The Galton–Watson Branching Process

Nathanael Marcus

July 10, 2026

Historical Motivation

A Question from 1873

Francis Galton asked a simple question.

Will family names eventually disappear?

Henry William Watson answered this question by introducing a probabilistic model of population growth.

This model became the first **branching process**.

Questions We Will Answer

Throughout this talk we will answer four fundamental questions.

1. How does a branching population evolve?
2. When is extinction certain?
3. How long does extinction take?
4. What happens to populations that survive?

The Galton–Watson Process

The Galton–Watson Process

Definition

A Galton–Watson process is a stochastic process

$$(Z_n)_{n \geq 0},$$

where

$$Z_n = \text{population size in generation } n.$$

The process begins with one ancestor,

$$Z_0 = 1.$$

Each individual independently produces offspring according to the same probability distribution.

Evolution of the Population

If

$$\xi_{n,i}$$

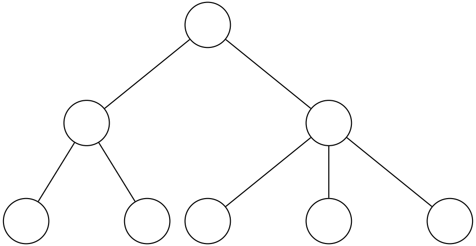
denotes the number of offspring produced by the ith individual in generation n, then

$$\boxed{Z_{n+1} = \sum_{i=1}^{Z_n} \xi_{n,i}.}$$

Key assumptions:

- offspring variables are independent,
- all individuals have the same offspring distribution,
- generations evolve recursively.

Family Tree Interpretation



Each node independently generates a random number of children.

Probability Generating Functions

Probability Generating Functions

Instead of working directly with the infinitely many probabilities

$$p_0, p_1, p_2, \ldots,$$

we encode the entire offspring distribution into one function.

$$\boxed{f(s) = E[s^\xi] = \sum_{k=0}^{\infty} p_k s^k.}$$

This function contains all information about the offspring distribution.

Important Properties

The generating function encodes the entire offspring distribution.

$$f(0) = p_0$$

Probability an individual has no offspring.

$$f(1) = 1$$

The probabilities sum to one.

$$f'(1) = \mu$$

The mean offspring number.

Higher derivatives of the PGF yield higher factorial moments of the distribution.

Function Iteration

The Population Generating Function

The offspring generating function describes one individual.

To study the entire population after n generations, we define

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

Our goal is to relate the unknown function f_n to the known offspring generating function f.

Composition Theorem

Theorem 1 (Composition Theorem)

For every $n \geq 0$,

$$\boxed{f_{n+1}(s) = f_n(f(s)).}$$

Repeated function composition completely determines the evolution of the branching process.

$$f_n = \underbrace{f \circ f \circ \cdots \circ f}_{n \text{ times}}.$$

Why Does Composition Appear?

Proof.

Condition on the current population.

If $Z_n = m$, then

$$Z_{n+1} = \xi_{n,1} + \cdots + \xi_{n,m}.$$

Since the offspring variables are independent,

$$E[s^{Z_{n+1}} \mid Z_n = m] = E[s^{\xi_{n,1}+\cdots+\xi_{n,m}}] = E[s^{\xi_{n,1}}] \cdots E[s^{\xi_{n,m}}] = (f(s))^m.$$

Finally, average over all possible values of m:

$$f_{n+1}(s) = E[s^{Z_{n+1}}] = \sum_{m=0}^{\infty} P(Z_n = m)\, E[s^{Z_{n+1}} \mid Z_n = m] = \sum_{m=0}^{\infty} P(Z_n = m)(f(s))^m = f_n(f(s)).$$

Expected Population Growth

Expected Population Size

Theorem 2

If

$$Z_0 = 1,$$

then

$$\boxed{E(Z_n) = \mu^n.}$$

The expected population depends only on the mean offspring number

$$\mu = E[\xi].$$

Proof Idea

Proof.

