Review

Not peer-reviewed version

The Branching Process: A General Conceptual Framework for Addressing Current Ecological and Evolutionary Questions

Xuhua Xia *

Posted Date: 17 November 2025

doi: 10.20944/preprints202511.1105.v1

Keywords: Galton–Watson process; offspring distribution; extinction probability; extinction time; fixation time; random process; evolution; population growth

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 November 2025 doi:10.20944/preprints202511.1105.v1

Review

The Branching Process: A General Conceptual Framework for Addressing Current Ecological and Evolutionary Questions

Xuhua Xia [1,2]

[1] Department of Biology, University of Ottawa, Marie-Curie Private, Ottawa, ON K1N 9A7, Canada; xxia@uottawa.ca; Tel.: +1-613-204-2347

[2] Ottawa Institute of Systems Biology, University of Ottawa, Ottawa, ON K1H 8M5, Canada

Abstract

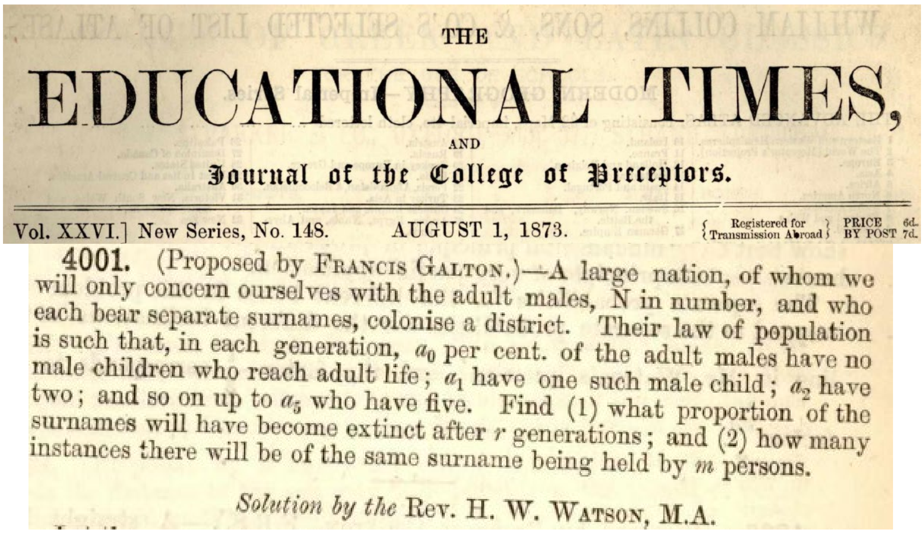
Classical branching-process theory, developed by Galton and Watson in the nineteenth century and later refined by Fisher and Haldane, provides the formal framework for quantifying the fate of new mutants, new viral and bacterial pathogens, new colonization of invasive species, etc. It is a powerful tool to quantify and predict the effect of differential reproductive success on the speciation potential of evolutionary lineages. Here, I revisit the conceptual framework of the branching process, detail its mathematical development over time, tie up a few historical loose strings, and highlight its potential applications in modern ecology and evolutionary biology.

Keywords: Galton–Watson process; offspring distribution; extinction probability; extinction time; fixation time; random process; evolution; population growth

---

1. Introduction

Branching process is not only "the beautiful theory" in the field of probability [1], but also a key methods in evolutionary studies since its formulation [2] (Figure 1). Galton asked two questions. Suppose we start with *N* adult males, each with a distinct surname. In each generation, each man produces *k* adult sons with the probability a_k, with $\sum a_k = 1$. All adults produce sons independently. What is the chance of a male whose descendants will perish after *t* generations? This chance is the same as the expected proportion of surnames surviving *t* generations. The second question is tricker. Galton wanted to have a distribution. That is, after *t* generations, how many surnames will have *m* (= 0, 1, 2, …,) bearers?



THE

EDUCATIONAL TIMES,

AND

Journal of the College of Preceptors.

Vol. XXVI.] New Series, No. 148. AUGUST 1, 1873. {Registered for Transmission Abroad} PRICE 6d. BY POST 7d.

