A Short Introduction to Branching Processes Part I

Gesine Reinert

Department of Statistics
University of Oxford
reinert@stats.ox.ac.uk

Michaelmas Term 2009
Outline

Some motivating examples

The Galton-Watson Branching Process

Limiting behaviour of the Galton-Watson process

The distribution of siblings

The total progeny of a branching process

Multi-type Galton-Watson processes

References
Examples which started the subject

- **Malthus (1817):** A population, when unchecked, grows exponentially.

- **Benoiston de Chateauneuf (1844, 1845), Bienaymé (1845):** Estimated the duration of noble families that were founded in the tenth to twelfth centuries to be 300 years; and Bienaymé derived the probability of extinction.

- **Galton (1869):** studied the decay of the English peerage and other families of ’men of note’; are distinguished families more likely to die out than ordinary ones? He came to the conclusion that one factor in lowering the rates was the tendency of peers to marry heiresses. An heiress, coming from a family with no sons, would be expected to have, by inheritance, a lower-than-ordinary fertility.
Examples which illustrate applications

- *Fisher* (1922, 1930): survival of genes
- Electron multiplies: An electron multiplier is a device for amplifying a weak current of electrons. Each electron, as it strikes the first in a series of plates, gives rise to a random number of electrons, which strike the next plate and produce more electrons, etc.
Frequently asked questions

- the probability of extinction
- When the process does not die out, what is its limiting behaviour?
- the distribution of siblings of a randomly chosen child
- the total progeny of a branching process
The model assumptions

Individuals or objects live for a single time period; at the end of the period, each individual produces a random number of offspring. All offspring are assumed to be equivalent.

If $Z_n$ is the size of the $n^{th}$ generation, for $n = 0, 1, \ldots$, then $Z_0, Z_1, \ldots$ form a Markov chain: if $Z_n$ is known, then $Z_{n+1}$ does not depend on $Z_0, \ldots, Z_{n-1}$.

The number of children born to an individual does not depend on how many other individuals are present.
Mathematical formulation

Unless otherwise stated, we assume that we start with a single individual, $Z(0) = 1$. The offspring distribution is described by a random variable $\xi$ having discrete non-negative probability distribution $(p_0, p_1, \ldots)$; here $p_k = P(\xi = k)$. Each individual has a random number of children in the next generation. These random variables are independent copies of $\xi$. Thus we have the branching process equation

$$Z_{n+1} = \sum_{i=1}^{Z_n} \xi_i,$$

where $\xi_1, \xi_2, \ldots$ are i.i.d. copies of $\xi$. 

Some motivating examples
The Galton-Watson Branching Process
Limiting behaviour of the Galton-Watson process
The distribution of siblings
The total progeny of a branching process
Multi-type Galton-Watson processes
References
Mean

Let $E\xi = \mu < \infty$ and $0 < \text{Var}(\xi) = \sigma^2 < \infty$. Let $M_n = EZ_n$. Then

$$M_{n+1} = E\left( \sum_{i=1}^{Z_n} \xi_i \right)$$

$$= \sum_k P(Z_n = k) E\left( \sum_{i=1}^{k} \xi_i \right)$$

$$= \sum_k P(Z_n = k) k \mu$$

$$= \mu M_n.$$

The solution of this recursion with $M_0 = 1$ is $M_n = \mu^n$. 
Let $V_n = \text{Var}(Z_n)$. Then we can derive that

$$V_{n+1} = M_n \sigma^2 + V_n \mu^2.$$ 

The solution for this recursion with $V_0 = 0$ is

$$V_n = \begin{cases} \sigma^2 n & \text{if } \mu = 1; \\ \sigma^2 \mu^{n-1} \frac{1-\mu^n}{1-\mu} & \text{if } \mu \neq 1. \end{cases}$$
So, if $\mu = 1$, the population mean is unchanged, but the variance increases linearly.

If $\mu > 1$, the population mean and variance grow geometrically.

If $\mu < 1$, the population mean and variance shrink geometrically.
The probability generating function

The probability generating function (p.g.f) of $\xi$ is

$$f(s) = \sum_{k=0}^{\infty} p_k s^k,$$

where $s$ is a complex variable.

The iterates of $f(s)$ are defined by

$$f_0(s) = s, \quad f_1(s) = f(s),$$

and for $n = 1, 2, \ldots,$

$$f_{n+1}(s) = f(f_n(s)).$$
Some facts about p.g.f.’s

The p.g.f. is strictly convex and increasing in $[0, 1]$. $f(0) = p_0; f(1) = 1.$
If $\mu \leq 1$ then $f(t) > t$ for $t \in [0, 1)$; if $\mu > 1$ then $f(t) = t$ has a unique root in $[0, 1)$.

