Population growth: Galton-Watson process

February 14, 2018

Population growth: Galton-Watson process

A simple branching process

Ultimate extinction

Probability law on the n-th generation

A simple branching process

Let us consider a simple stochastic model for the evolution of the size of a population.

(a) The population evolves in generations. Let Z_k be the number of members of the k-th generation. By assumption, $Z_0 = 1$.

(b) Each member of the k-th generation gives birth to a family, possibly empty, of members of the $(k+1)$-th generation.

- The family sizes form a collection of independent and identically distributed random variables.
- The size X of each of these families (i.e., the number of descendants of an individual) has probability function:

$$\mathrm{P}(X = i) = p_i, \quad i \geq 0$$

A simple branching process

Such a branching process

$$\{Z_0, Z_1, \dots, Z_n, \dots\}$$

is called a Galton-Watson process.

Example

For instance, suppose that

$$p_0 = \frac{1}{2}, \qquad p_1 = \frac{1}{4}, \qquad p_2 = \frac{1}{4}$$

Then

$$\mathrm{P}(Z_1 = 0) = \frac{1}{2}, \qquad \mathrm{P}(Z_1 = 1) = \frac{1}{4}, \qquad \mathrm{P}(Z_1 = 2) = \frac{1}{4}$$

Example

Suppose, for instance, that $Z_1 = 2$. Then

$$\mathrm{P}(Z_2 = 0 \mid Z_1 = 2) = \mathrm{P}(X_1 + X_2 = 0) = p_0^2 = \frac{1}{4}$$
$$\mathrm{P}(Z_2 = 1 \mid Z_1 = 2) = \mathrm{P}(X_1 + X_2 = 1) = p_0 p_1 + p_1 p_0 = \frac{1}{4}$$
$$\mathrm{P}(Z_2 = 2 \mid Z_1 = 2) = \mathrm{P}(X_1 + X_2 = 2) = p_0 p_2 + p_1 p_1 + p_2 p_0 = \frac{5}{16}$$
$$\mathrm{P}(Z_2 = 3 \mid Z_1 = 2) = \mathrm{P}(X_1 + X_2 = 3) = p_1 p_2 + p_2 p_1 = \frac{1}{8}$$
$$\mathrm{P}(Z_2 = 4 \mid Z_1 = 2) = \mathrm{P}(X_1 + X_2 = 4) = p_2^2 = \frac{1}{16}$$

Of course, we have

$$\sum_{n \geq 0} \mathrm{P}(Z_2 = n \mid Z_1 = 2) = 1$$

Example

Performing similar calculations and taking into account that

$$\mathrm{P}\left(Z_2 = n\right) = \sum_{r=0}^{2} \mathrm{P}(Z_2 = n \mid Z_1 = r)\mathrm{P}(Z_1 = r)$$

we obtain the probability function of Z_2:

$$\mathrm{P}(Z_2 = 0) = \frac{11}{16}, \quad \mathrm{P}(Z_2 = 1) = \frac{2}{16}, \quad \mathrm{P}(Z_2 = 2) = \frac{9}{64},$$
$$\mathrm{P}(Z_2 = 3) = \frac{1}{32}, \quad \mathrm{P}(Z_2 = 4) = \frac{1}{64}$$

Probability distribution of Z_{k+1}

In general,

$$\mathrm{P}(Z_{k+1} = n \mid Z_k = r) = \mathrm{P}(X_1 + X_2 + \cdots + X_r = n),$$

where the r.v. X_i are independent and identically distributed as X.

Of course,

$$\mathrm{P}(Z_{k+1} = 0 \mid Z_k = 0) = 1$$

By the TPT,

$$\mathrm{P}\left(Z_{k+1} = n\right) = \sum_{r \geq 0} \mathrm{P}(Z_{k+1} = n \mid Z_k = r)\mathrm{P}(Z_k = r)$$

Extinction probabilities

Let us calculate the probability that the population disappears in a number of generations $\leq m$,

$$d_m \equiv \mathrm{P}(Z_m = 0)$$

In the previous example,

$$d_0 = 0$$

$$d_1 = \frac{1}{2}$$

$$d_2 = \frac{11}{16}$$

$$\dots$$

Extinction probabilities

Let $D_m = \{Z_m = 0\}$. Then

$$\begin{aligned} d_m &= \mathrm{P}(D_m) \\ &= \sum_{k\geq 0} \mathrm{P}(D_m \mid Z_1 = k)\,\mathrm{P}(Z_1 = k) = \sum_{k\geq 0} (d_{m-1})^k\, \mathrm{P}(X = k) \end{aligned}$$