We proceed by induction.
Since $Z_0 = 1$, we have $E(Z_0) = 1 = \mu^0$.
Now suppose
$$E(Z_n) = \mu^n.$$
Conditioning on the current generation,
$$\begin{aligned} E(Z_{n+1}) &= E(E(Z_{n+1} \mid Z_n)) \\ &= E(\mu Z_n) = \mu E(Z_n) = \mu^{n+1}. \end{aligned}$$
Thus, by induction,
$$E(Z_n) = \mu^n.$$
□

Three Growth Regimes

Subcritical	Critical	Supercritical
$\mu < 1$	$\mu = 1$	$\mu > 1$
$E(Z_n) \to 0$	$E(Z_n) = 1$	$E(Z_n) \to \infty$
Population shrinks.	Population stays constant on average.	Population grows exponentially.

Why the Mean Is Not Enough

Consider two offspring distributions.

Distribution A

$$P(\xi = 1) = P(\xi = 3) = \frac{1}{2}. \qquad \mu = 2.$$

Every individual has at least one child.

Distribution B

$$P(\xi = 0) = P(\xi = 4) = \frac{1}{2}. \qquad \mu = 2.$$

The population can become extinct immediately.

Extinction Probability

It is clear that $E(Z_n)$ on its own is insufficient and does not fully describe the process.

Instead, we will now look at the probability of extinction.

Definition 3

The probability of extinction

$$q = P(\exists\, n \geq 0 : \; Z_n = 0).$$

The Fixed-Point Equation

Theorem 4

The extinction probability satisfies

$$\boxed{q = f(q).}$$

Thus computing the extinction probability reduces to solving a single equation instead of analyzing infinitely many possible family trees.

Why is $q = f(q)$?

Proof.

Condition on the initial ancestor having k offspring.

$$P(\text{extinction} \mid \xi = k) = q^k,$$

since each offspring independently starts a Galton–Watson process.
Using the Law of Total Probability,

$$\begin{aligned} q &= \sum_{k=0}^{\infty} P(\xi = k)\, P(\text{extinction} \mid \xi = k) \\ &= \sum_{k=0}^{\infty} p_k q^k = f(q). \end{aligned}$$

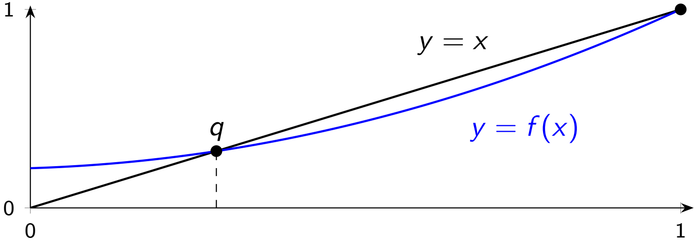
$\square$

Fixed Points

The equation $q = f(q)$ means that the extinction probability is a fixed point of the generating function.

Graphically, fixed points are the intersections of

$$y = f(x) \qquad \text{and} \qquad y = x.$$



Which Fixed Point?

Theorem 5 (Smallest Fixed Point)

The extinction probability is the smallest solution of

$$f(x) = x, \qquad x \in [0,1].$$

Idea.

Let $q_n = P(Z_n = 0)$.
Since extinction is permanent, $q_0 \leq q_1 \leq q_2 \leq \cdots \leq 1$, so $q_n \to q$.

Conditioning on the number of offspring of the initial ancestor gives $q_{n+1} = f(q_n)$.

Now let r satisfy $f(r) = r$.

Since $q_0 = 0 \leq r$, and f is increasing, $q_1 = f(q_0) \leq f(r) = r$.
Repeating this argument gives $q_n \leq r$ for every n.

Taking limits, $q \leq r$, so q is the smallest fixed point.

The Critical Threshold

Critical Threshold Theorem

Theorem 6

$$\boxed{\mu \leq 1 \Longrightarrow q = 1.}$$

$$\boxed{\mu > 1 \Longrightarrow 0 < q < 1.}$$

The mean offspring number determines whether extinction is certain.

Idea of the Proof

Idea.

We define $g(s) = f(s) - s$

Recall that fixed points satisfy $g(s) = f(s) - s = 0$.

