Chapter 7:
Extinction in Branching Processes

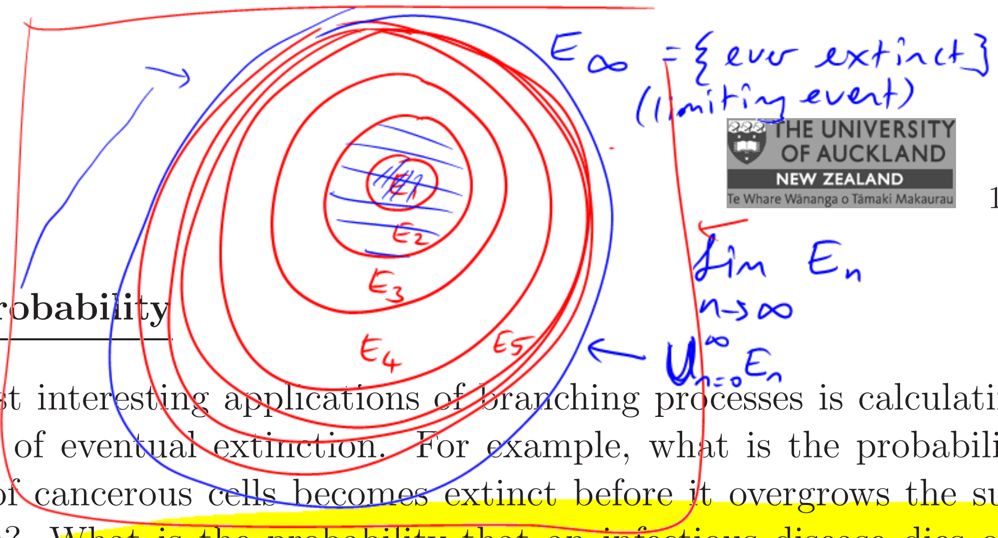
Revision: a branching process consists of reproducing individuals.

- 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 $\{Z_0 = 1, Z_1, Z_2, \ldots\}$.

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:

$$\boxed{G_n(s) = G_{n-1}\Big(G(s)\Big) = \underbrace{G\Big(G\Big(G\Big(\ldots G(s)}_{n \text{ times}}\ldots\Big)\Big)\Big) = G\Big(G_{n-1}(s)\Big).}$$



7.1 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 indivs at time n).


Extinction is Forever

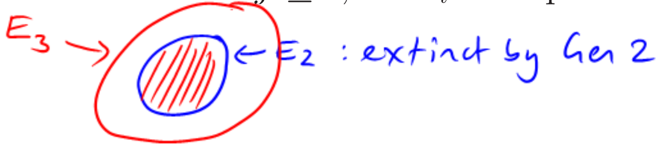
If $Z_n = 0$, then the population is extinct for ever:
ie. $Z_t = 0$ for all $t \geq n$.

Definition: Define event E_n to be the event

$E_n = \{ Z_n = 0 \}$: event that the popn is extinct by generation n.

Note: $\overset{\phi}{E_0} \subseteq E_1 \subseteq E_2 \subseteq E_3 \subseteq \ldots\ldots$

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

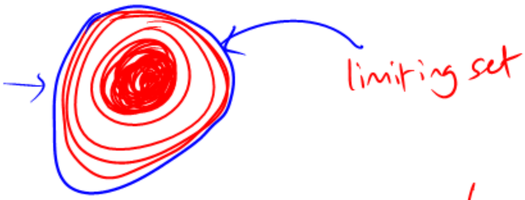
At the start of the branching process, we are interested in the probability of *ultimate extinction:* the probability that the population ever goes extinct, ie. extinct by time n for some unspecified value of n.