Hence, if

$$G_X(s) = \sum_{k\geq 0} s^k\, \mathrm{P}(X = k)$$

is the probability generating function of X, then

$$d_m = G_X(d_{m-1})$$

Extinction probabilities

In our example,

$$G_X(s) = \frac{1}{2} + \frac{1}{4}s + \frac{1}{4}s^2$$

Hence

$$d_0 = 0$$

$$d_1 = G_X(d_0) = G_X(0) = \frac{1}{2}$$

$$d_2 = G_X(d_1) = G_X\left(\frac{1}{2}\right) = \frac{1}{2} + \frac{1}{4}\cdot\frac{1}{2} + \frac{1}{4}\cdot\left(\frac{1}{2}\right)^2 = \frac{11}{16}$$

$$d_3 = G_X(d_3) = G_X\left(\frac{11}{16}\right) = \frac{1}{2} + \frac{1}{4}\cdot\frac{11}{16} + \frac{1}{4}\cdot\left(\frac{11}{16}\right)^2 = \frac{809}{1024}$$

$$\dots\dots$$

Ultimate extinction

Notice that

$$d_0 \leq d_1 \leq d_2 \leq \dots \leq d_m \leq \dots \leq 1$$

Hence the following limit exists:

$$d \equiv \lim_{m\to\infty} d_m$$

What is its meaning?

Ultimate extinction

We have $D_0 \subset D_1 \subset \cdots \subset D_m \subset D_{m+1} \subset \cdots$

Therefore we can consider the limit event

$$D = \lim_{m \to \infty} D_m = \bigcup_{k=0}^{\infty} D_k,$$

which corresponds to the ultimate extinction of the population.

Hence

$$d = \lim_{m \to \infty} d_m = \lim_{m \to \infty} \mathrm{P}(D_m) = \mathrm{P}\left(\lim_{m \to \infty} D_m\right) = \mathrm{P}(D)$$

is the probability of that ultimate extinction.

Ultimate extinction

From $d_m = G_X(d_{m-1})$ we get that $(m \to \infty)$:

$$d = G_X(d)$$

- So, the extinction probability d is a solution of the equation

$$s = G_X(s)$$

- Notice that $s = 1$ is always a root.

Ultimate extinction

Returning to the example,

$$s = G_X(s) \iff s = p_0 + p_1\, s + p_2\, s^2$$

Hence

$$p_2\, s^2 - (p_0 + p_2)\, s + p_0 = 0$$

The two roots are

$$s_1 = 1, \quad s_2 = \frac{p_0}{p_2}$$

Let m be the expected number of descendants per individual:

$$m = \mathrm{E}(X) = p_1 + 2p_2 = 1 - p_0 + p_2$$

Ultimate extinction

(1) Case $m < 1 \iff p_0 > p_2 \iff s_2 > 1$

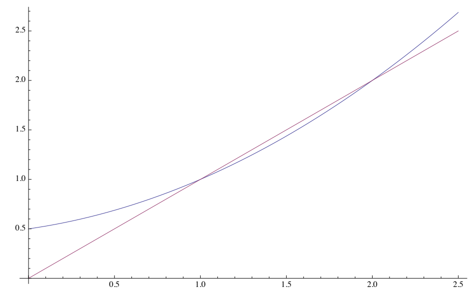

Figure: $p_0 = 1/2, \quad p_1 = 1/4, \quad p_2 = 1/4$

The iterations $d_m = G_X(d_{m-1})$ converge to $d = 1$.

Ultimate extinction

(2) Case $m = 1 \iff p_0 = p_2 \iff s_1 = s_2 = 1$

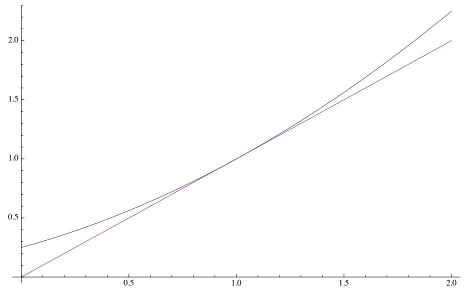

Figure: $p_0 = 1/4, \quad p_1 = 1/2, \quad p_2 = 1/4$

The iterations $d_m = G_X(d_{m-1})$ also converge to $d = 1$.

Ultimate extinction

(3) Case $m > 1 \iff p_0 < p_2 \iff s_2 < 1$

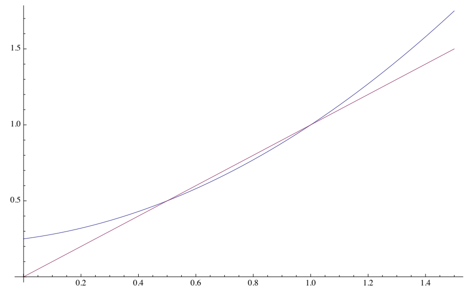

Figure: $p_0 = 1/4, \quad p_1 = 1/4, \quad p_2 = 1/2$

The iterations $d_m = G_X(d_{m-1})$ converge to $d = s_2 < 1$.