Fact: The p.g.f. of $Z_n$ is $f_n(s)$.

We could have used the p.g.f. to derive the mean and the variance of $Z_n$. 
If $f(s) = p_0 + p_1 s$ for some $0 < p_0 < 1$, $p_0 + p_1 = 1$, then in the associated branching process, in each period each individual dies independently, with probability $p_0$, and survives with probability $p_1 = 1 - p_0$. The process is a pure death process.

If $f(s) = p_0 + p_2 s^2$ for some $0 < p_0 < 1$, $p_0 + p_2 = 1$, then in the associated branching process, in each period each individual dies independently, with probability $p_0$, or is replaced by two progeny, with probability $1 - p_0$. 
The probability of extinction

By *extinction* we mean that the sequence \( \{ Z_n \} \) consists of zeros for all but a finite number of \( n \).
It is also the event that \( Z_n \to 0 \). Put \( q = P(Z_n \to 0) \), then

\[
q = P(Z_n = 0 \text{ for some } n)
= \lim_{n \to \infty} P(Z_n = 0)
= \lim_{n \to \infty} f_n(0).
\]

By conditioning on the first generation we can also show that

\[
q = \sum_{n} P(\xi = n)q^n = f(q).
\]
Consequences for the extinction probability

If $\mu > 1$ then $q = p^*$, where $p^*$ is the unique solution in $[0, 1)$ of $f(t) = t$. This is called the supercritical case. Actually one can show that if $\mu = 1 + \epsilon$, then the survival probability is

$$1 - q \approx 2\frac{\epsilon}{\sigma^2}.$$ 

If $\mu \leq 1$ then $q = 1$.

Note: If $\mu = 1$ then $q = 1$ but $E(\sum_{n=0}^{\infty} Z_n) = \sum_{n=0}^{\infty} 1 = \infty$. This case is called the critical case.

If $\mu < 1$ then $q = 1$ also; this is called the subcritical case.
Results for the supercritical case

In this case, both extinction and survival are possible with positive probability.

Note: conditional on extinction,

\[ P(Z_1 = i \mid \text{extinction}) = \frac{P(Z_1 = i; \text{extinction})}{P(\text{extinction})} \]

\[ = \frac{p_i q^i}{p^*} \]

\[ = : P(\hat{\xi} = i). \]

Then \( E\hat{\xi} < 1 \), and moreover, it holds that the distribution of the Galton-Watson process, conditioned on extinction, is the same as the distribution of the Galton-Watson process with offspring variable \( \hat{\xi} \).
Almost-sure convergence

Theorem

There is a positive random variable $W$ such that

$$\mu^{-n} Z_n \to W \quad \text{almost surely as } n \to \infty.$$ 

That is, either the process becomes extinct, or it grows as $W \mu^n$ for a random variable $W > 0$. The proof relies on the martingale limit theorem.

The distribution of $W$ is in general not easy to derive. It can be shown that its moment-generating function $\phi(s) = E e^{-sW}$ satisfies

$$\phi(\mu s) = f(\phi(s)), \quad \Re(s) \geq 0;$$

and $\phi'(0) = 1$. 
The Kesten-Stigum Theorem

If $E(\xi \log^+ \xi) < \infty$, then

$$\mu^{-n}EZ_n \to E(W) = 1,$$

and

$$P(W > 0) = 1 - q.$$

$W$ is a function of the first generation, and is known as the Founder effect.
A normal approximation

Conditional on survival, in distribution,

$$\mu_n^2 (\frac{Z_n - W}{\sqrt{W}}) \rightarrow N \left(0, \frac{\sigma^2}{\mu^2 - \mu} \right) \sqrt{W},$$

where the normal and $W$ are independent in the limit.
Yaglom’s Theorem for the subcritical case

Since $Z_n \to 0$ for $n \to \infty$ when $m < 1$, the limiting distribution of $Z_n$ is not interesting - but the distribution conditional on $Z_n \neq 0$ is.

Yaglom’s Theorem. For each $j = 1, 2, \ldots$,

$$\lim_{n \to \infty} P(Z_n = j | Z_n \neq 0) = b_j$$

exists, and $\sum_{j=1}^{\infty} b_j = 1$. Moreover, $g(s) = \sum_k b_k s^k$ satisfies the equation

$$g(f(s)) = \mu g(s) + 1 - \mu.$$
The critical case: $\mu = 1$

In this case one can show that

$$P(Z_n > 0) \approx \frac{2}{n\sigma^2}.$$  

Moreover the conditional mean is such that

$$E(Z_n|Z_n > 0) \approx \frac{n\sigma^2}{2},$$

growing at the rate of $n$.  

