MS-C2111 Stochastic Processes

Lecture 6

Branching processes

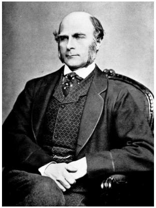
Jukka Kohonen
Aalto University

Contents

Galton's problem

PROBLEM 4001: A large nation, of whom we will only concern ourselves with the adult males, N in number, and who each bear separate surnames, colonise a district. Their law of population is such that, in each generation, p_0 per cent of the adult males have no male children who reach adult life; p_1 have one such male child; p_2 have two; and so on up to p_5 who have five.

Find what proportion of the surnames will have become extinct after t generations.



Sir Francis Galton
(1822–1911)

F Galton. Educational Times 1873.

H W Watson & F Galton. The Journal of the Anthropological Institute of Great Britain and Ireland 1875.

I-J Bienaymé. Soc. Philomat. Paris Extraits 1845.

Applications

- COVID-19, future epidemics
- Online social media, block chains
- Biology (bacteria, cell division, ecology)
  https://www.biointeractive.org/classroom-resources/bacterial-growth
- Particle physics

Contents

Contents

Branching process

Everyone in generation t independently produces a random number of children, and these children will form the next generation $t+1$

- $X_t =$ size of generation $t = 0, 1, \ldots$

Model parameters:

- Initial state X_0 (or initial distribution μ_0)
- Offspring distribution $(p_0, p_1, \ldots)$ where p_k is the probability of producing k children

The random sequence $(X_0, X_1, \ldots)$ is called a Galton–Watson process

Markov property

Because every individual in generation t produces children independently of others,

$$\mathbb{P}(X_{t+1} = j \,|\, X_t = i, X_{t-1} = i_{t-1}, \ldots, X_0 = i_0)$$
$$= \mathbb{P}(X_{t+1} = j \,|\, X_t = i)$$

- $(X_0, X_1, \ldots)$ is hence a Markov chain on state space $\mathbb{Z}_+$
- The transition matrix $P : \mathbb{Z}_+ \times \mathbb{Z}_+ \to [0,1]$ has entries

$$P(i,j) = \mathbb{P}(X_{t+1} = j \,|\, X_t = i) = \mathbb{P}(Y_1 + \cdots Y_i = j),$$

where $Y_1, Y_2, \ldots$ are independent and (p_k)-distributed random numbers representing the children counts of individuals in generation t

Transition matrix

$$P(i,j) \;=\; \mathbb{P}(Y_1 + \cdots Y_i = j)$$

- $P(0,j) = 0$ for $j \geq 1$ and $P(0,0) = 1$
- $P(1,j) = \mathbb{P}(Y_1 = j) = p_j$

$$\begin{aligned} P(2,j) = \mathbb{P}(Y_1 + Y_2 = j) &= \sum_{i=0}^{j} \mathbb{P}(Y_1 = i, Y_1 + Y_2 = j) \\ &= \sum_{i=0}^{j} \mathbb{P}(Y_1 = i, Y_2 = j - i) = \sum_{i=0}^{j} p_i p_{j-i} \end{aligned}$$

- $P(3,j) = \ldots$

Number of grandchildren

What is the number of grandchildren of a particular individual?

The descendants of an individual form a branching process with initial state $X_0 = 1$.

- The number of children X_1 is distributed as

$$\mathbb{P}(X_1 = k | X_0 = 1) \ = \ P(1, k) \ = \ p_k$$

- The number of grandchildren X_2 is distributed as

$$\mathbb{P}(X_2 = k | X_0 = 1) \ = \ P^2(1, k)$$

More generally, the number of descendants in the t-th generation is X_t is distributed as

$$\mathbb{P}(X_t = k | X_0 = 1) \ = \ P^t(1, k)$$

. . . Number of grandchildren

How do we compute the distribution of grandchildren $k \mapsto P^2(1, k)$ from the offspring distribution (p_k)?

- P is infinite so we cannot directly compute P^2 using a computer
- Sums of independent random variables are easy to treat using generating functions

Generating functions

- Characteristic function $s \mapsto \mathbb{E}e^{isY}$
- Moment generating function $s \mapsto \mathbb{E}e^{sY}$
- Cumulant generating function $s \mapsto \log \mathbb{E}e^{sY}$
- Probability generating function $s \mapsto \mathbb{E}s^Y$

When $Y \in \mathbb{Z}_+$, then $s \mapsto \mathbb{E}s^Y$ is usually the most convenient.