General case

The same analysis can be done in a more general situation.

Notice that if $s \geq 0$, then

$$G_X(s) = p_0 + p_1 s + p_2 s^2 + p_3 s^3 + \cdots \geq 0$$

$$G'_X(s) = p_1 + 2p_2 s + 3p_3 s^2 + \cdots \geq 0$$

$$G''_X(s) = 2p_2 + 6p_3 s + \cdots \geq 0$$

So, the plot of $G_X(s)$ is as in the case of a polynomial of degree 2.

General case

- In general, the plot of $G_X(s)$ will intersect in two points the line (with equation) s. So, the equation $s = G_X(s)$ will have two solutions.
- The iterations $d_m = G_X(d_{m-1})$ will converge to the smallest of the two solutions.
- Moreover,
$$m = \mathrm{E}(X) = G'_X(1)$$
is the slope of the tangent line of $G_X(s)$ at $s = 1$.
- Hence we have again the three cases considered above for G_X a polynomial of degree 2.

Ultimate extinction

Theorem

Let d be the smallest root of the equation $s = G_X(s)$ and let $m = \mathrm{E}(X)$. Then d is the probability of ultimate extinction of the population. Moreover,

- $m < 1 \Rightarrow d = 1$. *The extinction is sure.*
- $m = 1 \Rightarrow$ *The line s is tangent at $s = 1$ to $G_X(s)$ (double root $d = 1$). The extinction is sure.*
- $m > 1 \Rightarrow d < 1$. *In this case, there is a positive probability $1 - d > 0$ of non-extinction.*

Probability generating function of Z_n

Let $G_n(s)$ the probability generating function of Z_n.

We have

$$G_{n+1}(s) = \mathrm{E}\left(s^{Z_{n+1}}\right) = \sum_{k\geq 0} \mathrm{E}\left(s^{Z_{n+1}} \mid Z_n = k\right) \mathrm{P}(Z_n = k)$$

Moreover,

$$\mathrm{E}\left(s^{Z_{n+1}} \mid Z_n = k\right) = \mathrm{E}\left(s^{X_1+X_2+\cdots+X_k} \mid Z_n = k\right)$$

$$= \mathrm{E}\left(s^{X_1+X_2+\cdots+X_k}\right) = \left(\mathrm{E}\left(s^X\right)\right)^k = (G_X(s))^k$$

Probability generating function of Z_n

Therefore

$$G_{n+1}(s) = \sum_{k\geq 0} (G_X(s))^k \, \mathrm{P}(Z_n = k)$$

Recall that

$$G_n(z) = \sum_{k\geq 0} z^k \, \mathrm{P}(Z_n = k)$$

Then

$$G_{n+1}(s) = G_n\left(G_X(s)\right)$$

Probability generating function of Z_n

So, we have

$$G_1(s) = G_X(s)$$

$$G_2(s) = G_1(G_X(s)) = G_X(G_X(s)) = G_X^{(2)}(s)$$

$$G_3(s) = G_2(G_X(s)) = G_X(G_X(G_X(s))) = G_X^{(3)}(s)$$

....

$$G_n(s) = G_{n-1}(G_X(s)) = G_X(G_X(\cdots G_X(s)\cdots)) = G_X^{(n)}(s)$$

....

Probability generating function of Z_n

In the example,

$$G_X(s) = \frac{1}{2} + \frac{1}{4}s + \frac{1}{4}s^2$$

Then

$$G_2(s) = G_X(G_X(s)) = \frac{1}{2} + \frac{1}{4}G_X(s) + \frac{1}{4}\left(G_X(s)\right)^2$$

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

$$= \frac{11}{16} + \frac{1}{8}s + \frac{9}{64}s^2 + \frac{1}{32}s^3 + \frac{1}{64}s^4$$

Expected number of individuals

Let $m_n = \mathrm{E}(Z_n)$.

The derivative of $G_{n+1}(s) = G_n\left(G_X(s)\right)$ is

$$G'_{n+1}(s) = G'_n\left(G_X(s)\right)\, G'_X(s)$$

Taking $s = 1$ we have

$$G'_{n+1}(1) = G'_n\left(G_X(1)\right)\, G'_X(1) = G'_n(1)\, G'_X(1)$$

Hence $m_{n+1} = m_n\, m$. Therefore

$$m_n = m^n$$