Yaglom’s Theorem. For any $z \geq 0$,

$$\lim_{n \to \infty} P \left( \frac{Z_n}{n} > z | Z_n \neq 0 \right) = \exp \left\{ -2 \frac{z}{\sigma^2} \right\}.$$
The distribution of siblings

Let $X$ be the number of children in a randomly picked family, (thus $X$ has the same distribution as $Z_1$), and let $Y$ be the number of siblings in the family of a randomly picked child. With Bayes’ formula,

$$P(Y = i) = P(\text{pick child from family with } i+1 \text{ children})$$

$$= P(\text{pick child}|X = i+1)P(X = i+1)$$

$$= \frac{\sum_{k=0}^{\infty} P(\text{pick child}|X = k+1)P(X = k+1)}{(i+1)P(X = i+1)}$$

$$= \frac{\sum_{k=0}^{\infty} (k+1)P(X = k+1)}{\mu^{-1}(i+1)P(X = i+1)}.$$

Thus $X$ and $Y$ have the same distribution if and only if $X$ has a Poisson distribution.
The total progeny is \( Y_\infty = \sum_{n=0}^{\infty} Z_n \).

From the extinction probability, \( P(Y_\infty < \infty) = q \).

Putting \( Y_n = \sum_{k=0}^{n} Z_k \) we can calculate that

\[
EY_n = \frac{1 - \mu^{n+1}}{1 - \mu} \quad \text{if } \mu \neq 1,
\]

and if \( \mu = 1 \), then

\[
EY_n = n + 1.
\]

If \( p_{jk} = P(Z_{n+1} = k | Z_n = j) \), then it can be shown that for any \( j \leq k \)

\[
P[Y_\infty = k] = \frac{j}{k} p_{k,k-j}.
\]
The proof uses the *Ballot Theorem*: Let $X_1, X_2, \ldots$ be i.i.d taking values in the non-negative integers, and let $S_k = \sum_{i=1}^k X_i$. If $EX_1 < 1$, then

$$P(S_k \leq k \text{ for } 1 \leq k \leq n | S_n = s_n) = \left(1 - \frac{s_n}{n}\right)^+.$$  

It follows that

$$P(S_k < k \text{ for all } k) = 1 - EX_1.$$
Multi-type Galton-Watson processes

Now suppose that we can distinguish $r$ types of individuals, type $1, 2, \ldots, r$. Each individual has a unit life-length, and if it is of type $k$, then it splits into $k_1$ children of type 1, $k_2$ children of type 2, ..., up to $k_r$ children of type $r$, with probability $p_k(k_1, \ldots, k_r)$. 
Examples

- Normal and mutants in mitochondrial DNA. A normal can give birth to either two normals, or one normal and one mutant. Mutants can only give birth to mutants.

- Polymerase chain reaction (PCR): a given particle (single DNA strand) at any time either existed before the last PCR cycle or is newly created. The type space is “new” and “old”. The offspring distribution depends on the probability that a given molecule replicates successfully in a given PCR cycle.
The mean matrix

Let $Z_{1,j}^{(i)}$ be the number of type $j$ offspring of a single type $i$ particle in one generation. We assume that

$$0 < m_{i,j} = EZ_{1,j}^{(i)} < \infty$$

and we define the mean matrix

$$M = (m_{i,j})_{i,j=1,...,r}.$$ 

Let $Z_n = (Z_{n,1}, \ldots, Z_{n,r})$, or, when the process starts with one particle in state $i$, we denote it by $Z_n^{(i)}$. Then

$$E[Z_n|Z_0] = Z_0 M^n.$$
Higher moments

Put

\[ q_n^{(\alpha)}(i, j) = E\{Z_n^{(\alpha)} Z_n^{(\alpha)} - \delta_{ij} Z_n^{(\alpha)}\} = \frac{\partial^2 f^{(\alpha)}}{\partial s_i \partial s_j} (1) \]

and define the matrix

\[ Q_n^{(\alpha)} = \{ q_n^{(\alpha)}(i, j), i, j = 1, \ldots, r \}, \]

the quadratic form

\[ Q_n^{(\alpha)}[s] = \frac{1}{2} \sum_{i,j=1}^{r} s_i q_n^{(\alpha)}(i, j) s_j, \]

and the vectors of quadratic forms \( Q_n[s] = (Q_n^{(\alpha)}[s], \alpha = 1, \ldots, r); \) we put \( Q[s] = Q_1[s]. \)
The Frobenius Theorem

