MVE550 2023 Lecture 7
Dobrow Chapter 4
Introduction to branching processes
Probability generating functions

Petter Mostad

Chalmers University

November 17, 2023

Introduction

- Many real phenomena can be described as developing with a tree-like structure, for example
  - Growth of cells.
  - Spread of viruses or other pathogens in a population.
  - Nuclear chain reactions.
  - Spread of funny cat videos on the internet.
  - Spread of a surname over generations.
- The process with which one node gives rise to "children" can be described as random: We will assume the probabilistic properties of this process is the same for all nodes.
- We will assume all nodes are organized into *generations*.
- We are only concerned with the size of each generation.
- How large are the generations? How much does the size vary? Will the process become *extinct*?

Branching processes

A branching process is discrete Markov chain $Z_0, Z_1, \ldots, Z_n, \ldots$ where

- the state space is the non-negative integers
- $Z_0 = 1$
- 0 is an absorbing state
- Z_n is the sum $X_1 + X_2 + \cdots + X_{Z_{n-1}}$, where the X_j are independent random non-negative integers all with the same *offspring distribution*. In other words

$$Z_n = \sum_{i=1}^{Z_{n-1}} X_i.$$

- Connecting each of the Z_n individuals in generation n with their offspring in generation $n+1$ we get a tree illustrating the branching process.
- The offspring distribution is described by the probability vector $a = (a_0, a_1, \ldots,)$ where $a_j = \Pr(X_i = j)$.
- To focus on the interesting cases we assume $a_0 > 0$ and $a_0 + a_1 < 1$.

Expected generation size

- Note that the state 0 is absorbing: This absorbtion is called *extinction*.
- As $a_0 > 0$, all nonzero states are transient.
- Define $\mu = \mathrm{E}(X_i) = \sum_{j=0}^{\infty} j a_j$ (the expected number of children).
- Then we may compute that

$$\mathrm{E}(Z_n) = \mathrm{E}\left(\sum_{i=1}^{Z_{n-1}} X_i\right) = \mathrm{E}\left(\mathrm{E}\left(\sum_{i=1}^{Z_{n-1}} X_i \mid Z_{n-1}\right)\right) = \mathrm{E}(Z_{n-1})\,\mu.$$

- We get directly that

$$\mathrm{E}(Z_n) = \mu^n\,\mathrm{E}(Z_0) = \mu^n$$

- We subdivide Branching processes into three types:
  - *Subcritical* if $\mu < 1$. Then $\lim_{n\to\infty} \mathrm{E}(Z_n) = 0$.
  - *Critical* if $\mu = 1$. Then $\lim_{n\to\infty} \mathrm{E}(Z_n) = 1$.
  - *Supercritical* if $\mu > 1$. Then $\lim_{n\to\infty} \mathrm{E}(Z_n) = \infty$.
- We can prove that if $\lim_{n\to\infty} \mathrm{E}(Z_n) = 0$ then the probability of extinction is 1.

Variance of the generation size

- Continue with $\mu = \mathrm{E}(X_i)$ denoting the expected number of children and let $\sigma^2 = \mathrm{Var}(X_i)$ denote the variance of the number of children.
- Using the law of total variance, we get

$$\begin{aligned}
\mathrm{Var}(Z_n) &= \mathrm{Var}(\mathrm{E}(Z_n \mid Z_{n-1})) + \mathrm{E}(\mathrm{Var}(Z_n \mid Z_{n-1})) \\
&= \mathrm{Var}\left(\mathrm{E}\left(\sum_{i=1}^{Z_{n-1}} X_i \mid Z_{n-1}\right)\right) + \mathrm{E}\left(\mathrm{Var}\left(\sum_{i=1}^{Z_{n-1}} X_i \mid Z_{n-1}\right)\right) \\
&= \mathrm{Var}(\mu Z_{n-1}) + \mathrm{E}\left(\sigma^2 Z_{n-1}\right) \\
&= \mu^2 \mathrm{Var}(Z_{n-1}) + \sigma^2 \mu^{n-1}
\end{aligned}$$

- From this we prove by induction, for $n \geq 1$,

$$\mathrm{Var}(Z_n) = \sigma^2 \mu^{n-1} \sum_{k=0}^{n-1} \mu^k = \begin{cases} n\sigma^2 & \text{if } \mu = 1 \\ \sigma^2 \mu^{n-1}(\mu^n - 1)/(\mu - 1) & \text{if } \mu \neq 1 \end{cases}$$

The probability of extinction

Let X denote an offspring variable and e the probability of extinction.

- Using first step analysis:

$$\begin{aligned} e &= \sum_{k=0}^{\infty} \Pr(X = k) \Pr(\text{extinction of process with } k \text{ roots}) \\ &= \sum_{k=0}^{\infty} \Pr(X = k)\, e^k = \mathrm{E}_X[e^X] = G_X(e) \end{aligned}$$

  where we define the *Probability generating function* $G_X(s)$ as

$$G_X(s) = \mathrm{E}_X[s^X].$$

- Let e_n be the probability of extinction by generation n. Using first step analysis:

$$\begin{aligned} e_n &= \sum_{k=0}^{\infty} \Pr(X = k) \Pr(\text{extinction by gen. } n-1, \text{ with } k \text{ roots}) \\ &= \sum_{k=0}^{\infty} \Pr(X = k)\, e_{n-1}^k = \mathrm{E}_X[e_{n-1}^X] = G_X(e_{n-1}) \end{aligned}$$

Extinction probability theorem

- If $s = x$ is any positive solution to $G_X(s) = s$ then $e \leq x$:
- *Proof*: We have $0 = e_0 < x$. We have that G_X is an increasing function, as

$$s_0 < s_1 \Rightarrow \sum_{k=0}^{\infty} \Pr(X = k)\, s_0^k < \sum_{k=0}^{\infty} \Pr(X = k)\, s_1^k \Rightarrow G_X(s_0) < G_X(s_1).$$

  Thus applying G_X repeatedly we get $e_n < x$ and thus $e = \lim_{n\to\infty} e_n \leq x$.
- We have proved the following THEOREM:
  Let G be the probability generating function for the offspring distribution for a branching process. The probability of eventual extinction is the smallest positive root of the equation $s = G(s)$.

