Galton-Watson processes

Emilio Leonardi

Notes for the class on “Computer aided simulations labs”

Politecnico di Torino

November 2022

Outline

Section 1

Problem definition

GW processes

The structure of GW process

- Sample paths are trees generated as follows:
- The tree originates from an ancestor/root (generation-0 vertex)
- Every generation-i vertex v in the tree produces a random number Y_v of generation-$i+1$ vertices (the v-children).
- variables Y_v are i.i.d. random variables.
- the process stops if/when an empty generation is obtained

Properties of the tree

- Is it finite or infinite?
- how many vertices/generations?

Is it finite or infinite?

Possible behaviors

- Certain extinction (the tree is always finite)
- Probabilistic extinction (the tree may be either finite or infinite)
- Extinction does not occur (the tree is always infinite)

Trivial cases

- $Y_v \geq 1$;
- $Y_v = 0$.

Non trivial case

- $0 < \mathbb{P}(Y_v = 0) < 1$.

Some more detail

Distribution of Y_v

- Y_v is a discrete random variable
- $\mathbb{P}(Y_v = k) = p_k$ for $k = 0, 1, 2, \cdots$ is given

Independence

- Variables Y_v are independent!

Generations

- We denote with X_i the number of vertices in generation-i;
- by construction $X_0 = 1$

Section 2

Preliminary Analysis

computing X_i

- How to obtain X_i?
- Through a recursion: express X_i in terms of X_{i-1}
- given X_{i-1}:

Result

- We have X_{i-1} vertices in generation $i-1$;
- Every vertex in generation $i-1$ originates Y_i children, in a independent fashion; therefore

$$X_i = \sum_{j=1}^{X_{i-1}} Y_j$$

Extinction

- Extinction occurs when a generation contains no vertices
  - i.e., $X_i = 0$

Probability of extinction

- Probability of extinction within generation i:

$$q_i := \mathbb{P}(X_i = 0)$$

- What happens to q_i, as $i \to \infty$?

Asymptotic extinction probability

$$q := \lim_{i\to\infty} q_i$$

- Note that by construction $q_{i+1} \geq q_i$ (therefore q is well defined).

Asymptotic Extinction

- Our goal are:
  1. to relate the asymptotic extinction q to properties of Y_v
     - and in particular to the distribution of the number of the children
  2. to characterize qualitative properties of the process
     - discriminate between sure death and chance to survive forever!
- Some guess?

Average number of generation-i vertices

- how to compute $\mathbb{E}[X_i]$, i.e. the expected number of vertices belonging to generation-i?

generation-0

$\mathbb{E}[X_0] = 1$ since $X_0 = 1$

generation-1

- How to compute $\mathbb{E}[X_1]$?
  - by construction $X_1 = Y_{\text{root}}$
  - therefore $\mathbb{E}[X_1] = \mathbb{E}[Y_{\text{root}}] = m$
  - m is called average reproduction factor (finite and known)

Computing $\mathbb{E}[X_2]$

How to compute $\mathbb{E}[X_2]$?

Recalling that $X_2 = \sum_{j=1}^{X_1} Y_j$, we can compute $\mathbb{E}[X_2]$ by conditioning on X_1:

Tower property for expectations

$\mathbb{E}[X_2] = \mathbb{E}_{X_1}\mathbb{E}[X_2 \mid X_1]]$

- Now, in our case $X_2 \mid \{X_1 = k\} = \sum_1^k Y_j$,
    - therefore $\mathbb{E}[X_2 \mid X_1 = k] = \mathbb{E}[\sum_1^k Y_j] = mk$
    - i.e., $\mathbb{E}[X_2 \mid X_1] = mX_1$
    - and $\mathbb{E}[X_2] = \mathbb{E}_{X_1}\mathbb{E}[X_2 \mid X_1]] = \mathbb{E}_{X_1}[mX_1] = m\mathbb{E}[X_1] = m^2$

Computing $\mathbb{E}[X_i]$

By conditioning we can compute $\mathbb{E}[X_i]$

- Indeed:
  - Now, in our case $X_i \mid \{X_{i-1} = k\} = \sum_{j=1}^{k} Y_j$,
    - therefore $\mathbb{E}[X_i \mid X_{i-1} = k] = \mathbb{E}[\sum_1^k Y_j] = mk$
    - i.e., $\mathbb{E}[X_i \mid X_{i-1}] = mX_{i-1}$
    - and $\mathbb{E}[X_i] = \mathbb{E}_{X_{i_1}} \mathbb{E}[X_i \mid X_{i-1}]] = \mathbb{E}_{X_{i-1}}[mX_{i-1}] = m\mathbb{E}[X_{i-1}]$

Therefore by induction over i we obtain $\mathbb{E}[X_i] = m^i$

Section 3

Regimes

Different possible behaviors

Depending on m, $\mathbb{E}[X_i]$ exhibits three possible behaviors:

- if $m < 1$ $\rightarrow \mathbb{E}[X_i] \rightarrow 0$ **subcritical**
- if $m = 1$ $\rightarrow \mathbb{E}[X_i] = 1, \forall i$ **critical**
- if $m > 1$ $\rightarrow \mathbb{E}[X_i] \rightarrow \infty$ **supercritical**

Section 4

Analysis of the subcritical regime

Subcritical regime

If $m < 1$, what happens to $q = \lim_{i \to \infty} q_i$?