A strictly positive matrix $M$ has a maximal eigenvalue $\rho$ which is positive, simple, and has associated right and left eigenvectors $u$ and $v$. If these are normalised such that $u \cdot v = \sum_{i=1}^{r} u_i v_i = 1$ and $u \cdot 1 = \sum_{i=1}^{r} u_i = 1$, then one can write

$$M^n = \rho^n P + R^n,$$

where $P$ is the matrix whose $(i, j)$ entry is $u_i v_j m$ and where

$$PR = RP = 0$$

and $r_{i,j}^{(n)} \leq c \rho_0^n$, for $i, j = 1, \ldots, r$, for some $c < \infty$, and $0 < \rho_0 < \rho$. 
Extinction probabilities

We assume that the process is *nonsingular*, that is, it is not the case that each particle has exactly one offspring. Let \( q^{(i)} \) be the probability of eventual extinction of the process initiated with a single particle of type \( i \), and put \( q = (q^{(1)}, \ldots, q^{(r)}) \). Let \( \rho \) be the maximal eigenvalue of the strictly positive mean matrix \( M \). Let \( f_n \) denote the generating function of \( Z_n \):

\[
  f^{(i)}(s) = \sum_{j_1, \ldots, j_r} p^{(i)}(j_1, \ldots, j_r) s_1^{j_1} \cdots s_r^{j_r}
\]

and \( f(s) = (f^{(1)}(s), \ldots, f^{(r)}(s)) \).
Theorem

1. If $\rho \leq 1$ then $q = 1$. If $\rho > 1$ then $q < 1$.
2. $\lim_{n \to \infty} f_n(s) = q$
3. The only solution of $f(s) = s$ with $s$ in the unit cube is $q$.

If $\rho > 1$ the process is called supercritical, if $\rho = 1$ it is called critical, and if $\rho < 1$ it is called subcritical.
Results for the supercritical case

**Theorem**

*There is a nonnegative random variable $W$ such that*

$$\lim_{{n \to \infty}} \rho^{-n}Z_n = vW \quad \text{almost surely}.$$  

*Furthermore $P(W > 0) > 0$ if and only if*

$$E\{Z_{1j}^{(i)} \log Z_{1j}^{(i)}\} < \infty \quad \text{for all } i, j = 1, \ldots, r.$$
Results for the subcritical case

Again we condition.

Theorem

$$\lim_{n \to \infty} P(Z_n = j | Z_0 = i, Z_n \neq 0) = b(j)$$

exists, is independent of $i$, and is a probability measure.
Results for the critical case

Here $\| \cdot \|$ denotes the Euclidean norm.

**Theorem**

If $E\|Z_1\|^2 < \infty$ and if $w \cdot v > 0$, then $\frac{1}{n} Z_n \cdot w$, conditioned on $Z_n \neq 0$, converges in distribution to the random variable with density

$$f(x) = \frac{1}{\gamma_1} e^{-\frac{x}{\gamma_1}}, \quad x \geq 0,$$

where

$$\gamma_1 = \frac{v \cdot w}{v \cdot Q[u]}.$$
If the vector $w$ is orthogonal to $v$, then the limit law degenerates, $\gamma_1 = 0$. Instead,

**Theorem**

If $E\|Z_1\|^2 < \infty$ and if $w \cdot v = 0$, then $\frac{1}{\sqrt{n}} Z_n \cdot w$, conditioned on $Z_n \neq 0$, converges in distribution to the random variable with density

$$f(x) = \frac{1}{\gamma_2} e^{-\frac{|x|}{\gamma_2}}, \quad -\infty < x < \infty,$$

for some $\gamma_2 > 0$. 
Extensions of the Galton-Watson process:

1. Bellman-Harris process: independent individuals lead a life of random length, then give birth to a random number of children, independently of the mother’s life span.
2. Sevast’yanov: allow for dependence between life span and reproduction.
3. Crump-Mode-Jagers: individuals have random life spans during which births occur as a point process, time is continuous, and any type of dependence between reproduction and life is allowed.
5. Branching processes with immigration
6. Branching processes in random environments
Outlook: Part 2

Continuous time
The general branching process
Branching Process Approximations
Some motivating examples
The Galton-Watson Branching Process
Limiting behaviour of the Galton-Watson process
The distribution of siblings
The total progeny of a branching process
Multi-type Galton-Watson processes

References

References continued