for } r = 1, 2, \dots$$

(Exercise)

This is very simple and powerful: we can substitute the values of  $A, B, C$ , and  $D$  to find  $\mathbb{P}(Z_n = r)$  or  $\mathbb{P}(Z_n \leq r)$  for any  $r$  and  $n$ .

**Note:** A Java applet that simulates branching processes can be found at:

[http://www.dartmouth.edu/~chance/teaching\\_aids/books\\_articles/probability\\_book/bookapplets/chapter10/Branch/Branch.html](http://www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/bookapplets/chapter10/Branch/Branch.html)