Contents

Probability generating function

The probability generating function (pgf) of $Y \in \mathbb{Z}_+$ is defined by

$$\phi_Y(s) = \mathbb{E}s^Y = \sum_{k=0}^{\infty} s^k \mathbb{P}(Y = k)$$

for those s where the sum on the right converges.

Note

- $\phi_Y(s)$ is defined for all $s \in [-1, 1]$ because

$$\sum_{k=0}^{\infty} |s|^k \mathbb{P}(Y = k) \ \leq \ \sum_{k=0}^{\infty} \mathbb{P}(Y = k) \ = \ 1 \ < \ \infty.$$

- Y and Y' have the same law $\implies$ Y and Y' have the same pgf
- Y and Y' have the same pgf $\implies$ Y and Y' have the same law

Probability generating function — properties

$$\phi_Y(s) = \mathbb{E}s^Y = \sum_{k=0}^{\infty} s^k \mathbb{P}(Y = k)$$

For every random integer $Y \in \mathbb{Z}_+$:

- ϕ_Y is continuous, nondecreasing, and convex on $[0, 1]$.
- $\phi_Y(0) = \mathbb{P}(Y = 0)$ and $\phi_Y(1) = 1$.
- $\mathbb{P}(Y = k) = \phi_Y^{(k)}(0)/k!$ for all $k \in \mathbb{Z}_+$

If $\phi_Y(r)$ exists for some $r > 1$:

- ϕ_Y is infinitely differentiable on $(-r, r)$
- $\mathbb{E}Y^k < \infty$ for all $k = 1, 2, \ldots$
- $\phi'_Y(1) = \mathbb{E}Y$
- $\phi''_Y(1) = \mathbb{E}Y^2 - \mathbb{E}Y$
- $\mathrm{var}(Y) = \mathbb{E}Y^2 - (\mathbb{E}Y)^2 = \phi''_Y(1) + \phi'_Y(1) - \phi'_Y(1)^2$

Pgf of a sum

Theorem

Let $Y = \sum_{i=1}^{n} Y_i$ *where the random summands* $Y_1, \dots, Y_n \in \mathbb{Z}_+$ *are independent. Then for all* $s \in [-1, 1]$*,*

$$\phi_Y(s) = \phi_{Y_1}(s) \cdots \phi_{Y_n}(s).$$

Proof.

By independence,

$$\phi_Y(s) = \mathbb{E}s^Y = \mathbb{E}(s^{Y_1} \cdots s^{Y_n}) = \mathbb{E}(s^{Y_1}) \cdots \mathbb{E}(s^{Y_n}) = \phi_{Y_1}(s) \cdots \phi_{Y_n}(s).$$

Note

If $Y_1, \dots, Y_n$ are IID,

$$\phi_Y(s) = \phi_{Y_1}(s)^n$$

$\square$

Pgf of a sum — random sum index

What happens when the sum index N is random?

Theorem

Let $Y = \sum_{i=1}^{N} Y_i$, *where* $N, Y_1, Y_2, \dots \in \mathbb{Z}_+$ *are independent and the summands* $Y_1, Y_2, \dots$ *are identically distributed. Then for all* $s \in [-1, 1]$,

$$\phi_Y(s) = \phi_N(\phi_{Y_1}(s)).$$

Pgf of a sum — random sum index

Proof.

By conditioning on N we find that

$$
\begin{aligned}
\phi_Y(s) &= \mathbb{E}\left(s^{\sum_{i=1}^{N} Y_i}\right) \\
&= \sum_{n=0}^{\infty} \mathbb{E}\left(s^{\sum_{i=1}^{n} Y_i} \mid N = n\right) \mathbb{P}(N = n) \\
&= \sum_{n=0}^{\infty} \mathbb{E}\left(s^{\sum_{i=1}^{n} Y_i}\right) \mathbb{P}(N = n) \\
&= \sum_{n=0}^{\infty} \left(\mathbb{E} s^{Y_1}\right)^n \mathbb{P}(N = n) \\
&= \sum_{n=0}^{\infty} \phi_{Y_1}(s)^n \, \mathbb{P}(N = n) \\
&= \phi_N(\phi_{Y_1}(s)).
\end{aligned}
$$

□

Contents

Transition matrix of a branching process

Theorem

The entries of the transition matrix on row i satisfy

$$\sum_{j=0}^{\infty} P(i,j)\, s^j \;=\; \phi(s)^i$$

where $\phi(s) = \sum_{k=0}^{\infty} p_k s^k$ is the pgf of the offspring distribution (p_k).