4001. (Proposed by FRANCIS GALTON.)—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, a_0 per cent. of the adult males have no male children who reach adult life; a_1 have one such male child; a_2 have two; and so on up to a_5 who have five. Find (1) what proportion of the surnames will have become extinct after r generations; and (2) how many instances there will be of the same surname being held by m persons.

Solution by the Rev. H. W. WATSON, M.A.

Figure 1. Francis Galton's question and Henry Watson's solution.

The branching process, which was used by Watson to provide a partial solution to Galton's questions, has since found applications in ecology and evolution, through the effort of Fisher [3], Haldane [4] and many others. The basic input to the branching processes is a probability mass function specifying the reproductive potential of individuals, and the output consists of extinction probability Q, extinction time T and the variance of T [5,6]. This paper focuses on the details between the input and the output, and highlights its relevance to current research problems.

2. A Branching Process with Empirical Reproductive Success

Suppose we follow the queen and the number of her daughter queens of an invasive giant Asian hornet, *Vespa mandarinia,* that almost succeeded in invading North America [7] or yellow-legged Asian hornet, *V. velutina* that successfully invaded France and other European countries [8]. This example is also relevant to the release of animals to rescue a declining population [9] or to create a game population for hunting. Will the individuals in the new environment go extinct? How long will it take to go extinct?

Of 100 empirically monitored hornet queens, 50 did not produce any reproducing daughter queens, 30 produced one reproducing queen and 20 produced two reproducing queens. Thus, a queen produces 0, 1, and 2 daughter queens with the corresponding probabilities 0.5, 0.3 and 0.2, respectively. The probability mass function (PMF), which is often called reproduction law in branching processes, is then:

$$\begin{aligned} f(x=0) &= 0.5 \\ f(x=1) &= 0.3 \\ f(x=2) &= 0.2 \end{aligned} \tag{1}$$

We assume that this PMF does not change with time in a branching process and is applicable to every replicating queen. However, one could have branching processes to accommodate queens with different reproducing potentials.

A branching process and its analysis are anchored on a probability generating function (PGF) derived from the PFM. PGF is defined as $G_X(s) = \sum_{k=0}^{Max\,k} f(k)s^k$, where k is the value that X can take, and the subscript X can often be omitted without confusion. Given the distribution specified in Equation (**1**), the PGF is therefore

$$G(s) = 0.5s^0 + 0.3\,s^1 + 0.2\,s^2 = 0.5 + 0.3s + 0.2s^2 \tag{2}$$

Note that $G_X(1) = 1$. The mean and variance of X based on the PGF are

$$E(X) = G'_X(1) = \mu \tag{3}$$

$$Var(X) = G''_X(1) + G'_X(1) - [G'_X(1)]^2 \tag{4}$$

For the PGF in Equation (**2**), we have

$$G'_X(s) = 0.3 + 0.4s;\ G''_X(s) = 0.4 \tag{5}$$

$$E(X) = G'_X(1) = 0.3 + 0.4 = 0.7 \tag{6}$$

$$Var(X) = G''_X(1) + G'_X(1) - [G'_X(1)]^2 = 0.4 + 0.7 - 0.7^2 = 0.61 \tag{7}$$

In ecological and evolutionary studies, a stable wild-type population with $\mu = E(X) = 1$ is often implicitly assumed.

2.1. Extinction Probability Q

What is the probability that an individual who produces offspring according to Equation (**1**) will leave no descendants in the future? The extinction probability Q was derived [2] as the smallest non-negative solution to the following fixed-point equation:

$$\boldsymbol{Q = G(Q)} \tag{8}$$

From Equation (**2**), we have

$$\boldsymbol{Q = G(Q) = 0.5 + 0.3\,Q^1 + 0.2\,Q^2} \tag{9}$$

$$\boldsymbol{0.2Q^2 - 0.7Q + 0.5 = 0} \tag{10}$$

This quadratic equation has two roots, $r_1 = 2.5, r_2 = 1$. Therefore, $Q = 1$, i.e., any population started from an individual who reproduces according to Equation (**1**) will certainly go extinct. Fisher [3] recognized the dependence of Q on μ. If $\mu \leq 1$, the mutant will ultimately die out. If $\mu > 1$, then there is a non-zero chance of survival. Such findings led to the subsequent classification of subcritical ($\mu < 1$), critical ($\mu = 1$) and supercritical ($\mu > 1$) conditions [5,6]. Under subcritical and critical conditions, $Q = 1$. Under supercritical conditions, $Q < 1$.