- Intuitively, since $\mathbb{E}[X_i] \to 0$ we may expect that $q = 1$;
- we can prove formally it by exploiting Markov inequality.

Markov inequality

Markov inequality

Given a R.V. $X \geq 0$ with $\mathbb{E}[X] < \infty$ then:

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a} \qquad \forall a > \mathbb{E}[X].$$

Application of the Markov inequality

Setting $a = 1$, we have:

$$\mathbb{P}(X_i \geq 1) \leq \mathbb{E}[X_i] = m^i$$

Letting $i \to \infty$ we obtain:

$$q := \lim_{i\to\infty} q_i = \lim_{i\to\infty} \mathbb{P}(X_i = 0) = \lim_{i\to\infty} 1 - \mathbb{P}(X_i \geq 1) \geq 1 - m^i \to 1.$$

Section 5

Analysis of the supercritical regime

Preliminary considerations

What happens to $q = \lim_{i \to \infty} q_i$ in the supercritical regime?

- when $m > 1$, $\mathbb{E}[X_i] \to \infty$,
- we would be tempted to bet on $q = 0$, but:

Caution!

Note that: $q \geq q_1 = \mathbb{P}(X_1 = 0) = \mathbb{P}(Y_{\text{root}} = 0) = p_0 > 0$, therefore extinction may occur!

Moment/Probability Generating Function

The moment/probability generation function of a discrete R.V. $X \geq 0$ is:

$$\phi_X(z) = \mathbb{E}[z^X] = \sum_{k=0}^{\infty} \mathbb{P}(X = k) z^k$$

- a.k.a. it is the z-trasform of its distribution!

It allows the manipulation of probabilities and the computation moments of distributions. For example

$$\phi_X(0) = \mathbb{P}(X = 0) \qquad \frac{1}{k!} \frac{\mathrm{d}^k \phi_X(z)}{\mathrm{d}z^k} \,|_{z=0} = \mathbb{P}(X = k)$$

MGF: important properties

- Given $S = X_1+, \cdots, +X_k$, with $\{X_i\}_i$ i.i.d. we have:

$$\begin{aligned}\phi_S(z) =& \mathbb{E}[z^S] = \mathbb{E}[z^{X_1+X_2+,\cdots,+X_k}] = \mathbb{E}[z^{X_1} \cdots z^{X^2} \cdots \cdots z^{X_k}] \\ =& \mathbb{E}[z^{X_1}]\mathbb{E}[z^{X_2}] \cdots \mathbb{E}[z^{X_k}] = \phi_X^k(z)\end{aligned}$$

- i.e., the MGF of a sum of k *i.i.d* r.v.s is equal to the k-th power of the MGF of each individual r.v.

MGF: important properties (cnt)

- Given $S = X_1 +, \cdots, +X_k$, with $\{X_i\}_i$ i.i.d. and given r.v. K independent from $\{X_i\}_i$, defined $S = \sum_{i=1}^{i=K} X_i$ we have:

$$\phi_S(z) = \phi_K(\phi_X(z))$$

- The proof can be easily obtained by conditioning on K (i.e. applying the tower property).

Supercritical regime

Getting back to our problem, we have:

$$X_i = \sum_{i=1}^{X_{i-1}} Y_j$$

Therefore

$$\phi_{X_i}(z) = \phi_{X_{i-1}}(\phi_Y(z)) \quad \text{with} \quad \phi_{X_1}(z) = \phi_Y(z)$$

hence $\phi_{X_2}(z) = \phi_Y(\phi_Y(z))$ and by induction over i

$$\phi_{X_i}(z) = \phi_Y(\phi_Y(\phi_Y(\cdots))) = \phi_Y(\phi_{X_{i-1}}(z))$$

Asymptotic Extinction in the supercritical regime

- By definition $q_i = \mathbb{P}(X_i = 0) = \Phi_{X_i}(0)$
- and since $\Phi_{X_i}(0) = \Phi_Y(\Phi_{X_{i-1}}(0))$ item We have $q_i = \phi_Y(q_{i-1})$

Therefore, observing that $\phi_Y(z)$ is surely smooth (continuous and indefinitely derivable) for $|z| < 1$ (under mild assumptions is smooth on a larger domain).