Note

- $P(i,j)$ is hence the j-th term of the power series of $\phi(s)^i$
- $P(i,j)$ is obtained by differentiating $\phi(s)^i$ j times at zero:

$$P(i,j) = \frac{\left[\left(\frac{d}{ds}\right)^j \phi(s)^i\right]_{s=0}}{j!}$$

Transition matrix of a branching process

Proof.

The (i,j)-entry of the transition matrix can be written as

$$P(i,j) \;=\; \mathbb{P}(S=j),$$

where $S = Y_1 + \cdots + Y_i$ is the sum of independent (p_k)-distributed random numbers and Y_i represents the number of children of individual i in generation zero.

Hence

$$\sum_{j=0}^{\infty} P(i,j)s^j \;=\; \sum_{j=0}^{\infty} s^j \mathbb{P}(S=j) \;=\; \phi_S(s) \;=\; \phi_{Y_1}(s)^i \;=\; \phi(s)^i.$$

□

Example

Initially there are two individuals, each producing 3 children with probability $a = 0.1$ and 0 children otherwise. What is the probability that the next generation contains 6 individuals?

The pgf $\phi(s)$ of the offspring distribution satisfies

$$\begin{aligned} \phi(s) &= (1-a) + as^3, \\ \phi^2(s) &= (1-a)^2 + 2(1-a)as^3 + a^2 s^6. \end{aligned}$$

By the previous theorem,

$$\sum_{j=0}^{\infty} P(2,j)s^j = (1-a)^2 + 2(1-a)as^3 + a^2 s^6,$$

so that $\mathbb{P}(X_1 = 6 \mid X_0 = 2) = P(2,6) = a^2 = 0.01$.

Number of grandchildren

The descendants of any particular individuals form a branching process with $X_0 = 1$.

- Number of children X_1: $\mathbb{P}(X_1 = k) \ = \ P(1,k) \ = \ p_k$
- Number of grandchildren X_2: $\mathbb{P}(X_2 = k) \ = \ P^2(1,k) = ?$

The number of grandchildren can be represented as

$$X_2 \ = \ \sum_{i=1}^{X_1} Y_i,$$

where $X_1, Y_1, Y_2, \ldots$ are IID and (p_k)-distributed. Hence the pgf of X_2 is

$$\phi_{X_2}(s) = \phi_{X_1}(\phi_{Y_1}(s)) = \phi(\phi(s)).$$

Example

Initially there are two individuals, each producing 3 children with probability $a = 0.1$ and 0 children otherwise. What is the probability that one of the initial individuals gets 6 grandchildren?

The pgf of the offspring distribution $\phi(s) = (1-a) + as^3$ satisfies

$$\begin{aligned}\phi(\phi(s)) &= (1-a) + a\phi(s)^3 \\ &= (1-a) + a(1-a)^3 + 3a^2(1-a)^2 s^3 + 3a^3(1-a)s^6 + a^4 s^9.\end{aligned}$$

By the previous slide,

$$\phi_{X_2}(s) = \sum_{k=0}^{\infty} P^2(1,k)s^k = \phi(\phi(s)),$$

so that $\mathbb{P}(X_2 = 6 \mid X_0 = 1) = P^2(1,6) = 3a^3(1-a) = 0.0027$.

Number of descendants in t-th generation

Theorem

The pgf of the size of generation t for a branching process starting with $X_0 = 1$ is

$$\phi_{X_t}(s) \; = \; (\underbrace{\phi \circ \cdots \circ \phi}_{t})(s).$$

Proof.

The size of generation $t+1$ can be written as $X_{t+1} = \sum_{i=1}^{X_t} Y_i$, where $X_t, Y_1, Y_2, \ldots$ are independent and $Y_1, Y_2, \ldots$ are (p_k)-distributed. Hence

$$\phi_{X_{t+1}}(s) = \phi_{X_t}(\phi_{Y_1}(s)) = \phi_{X_t}(\phi(s)).$$

The claim follows by induction (case $t = 1$ is clearly OK). $\square$

Contents

Expected population size

Theorem

A branching process with initial state $X_0 = i$ satisfies

$$\mathbb{E}_i(X_t) = im^t, \quad t = 0, 1, 2, \ldots,$$

where $m = \sum_{k=0}^{\infty} kp_k$ is the mean of the offspring distribution.