The calculation above allows us to conclude that, even if the invasive hornet were lucky to produce multiple queens in certain years, it will certainly go extinct. However, how long will it take for the invasive hornet to go extinct?

2.2. Extinction Time T and Var(T)

A more difficult problem than the extinction probability is the extinction time. The individual (and its descendants) could go extinct at time t_1, t_2, t_3, etc., with their corresponding probabilities, so this extinction time should be characterized as a distribution.

Designate population size at time t as Z_t. and the extinction probability of the population by generation t as $q_t = \Pr(Z_t = 0)$. We start the process with a single individual at time 0, so $\boldsymbol{Z_0 = 1}$. We can compute q_t by iteration using the PGF in Equation (**9**):

$$\boldsymbol{q_{t+1} = G(q_t)} \tag{11}$$

For the first few generations:

$$\boldsymbol{q_1 = G(q_0) = G(0) = 0.5} \tag{12}$$

$$\boldsymbol{q_2 = G(q_1) = G(0.5) = 0.5 + 0.3 \times 0.5 + 0.2 \times 0.5^2 = 0.7} \tag{13}$$

$$\boldsymbol{q_3 = G(q_2) = G(0.7) = 0.808} \tag{14}$$

$$\boldsymbol{q_4 = G(q_3) \approx 0.876} \tag{15}$$

The functions above introduce us to a mathematical term called composition, e.g., q_4 is a composition of the function G to itself 4 times. A more descriptive term is iterate [5], i.e., q_4 is an iterate of q_3 or the 4th iterate of q_0. As t increases, q_t will increase monotonously towards Q. Under subcritical conditions, $Q = 1$, so q_t will increase monotonously towards 1.

What is the probability that extinction happens exactly at time t, i.e., $\mathbf{Pr}(\boldsymbol{T = t})$? This $\mathbf{Pr}(\boldsymbol{T = t})$ is simply $q_t - q_{t-1}$. Thus

$$\mathbf{Pr}(\boldsymbol{T = 1}) = \boldsymbol{q_1 - q_0 = 0.5 - 0 = 0.5} \tag{16}$$

$$\mathbf{Pr}(\boldsymbol{T = 2}) = \boldsymbol{q_2 - q_1 = 0.7 - 0.5 = 0.2} \tag{17}$$

$$\mathbf{Pr}(\boldsymbol{T = 3}) = \boldsymbol{q_3 - q_2 = 0.808 - 0.7 = 0.108} \tag{18}$$

$$\mathbf{Pr}(\boldsymbol{T = 4}) = \boldsymbol{q_4 - q_3 = 0.876 - 0.808 = 0.068} \tag{19}$$

We can calculate T by the following equation:

$$\boldsymbol{T = \sum_{t=1}^{\infty} t \ \Pr(T = t) = 1 \times 0.5 + 2 \times 0.2 + \cdots = 2.3954} \tag{20}$$

This equation works because $\sum_{T=1}^{\infty} \Pr(T) = 1$. If the extinction probability $Q < 1$, then we should use the weighted average:

$$\boldsymbol{T = \frac{\sum_{t=1}^{\infty} t \ \Pr(T = t)}{Q} = \frac{\sum_{t=1}^{\infty}[t\,(q_t - q_{t-1})]}{\sum_{t=1}^{\infty}(q_t - q_{t-1})}} \tag{21}$$

An alternative way of calculating T under subcritical conditions is

$$\boldsymbol{T = \sum_{t=1}^{\infty}(1 - q_t) = (1 - 0) + (1 - 0.5) + (1 - 0.7) + \cdots = 2.3954} \tag{22}$$

The variance of T can be computed by the general equation:

$$\boldsymbol{Var(T) = E(T^2) - [E(T)]^2} \tag{23}$$

We already know $E(T) = 2.3954$, so $[E(T)]^2 = 5.7379$. The first term can be calculated in either of two ways below.