We can express this probability in different ways:

$$\mathbb{P}(\text{ultimate extinction}) = \mathbb{P}\left(\bigcup_{n=0}^{\infty} E_n \right) \quad \text{ie. } \mathbb{P}\left(\begin{array}{llll} \text{extinct} & \text{by} & \text{time } 0 & \text{OR} \\ \text{"} & \text{"} & \text{"} \ 1 & \text{OR} \\ \text{"} & \text{"} & \text{"} \ 2 & \text{OR} \\ & & \ddots & \end{array} \right)$$

Or: $\mathbb{P}(\text{ultimate extinction}) = \mathbb{P}\left(\lim_{n\to\infty} E_n \right)$ ie. $\mathbb{P}$(extinct by generation ∞)

sequence of real numbers



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

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

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

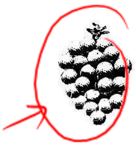
limit of real numbers, $\mathbb{P}(E_0), \mathbb{P}(E_1), \mathbb{P}(E_2), \ldots$

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

$\gamma_n = \mathbb{P}(E_n) = \mathbb{P}(\text{extinct by generation } n)$

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

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



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

γ

Extinction is Forever

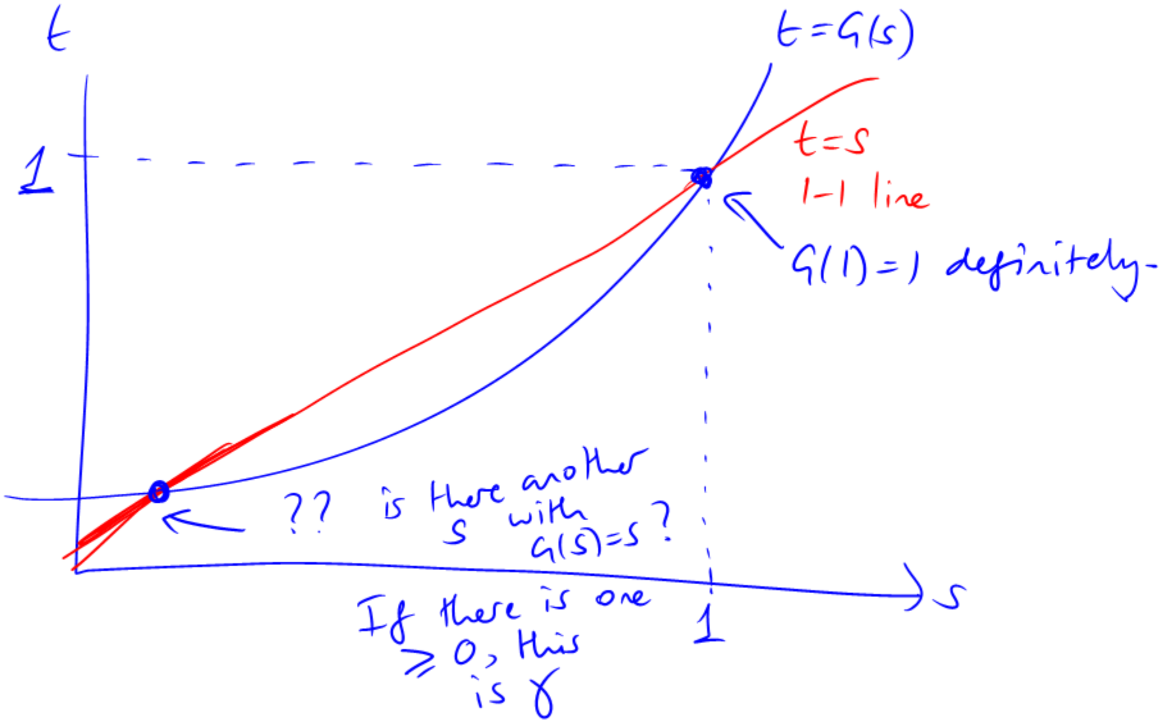
Theorem 7.1: Let γ be the probability of ultimate extinction. Then

γ 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 the values of s that satisfy $G(s) = s$.
- find the smallest of these values that is ≥ 0 ; this is the required value γ.

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


t
t = G(s)
1
t = s
1-1 line
G(1) = 1 definitely.
?? is there another s with G(s) = s ?
s
1
If there is one ≥ 0, this is γ

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

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

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

The required value γ is the smallest such solution ≥ 0.