Note

- If $m < 1$, then $\mathbb{E}X_t \to 0$ exponentially fast.
- If $m = 1$, then $\mathbb{E}X_t = i$ for all t.
- If $m > 1$, then $\mathbb{E}X_t \to \infty$ exponentially fast.

Infectious diseases: *Basic reproduction number R_0 = Average number of infections caused by a typical infected individual during early stage of outbreak.*

Proof.

The size of generation $t+1$ can be written as

$$X_{t+1} \;=\; \sum_{i=1}^{X_t} Y_i,$$

where $X_t, Y_1, Y_2, \ldots$ are independent and $Y_1, Y_2, \ldots$ are (p_k)-distributed. By conditioning on X_t,

$$\begin{aligned}
\mathbb{E}X_{t+1} &= \sum_{k=0}^{\infty} \mathbb{P}(X_t = k)\mathbb{E}(X_{t+1} \,|\, X_t = k) \\
&= \sum_{k=0}^{\infty} \mathbb{P}(X_t = k)\mathbb{E}(\sum_{i=1}^{k} Y_i) \\
&= \sum_{k=0}^{\infty} \mathbb{P}(X_t = k)km = m\mathbb{E}X_t.
\end{aligned}$$

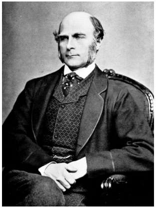
Induction $\implies$ $\mathbb{E}X_t = m^t \mathbb{E}X_0 = im^t$. $\square$

Contents

Galton's problem

PROBLEM 4001: A large nation, of whom we will only concern ourselves with the adult males, N in number, and who each bear separate surnames, colonise a district. Their law of population is such that, in each generation, p_0 per cent of the adult males have no male children who reach adult life; p_1 have one such male child; p_2 have two; and so on up to p_5 who have five.

Find what proportion of the surnames will have become extinct after t generations.



Sir Francis Galton
(1822–1911)

F Galton. Educational Times 1873.

H W Watson & F Galton. The Journal of the Anthropological Institute of Great Britain and Ireland 1875.

I-J Bienaymé. Soc. Philomat. Paris Extraits 1845.

Extinction in finite time horizon

What is the probability that an individual in generation 0 has no descendants in generation t?

Because the descendants form a branching process (X_t) with initial state $X_0 = 1$, this probability is

$$\eta_t \;=\; \mathbb{P}(X_t = 0) \;=\; P^t(1,0).$$

Recall that

- $\mathbb{P}(X_t = 0)$ is the constant term in the power series of $\phi_{X_t}(0)$
- The number of descendants in the t-th generation has pgf $\phi_{X_t}(s) \;=\; (\underbrace{\phi \circ \cdots \circ \phi}_{t})(s)$.

The probabilities η_t are obtained recursively from

- $\eta_1 = \phi(0)$,
- $\eta_2 = \phi(\phi(0)) = \phi(\eta_1)$, ...
- $\eta_{t+1} \; = \phi(\eta_t)$ for all $t \geq 0$.

Extinction eventually

Is it possible for an individual to have infinitely many descendants, or does the family line eventually become extinct?

The probability of eventual extinction is

$$\eta \ = \ \mathbb{P}(X_t = 0 \text{ for some } t \in \mathbb{Z}_+).$$

Note

- $\eta = \mathbb{P}_1(T_0 < \infty)$ is the hitting probability of state 0 for the Markov chain starting at state 1.

Extinction eventually

Theorem

Extinction probability η = *Smallest fixed point of the pgf* $\phi(s)$ *in* $[0, 1]$.

Proof.

$$\eta = \mathbb{P}(\cup_{t=0}^{\infty}\{X_t = 0\}) = \lim_{t\to\infty} \mathbb{P}(X_t = 0) = \lim_{t\ \to\infty} \eta_t.$$

Because ϕ is continuous on $[0, 1]$, the recursive equation $\eta_t = \phi(\eta_{t-1})$ implies $\eta = \phi(\eta)$. Hence η is a fixed point of ϕ.