$$q = \lim_{i\to\infty} q_i = \lim_{i\to\infty} \phi_Y(q_{i-1}) = \phi_Y(\lim_{i\to\infty} q_{i-1}) = \phi_Y(q)$$

i.e. $q \in (0,1)$ satisfies the equation:

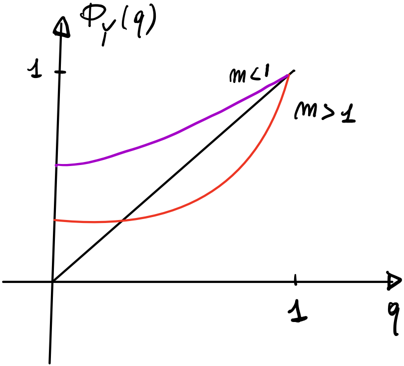
$$q = \phi_Y(q)$$

Properties of $\phi_Y(z)$ and their consequences

Given that:

- $\phi_Y(0) = \mathbb{P}(Y = 0) = p_0 > 0$
- $\phi_Y(1) = \sum_{k=0}^{\infty} \mathbb{P}(Y = k) = 1$
- $\frac{\mathrm{d}\phi_Y(x)}{\mathrm{d}z} \mid_{z=1} = \mathbb{E}[Y] = m$
- $\phi_Y(z)$ is strictly increasing and convex for $z \in [0, 1]$

Equation $q = \phi_Y(q)$ has exactly one solution q^* in $(0, 1)$ if $m > 1$ and no solutions in $(0, 1)$ if $m \leq 1$.



$\Phi_Y(q)$
1
$m<1$
$m>1$
1
q

Further considerations for $m > 1$

Since $\phi_Y(z)$ is strictly increasing, note that:

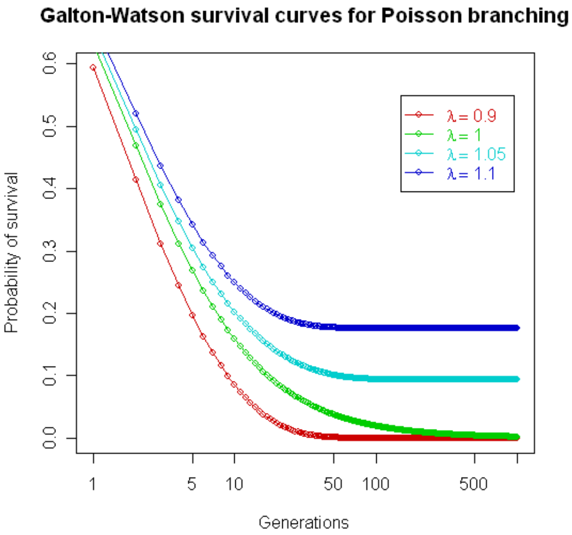
- $q_1 = \phi_Y(0) \leq \phi_Y(q^*) = q^*$
- assume $q_{i-1} < q^*$ then $q_i = \phi_Y(q_{i-1}) \leq \phi_Y(q^*) = q^*$

By induction $q_i < q_*$ for any $i \in \mathbb{N}$.

In conclusion, since on the one hand $q = \lim_{i\to\infty} q_i = \sup_i q_i \leq q^*$, and on the other hand, as already shown, q must satisfy the equation $q = \phi_Y(q)$, necessarily we have:

for $m > 1$

$$q = q^*.$$


Galton-Watson survival curves for Poisson branching
Probability of survival
0.0
0.1
0.2
0.3
0.4
0.5
0.6
Generations
1
5
10
50
100
500
λ = 0.9
λ = 1
λ = 1.05
λ = 1.1

Self-similar/fractal structure

Compare the structure of the whole GW tree with the structure of a subtree routed in a given vertex v:

self-similarity/ fractal structure

They are identical! Tree and sub-tree are generated according to exactly the same algorithm!

We can exploit this property to derive the MGF of the total number of vertices in the tree.
Let N_0 denote the number of vertices in the whole tree. Let N_i denote the number of nodes in the sub-tree rooted in i-th children of the root. By construction

$$N_0 = 1 + \sum_{i=1}^{Y_{root}} N_i$$

MGF for N_0 ($m < 1$)

Therefore

$$\phi_{N_0}(z) := \mathbb{E}[z^{N_0}] = \mathbb{E}[z^{1+\sum_{i=1}^{Y_{root}} N_i}] = z\mathbb{E}[z^{\sum_{i=1}^{Y_{root}} N_i}] = z\phi_Y(\phi_{N_1}(z))$$

given that N_i are obliviously i.i.d. and independent from Y_{root}.
At last $\phi_{N_i}(z) = \phi_{N_0}(z)$, because of the self-similarity, and therefore $\phi_{N_0}(z)$ satisfies the following functional equation:

$$\phi_{N_0}(z) = z\phi_Y(\phi_{N_0}(z))$$