→ Before proving Theorem 7.1 we prove the following Lemma.

general property of PGFs: Ch 4

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 (last lecture), then $\mathbb{P}(Z_n=0)=G_n(0)$

So $\gamma_n = G_n(0)$.

Similarly, $\gamma_{n-1} = G_{n-1}(0)$.

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

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

Proof of Theorem 7.1: We need to prove:

→ (i) $G(\gamma) = \gamma$; ✓

→ (ii) γ is the smallest non-negative value for which $G(\gamma) = \gamma$.
That is, if $s \geq 0$ and $G(s) = s$, then $\gamma \leq s$.

Proof of (i):

From  overleaf, $\gamma = \lim_{n\to\infty} \gamma_n$

$= \lim_{n\to\infty} G(\gamma_{n-1})$ by Lemma

$= G(\underbrace{\lim_{n\to\infty} \gamma_{n-1}}_{=\gamma})$

allowed because G is continuous; ie. if γ_{n-1} and γ_n are very close, then $G(\gamma_{n-1})$ and $G(\gamma_n)$ are very close.

$\Rightarrow \gamma = G(\gamma)$. $\square$

Proof of (ii):

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

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

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

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

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:

Setting up an "imposter" solution, s.

$$\begin{aligned} 0 \leq s \quad &\Rightarrow \quad 0 = \gamma_0 \leq s && \text{(because } \gamma_0 = 0\text{)} \\ &\Rightarrow \quad G(\gamma_0) \leq G(s) && \text{(by } \clubsuit\text{)} \\ \text{i.e.} \quad & \qquad \gamma_1 \leq s \\ &\Rightarrow \quad G(\gamma_1) \leq G(s) && \text{(by } \clubsuit\text{)} \\ \text{i.e.} \quad & \qquad \gamma_2 \leq s \\ & \qquad \vdots \end{aligned}$$

$$\text{Thus} \qquad \gamma_n \leq s \quad \text{for all } n.$$

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

All of Ass 4 except Q2(d, e) (read ahead)

Example 1: Let $\{Z_0 = 1, Z_1, Z_2, \ldots\}$ 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.

$\mathbb{E}Y = \frac{1}{2}$ $\qquad Y = 0 \quad 1 \quad 2$

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

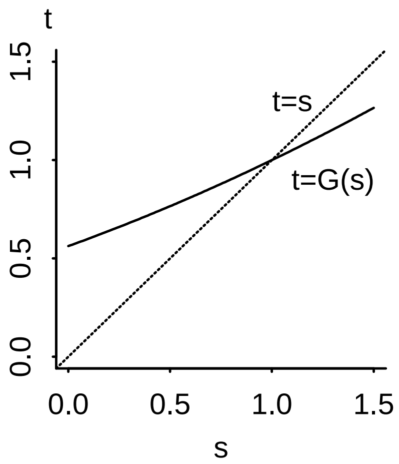
Require $\mathbb{P}(\text{ultimate extinction}) = \gamma$.

Know γ is the smallest solution ≥ 0 to the equation $G(s) = s$.

Now $Y \sim Bin(2, \frac{1}{4})$, so the PGF is

$G(s) = (ps + q)^n$ for $Bin(n, p)$

(Ch 4, § 4.2)



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

Solve $G(s) = s$

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

$$\frac{1}{16}s^2 + \frac{6}{16}s + \frac{9}{16} = s$$

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

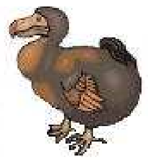
$$\Rightarrow s^2 - 10s + 9 = 0$$

Trick: we know $s=1$ is always a solution for $G(s) = s$. So use this to factorize quickly:

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

$\Rightarrow$ Solutions $s = 1$, $s = 9$.

$\swarrow$ smallest soln ≥ 0.



The smallest soln ≥ 0 is $s = 1$

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

Extinction is definite when $Y \sim Bin(2, \frac{1}{4})$.