Probability generating functions

- For *any* discrete random variable X taking values in $\{0, 1, 2, \ldots, \}$ define the probability generating function $G(s)$, or $G_X(s)$, as

$$G(s) = \mathrm{E}\left(s^X\right) = \sum_{k=0}^{\infty} s^k \Pr\left(X = k\right).$$

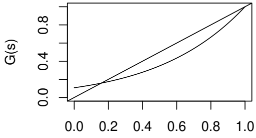
- The series converges absolutely for $|s| \leq 1$. We assume s is a real number in $[0, 1]$.
- We get a 1-1 correspondence between probability vectors on $\{0, 1, 2, \ldots, \}$ and functions represented by a series where the non-negative coefficients sum to 1.
- Specifically, if $G_X(s) = G_Y(s)$ for all s for random variables X and Y then X and Y have the same distribution.
- The correspondence of X with $G_X(s)$ provides an important and useful computational tool.

What does $G_X(s)$ look like?

- $G_X(1) = 1$ and $G_X(0) = \Pr(X = 0)$.
- We get

$$\begin{aligned}
G'(s) &= \sum_{k=1}^{\infty} k s^{k-1} \Pr(X = k) = \mathrm{E}\left(X s^{X-1}\right) \\
G''(s) &= \sum_{k=2}^{\infty} k(k-1) s^{k-2} \Pr(X = k) = \mathrm{E}\left(X(X-1) s^{X-2}\right) \\
G'''(s) &= \sum_{k=3}^{\infty} k(k-1)(k-2) s^{k-3} \Pr(X = k) = \mathrm{E}\left(X(X-1)(X-2) s^{X-3}\right)
\end{aligned}$$

- So the derivatives are non-negative, and $G'(s)$ and $G''(s)$ are positive for $s \in (0, 1)$.
- Below: $G_X(s)$ when $X \sim \text{Binomial}(10, 0.2)$. (Diagonal added)



Some properties of probability generating functions

- To go from X to $G_X(s)$: Compute the infinite (or finite) sum.
- To go from $G_X(s)$ to X: Use that we have

$$P(X = j) = \frac{G^{(j)}(0)}{j!}.$$

- If X and Y are independent,

$$G_{X+Y}(s) = \mathrm{E}\left(s^{X+Y}\right) = \mathrm{E}\left(s^X s^Y\right) = \mathrm{E}\left(s^X\right)\mathrm{E}\left(s^Y\right) = G_X(s)G_Y(s)$$

- $\mathrm{E}(X) = G'(1)$
- $\mathrm{E}(X(X-1)) = G''(1)$.
- As a consequence, $\mathrm{Var}(X) = G''(1) + G'(1) - G'(1)^2$.

Probability generating functions for Branching processes

Assume we have a Branching process $Z_0, Z_1, \ldots,$ with independent random variables X counting the offspring at each node.

- Write $G_n(s) = G_{Z_n}(s) = \mathrm{E}\left(s^{Z_n}\right)$ and $G(s) = G_X(s) = \mathrm{E}\left(s^X\right)$.
- We get

$$\begin{aligned} G_n(s) &= \mathrm{E}\left(s^{\sum_{k=1}^{Z_{n-1}} X_k}\right) = \mathrm{E}\left(\mathrm{E}\left(s^{\sum_{k=1}^{Z_{n-1}} X_k} \mid Z_{n-1}\right)\right) \\ &= \mathrm{E}\left(\mathrm{E}\left(\prod_{k=1}^{Z_{n-1}} s^{X_k} \mid Z_{n-1}\right)\right) = \mathrm{E}\left(G(s)^{Z_{n-1}}\right) = G_{n-1}(G(s)). \end{aligned}$$

- As $G_0(s) = \mathrm{E}\left(s^{Z_0}\right) = s$, it follows that $G_n(s) = G(G(G(\ldots G(s)\ldots)))$, with n iterations of the G function.
- This result can be applied (e.g., numerically) to compute $G_n(s)$.

Extinction probability theorem, with addition

- THEOREM
  Let G be the probability generating function for the offspring distribution for a branching process. The probability of eventual extinction is the smallest positive root of the equation $s = G(s)$.
  Also, if the process is critical ($\mu = 1$) then the extinticion probability is 1.
- *Proof of last sentence*: In the critical case,

  $$G'(1) = \mathrm{E}(X) = \mu = 1.$$

  As $G''(s) > 0$ for $s \in (0, 1)$, we get that $G'(s) < 1$ for $s \in (0, 1)$, and $\frac{d}{ds}(G(s) - s) < 0$ for $s \in (0, 1)$.
  As $G(1) - 1 = 0$ for any probability generating function, we get that $G(s) = s$ has its smallest positive root at 1.