Let us show that η is the smallest fixed point.
If $\phi(a) = a$ for some $a \in [0, 1]$, then because ϕ is nondecreasing:
$a \geq 0 \implies a = \phi(a) \geq \phi(0) = \eta_1 \implies a \geq \eta_1$
$a \geq \eta_1 \implies a = \phi(a) \geq \phi(\eta_1) = \eta_2 \implies a \geq \eta_2$
Hence $a \geq \eta_t$ for all t, so that $a \geq \eta$. $\square$

Example

In a population every individual produces twins with probability a and otherwise no children. What is the probability of eventual extinction of the family line of an individual?

The pgf of the offspring distribution is $\phi(s) = (1-a) + as^2$.
The fixed points of ϕ are the zeros of $as^2 - s + (1-a) = 0$:

$$s = \frac{1 \pm \sqrt{1-4a(1-a)}}{2a} = \frac{1 \pm \sqrt{(1-2a)^2}}{2a} = \begin{cases} (1-a)/a, \\ 1. \end{cases}$$

The extinction probability is hence

$$\eta = \begin{cases} 1, & \text{if } a \leq 1/2, \\ \frac{1-a}{a}, & \text{if } a > 1/2. \end{cases}$$

Sure extinction for $m < 1$

Theorem

If the mean of the offspring distribution is $m < 1$, then the branching process becomes extinct with probability one.

Proof.

The probability that the population is alive at time t is

$$1 - \eta_t = \mathbb{P}(X_t \geq 1) = \sum_{k=1}^{\infty} \mathbb{P}(X_t = k) \leq \sum_{k=1}^{\infty} k\mathbb{P}(X_t = k) = \mathbb{E}X_t.$$

Because $\mathbb{E}X_t = m^t \mathbb{E}X_0 \to 0$, we see from this that

$$1 - \eta = \lim_{t\to\infty} (1 - \eta_t) \leq \lim_{t\to\infty} \mathbb{E}X_t = 0.$$

Hence $\eta = 1$. $\square$

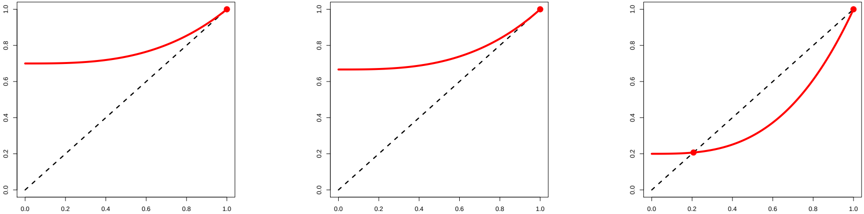
Sure extinction — General characterization

Theorem

For every nontrivial offspring distribution ($0 < p_0 < 1$), $\eta = 1$ if and only if $m \leq 1$.

Proof idea.

The slope of $\phi(s)$ at $s = 1$ is $m = \phi'(1)$.



$m < 1$ $\quad$ $\eta = 1$

$m = 1$ $\quad$ $\eta = 1$

$m > 1$ $\quad$ $\eta \approx 0.207$

□

No sustainability

- If $m < 1$, the population surely dies out eventually.
- If $m = 1$, the population surely dies out eventually, but the expected population size remains constant:

$$\mathbb{E}X_t = m^t \mathbb{E}X_0 = \mathbb{E}X_0 \quad \text{for all } t \geq 0.$$

- If $m > 1$, the population may survive in the long run, and has exponential mean growth:

$$\mathbb{E}X_t = m^t \mathbb{E}X_0 \to \infty.$$

Thomas Robert Malthus
(1766–1834)

T R Malthus. An Essay on the Principle of Population. St. Paul's Church-yard 1789.

References

Irénée-Jules Bienaymé.
De la loi de multiplication et de la durée des familles.
Soc. Philomat. Paris Extraits, 13:131–132, 1845.

Richard Durrett.
Essentials of Stochastic Processes.
Springer, second edition, 2012.

Francis Galton.
Problem 4001.
Educational Times, page 17, 1873.

Thomas Robert Malthus.
An Essay on the Principle of Population.
St. Paul's Church-yard, 1798.

Henry William Watson and Francis Galton.
On the probability of the extinction of families.
The Journal of the Anthropological Institute of Great Britain and Ireland, 4:138–144, 1875.

Sources

Photos used in the presentation:

1. Image courtesy of think4photop at FreeDigitalPhotos.net