Example 2: Let $\{Z_0 = 1, Z_1, Z_2, \ldots\}$ 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: For $Y \sim \text{Geo}(\frac{1}{4})$, the PGF is $G(s) = \frac{p}{1-qs}$ (Ch 4)

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

Need to solve $G(s) = s$:

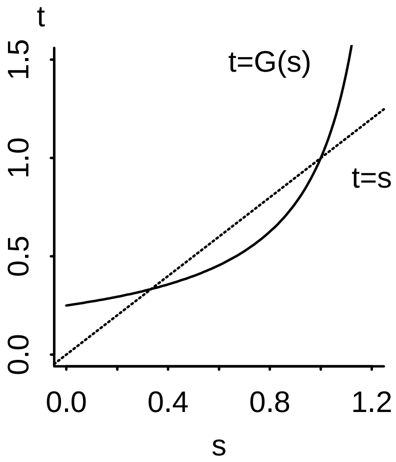
$$\frac{1}{4-3s} = s$$

$$\Rightarrow 3s^2 - 4s \underset{\sim}{+1} = 0$$

Trick: $(s-1)$ must be a factor.

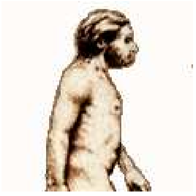
$$\Rightarrow (s-1)(3s-1) = 0$$

$$\Rightarrow s = 1, \text{ or } s = \frac{1}{3} \swarrow \text{this is } \gamma$$



$$\therefore \mathbb{P}(\text{ultimate extinction}) = \gamma = \frac{1}{3} \quad \text{when } Y \sim \text{Geo}(\tfrac{1}{4})$$

Extinction is possible, but not definite, when $Y \sim \text{Geo}(\frac{1}{4})$



7.2 Conditions for ultimate extinction

It turns out that the probability of extinction depends crucially on the value of μ, where $\mu = \mathbb{E}Y$ is the mean of the family-size distribution. Some values of μ *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 $\mu = 1$.

$\gamma = 1$

$\gamma < 1$

If the mean number of offspring per individual μ 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 μ 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 $\mu = 1$.

$\mu = \mathbb{E}Y = 1$

Theorem 7.2: Let $\{Z_0 = 1, Z_1, Z_2, \ldots\}$ be a branching process with family size distribution Y. Let $\mu = \mathbb{E}(Y)$ be the mean family size distribution, and let γ be the probability of ultimate extinction. Then

(i) If $\mu > 1$, then $\gamma < 1$: extinction is NOT GUARANTEED if $\mu > 1$.

(ii) If $\mu < 1$, then $\gamma = 1$: " IS GUARANTEED if $\mu < 1$.

(iii) If $\mu = 1$, then $\gamma = 1$ UNLESS the family size is always constant at $Y = 1$.

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 ≥ 2.

banana shaped curve.

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 $0 < s < 1$, So $G'(s) > 0$ for $0 < s < 1$.

because all terms are ≥ 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 $0 < s < 1$. □

$\Rightarrow G'(s) > 0$ and $G''(s) > 0$ for $0 < s < 1$.

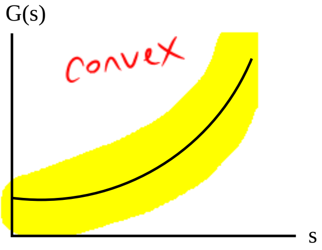
$\Rightarrow G$ and G' are both strictly increasing for $0 < s < 1$.

positive 2nd derivate (as we have here)

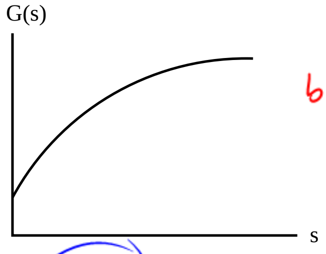
Note: When $G''(s) > 0$ for $0 < s < 1$, the function G is said to be **convex** on that interval.

banana-shaped

$G'(s) > 0$ and $G''(s) > 0$



Convex: G''(s) > 0



Concave: G''(s) < 0

$G'(s) > 0$ but $G''(s) < 0$

Concave

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

Need to understand / reproduce these diagrams.

Proof of Theorem 7.2: This is usually done graphically.