Suppose a second fixed point $q < 1$ exists. Then $g(q) = g(1) = 0$.
Thus, by the Mean Value Theorem, there exists $c \in (q, 1)$ such that $g'(c) = 0$, so

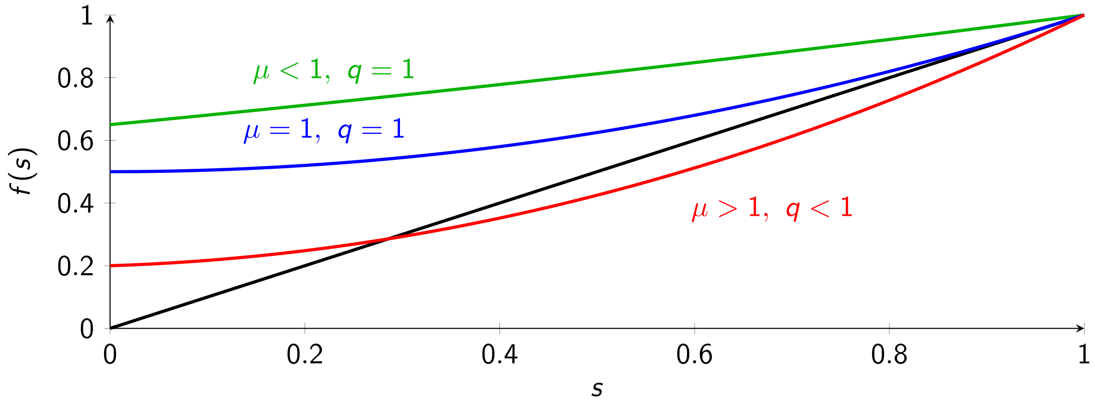
$$f'(c) = 1.$$

Since f is convex, f' is increasing, giving $1 = f'(c) < f'(1) = \mu$.
Thus, a second fixed point can exist only when

$$\boxed{\mu > 1.}$$

Conversely, when $\mu > 1$, the Intermediate Value Theorem guarantees a second fixed point. □

Geometric Interpretation



As μ passes through 1, a second fixed point appears in $(0, 1)$, giving a positive probability of survival.

Extinction Time

Extinction Time

The extinction probability answers

Will the population die out?

The extinction time answers

When will it die out?

Definition

The extinction time is

$$T = \min\{n \geq 0 : Z_n = 0\},$$

with $T = \infty$ if extinction never occurs.

Unlike the extinction probability, no fixed-point equation determines T.

A Surprising Result

Theorem 7

If

$$\mu = 1,$$

and the offspring distribution has finite positive variance,
then

$$\boxed{E(T) = \infty.}$$

Even though

$$P(\text{eventual extinction}) = 1.$$

Why?

Proof.

Since T is a non-negative integer-valued random variable

$$E(T) = \sum_{n=0}^{\infty} P(T > n).$$

In the critical case,

$$P(T > n) \sim \frac{2}{\sigma^2 n}.$$

Therefore,

$$E(T) \sim \sum_{n=1}^{\infty} \frac{1}{n},$$

which diverges.

Long-Term Growth

Normalizing Population Growth

For a supercritical process,

$$E(Z_n) = \mu^n.$$

Remove this deterministic growth by defining

$$W_n = \frac{Z_n}{\mu^n}.$$

Theorem 8

$$(W_n)$$

is a martingale.

The normalized process has no expected upward or downward drift.

Martingale Convergence

A fundamental theorem states that

$$W_n \longrightarrow W$$

almost surely.

Consequently,

$$\boxed{Z_n \sim W\mu^n.}$$

Every surviving population grows at the same exponential rate, but the constant W is random.

Modern Branching Processes

Beyond the Classical Model

Many modern models extend the original Galton–Watson process.

- Multi-type branching processes
- Continuous-time branching processes
- Branching processes conditioned on survival

These models describe far more realistic populations while preserving the core branching idea.

Applications

Branching processes now appear throughout mathematics and science.

- Epidemic modeling
- Population genetics
- Cell division
- Ecology
- Information spread
- Social networks
- Computer viruses
- Randomized algorithms

Takeaways

1. Probability generating functions encode the entire branching process.
2. Extinction is determined by fixed points of the generating function.
3. The critical threshold occurs at

$$\mu = 1.$$

4. Martingales explain the long-term growth of surviving populations.

Questions?