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

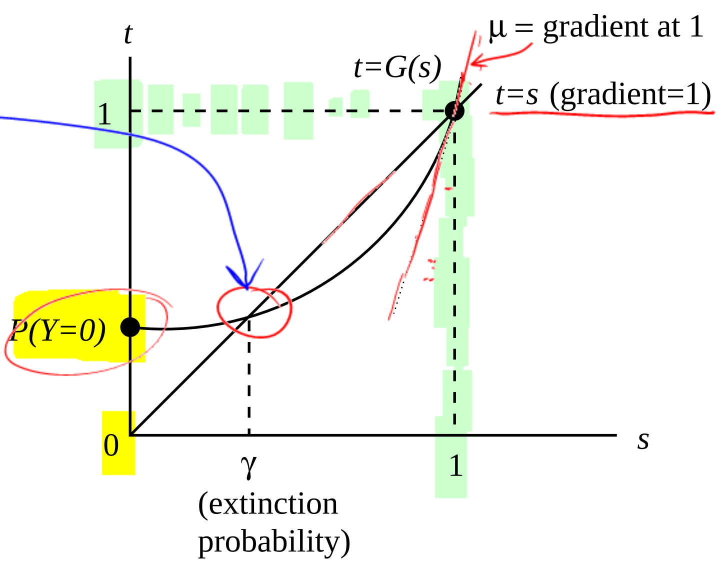
banana → 1. G(s) is increasing and strictly convex (as long as Y can be ≥ 2)

2. $G(0) = \mathbb{P}(Y=0) \geq 0$ (true for all PGFs)

3. $G(1) = 1$ ← (true for any non-defective r.v.)

4. $G'(1) = \mu$, so the slope of G(s) at $s=1$ gives the value μ.

5. The extinction probability, γ, is the smallest value ≥ 0 for which $G(s) = s$ (ie. the smallest non-negative fixed point of G.)

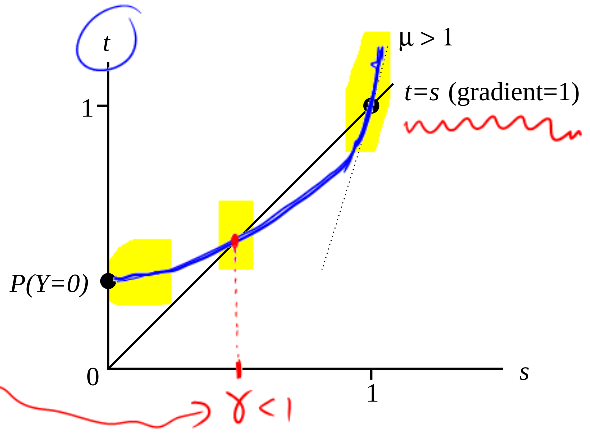


Case (i): $\mu > 1$

When $\mu > 1$, the curve $G(s)$ is forced beneath the line $t = s$ at $s = 1$.

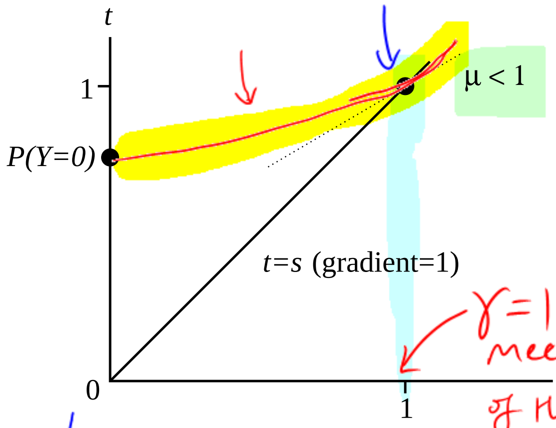
The curve $G(s)$ has to cross the line $t = s$ again to meet the t-axis at $P(Y=0)$.

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



$\gamma = 1$ only meeting of the curves in $0 \leq s \leq 1$.

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 $\gamma = 1$.

$$G(s) = \mathbb{E}(s^Y) = \mathbb{P}(Y=0)s^0 + \mathbb{P}(Y=1)s^1 = p_0 + p_1 s$$

can't do this with either banana or straight line!

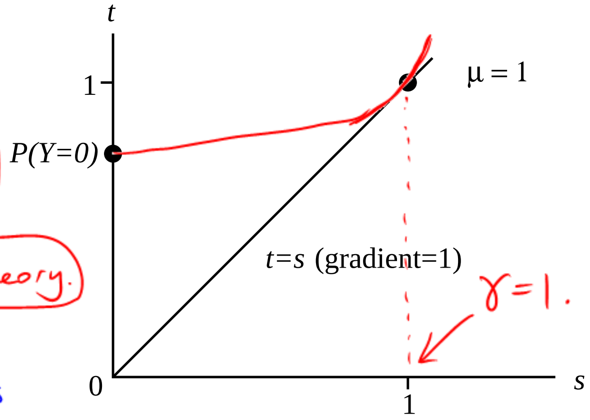
Case (iii): $\mu = 1$

When $\mu = 1$, the situation is the same as for $\mu < 1$. So $\gamma = 1$ in general.

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 ≥ 0 is $\gamma = 0$.

Only exception to general theory.

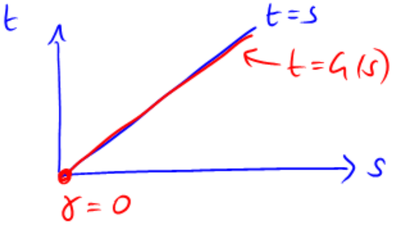
Thus extinction is guaranteed for $\mu = 1$, ***unless*** $Y = 1$ with probability 1.



$\mathbb{P}(Y=1) = 1$ ←

So $\mathbb{E}(s^Y) = s^1$

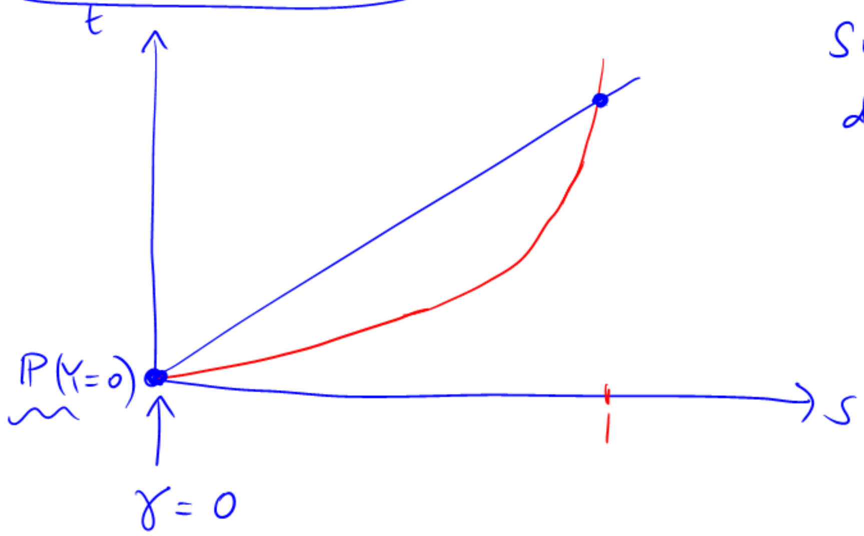
so $G(s) = s$ for all s



$\boxed{\mathbb{P}(Y=0) = 0}$ ← everyone has to have children!

No prospect of extinction.

Sure enough, $\gamma = 0$ from diagram. ☺



Example 1: Let $\{Z_0 = 1, Z_1, Z_2, \ldots\}$ be a branching process with family size distribution $Y \sim \text{Binomial}(2, \frac{1}{4})$, as in Section 7.1. Find the probability of eventual extinction.

Solution: Consider $Y \sim Bin(2, \frac{1}{4})$, then $\mu = \mathbb{E}Y = 2 * \frac{1}{4} = \frac{1}{2} < 1$.

So we know immediately that $\gamma = \mathbb{P}(\text{eventual extinction}) = 1$.

Same answer as in §7.1, but much faster!
(No need to do the longer calculation.)

Example 2: Let $\{Z_0 = 1, Z_1, Z_2, \ldots\}$ be a branching process with family size distribution $Y \sim \text{Geometric}(\frac{1}{4})$, as in Section 7.1. Find the probability of eventual extinction.

Solution: For $Y \sim Geo(\frac{1}{4})$, $\mu = \mathbb{E}Y = \frac{3/4}{1/4} = 3 > 1$.

So we only know $\gamma < 1$.

We still need to solve $G(s) = s$ and find smallest soln ≥ 0 to find what γ is. [See §7.1 : answer was $\gamma = \frac{1}{3}$.]

Note: The mean μ 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 $\mathbb{E}(Z_n) = \mu^n \to 0$ as $n \to \infty$.

all processes in nature aim at this: $\mu = 1$

- 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 $\mathbb{E}(Z_n) = \mu^n \to \infty$ as $n \to \infty$.

hints at the fact that $\mathbb{E}(\text{time to extinction}) = \infty$ even though $\mathbb{P}(\text{extinction}) = 1$.

But how long have you got. . . ?

7.3 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:

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

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

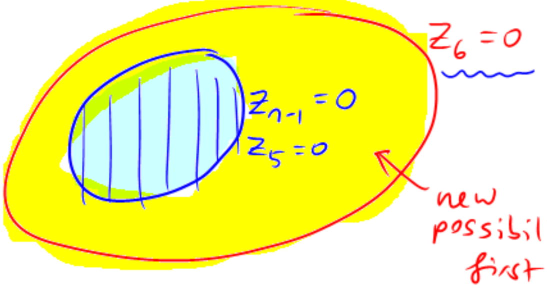
There is no guarantee that the PGF $G_n(s)$ or the value $G_n(0)$ can be calculated easily. However, we can build up $G_n(0)$ in steps:

e.g. $G_2(0) = G(G(0))$ easy

then $G_3(0) = G(G_2(0))$ easy

or $G_4(0) = G_2(G_2(0))$.

e.g. $n-1=5$



new possibility: extinct for the first time at time 6

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:

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

ie. $Z_n=0 \cap Z_{n-1}>0$. ← yellow shading above.

Now by the Partition Rule,

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

what we want

blue shading above: just equal to $\mathbb{P}(Z_{n-1}=0)$

both blue plus yellow shading

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

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

Similarly, $\mathbb{P}(Z_n=0) = G_n(0)$.

So $(\star)$ gives:

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

outer circle minus the blue ring

Relevant to A4 Q2d, 2e and 3c.

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

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

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

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

↓ mathematical induction

$$\vdots$$

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

Thus time to extinction, T, satisfies

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

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 σ^2 is finite;*
- *infinite if $\mu > 1$ (because with positive probability, extinction never happens).*

(Results not proved here.)

For Reference / interest.

7.4 Case Study: Geometric Branching Processes

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

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 144), 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 (γ).
- Find the **full probability distribution of Z_n:** probabilities $\mathbb{P}(Z_n = 0)$, $\mathbb{P}(Z_n = 1)$, $\mathbb{P}(Z_n = 2)$, ...

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

$$G_n(s) = \mathbb{E}\left(s^{Z_n}\right) = \begin{cases} \dfrac{n-(n-1)s}{n+1-ns} & \text{if} \quad p = q = 0.5, \\[2ex] \dfrac{(\mu^n - 1) - \mu(\mu^{n-1}-1)s}{(\mu^{n+1}-1) - \mu(\mu^n - 1)s} & \text{if} \quad p \neq q, \ \text{where } \mu = \frac{q}{p}. \end{cases}$$

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 6),

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

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

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

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

---

2. Exact time of extinction, *T*

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

From Section 7.3,

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

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.

Ass 3 Q 3-④

Shows similar result, proved by induction, for Total Progeny of a BP when $Y \sim$ Geo.

3. Whole distribution of Z_n

From Chapter 4, $\mathbb{P}(Z_n = r) = \frac{1}{r!} G_n^{(r)}(0)$.

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

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

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

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

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

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

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

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

$$\vdots$$

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

(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`