$$\boldsymbol{E(T^2) = \frac{\sum_{t=1}^{\infty} t^2 \ \Pr(T = t)}{\sum_{t=1}^{\infty} \Pr(T = t)} = \sum_{t=1}^{\infty} t^2 \ \Pr(T = t) = 1^2 \cdot 0.5 + 2^2 \cdot 0.2 + \cdots \approx 10.7818} \tag{24}$$

$$\boldsymbol{E(T^2) = \sum_{t=0}^{\infty}(2t + 1)(1 - q_t) \approx 10.7818 \ (under\ subcriticla\ condition)} \tag{25}$$

$$\boldsymbol{Var(T) = E(T^2) - [E(T)]^2 = 10.7818 - 5.7378 = 5.0439} \tag{26}$$

Equation (**24**) is generally applicable. Equation (25) can be used only under subcritical conditions. Equation (**22**) and Equation (25) can be used only under subcritical conditions. In contrast, Equation (**21**) and Equation (**24**) are general and consequently can be used in subcritical, critical and supercritical conditions.

This calculation of T and its variance allows us to have more informative predictions on the fate of the invasive hornet. Not only will the invasive hornet will go extinct, but it has a probability of 0.95 of going extinct in $2.3954 + 1.96 \times \sqrt{5.0439}$ generations. The Asian hornets, the number of generations is equivalent to the number of years.

The calculation above is summarized in Table 1, where t is the generation, q_t in the second column is the probability of extinction by generation t, $q_{t+1} - q_t$ in the third column is the probability of extinction that occurs exactly in generation t, i.e., $\Pr(T = t)$. Note that $(q_t - q_{t-1})$ is the generation-specific extinction probability, whereas q_t is the cumulative probability of extinction. T can be calculated either as the weighted average of t (weighted by $\Pr(T = t)$, i.e., the third column in Table 1) as shown in Equation (**20**), or as the summation of $1 - q_t$ in the fourth column in Table 1, as shown in Equation (**22**). The last column is for calculating $E(T^2)$ needed for computing $Var(T)$.

Table 1. Numerical method for computing extinction time T and its variance.

T	qt	qt–qt-1	1-qt	(2t+1)(1-qt)
0	0		1	1
1	0.5	0.5	0.5	1.5
2	0.7	0.2	0.3	1.5
3	0.808	0.108	0.192	1.344
4	0.872973	0.064973	0.127027	1.143245
5	0.914308	0.041335	0.085692	0.94261
6	0.941484	0.027176	0.058516	0.760704
..	…	…	…	…
37	0.999999	3.75E-07	8.74E-07	6.56E-05

38	0.999999	2.62E-07	6.12E-07	4.71E-05
39	1	1.84E-07	4.28E-07	3.38E-05
40	1	1.28E-07	3E-07	2.43E-05
..	…	…	…	…
92	1	1.22E-15	2.55E-15	4.72E-13
93	1	0	1.89E-15	3.53E-13
94	1	0	0	0

What would be the extinction time T and its $Var(T)$ under supercritical conditions? T can be calculated the same way by using Equation (**21**), i.e., it is conditional on the extinction probability. $Var(T)$ can also be calculated as before:

$$\boldsymbol{Var(T) = E(T^2) - [E(T)]^2 = \frac{\sum_{t=1}^{\infty} t^2 \Pr(T=t)}{\sum_{t=1}^{\infty} \Pr(T=t)} - T^2} \tag{27}$$

For example, in the supercritical case above with $\boldsymbol{E(X)} = 1.1$, we have $T \approx 5.229, Var(T) = 43.826$. Note that T under the supercritical condition could be small when $\boldsymbol{E(X)}$ is large. This is because such a population will either have bad luck and go extinct quickly (i.e., a small T), or increase in population size and never go extinct. For example, if the probabilities of producing 0, 1, or 2 offspring per generation are 0.1, 0.4, and 0.5, respectively, then $\boldsymbol{E}(X) = 1.5$. The resulting $T = 2.159, Var(T) = 3.026$.

3. A branching Process with Reproductive Success Following a Poisson Distribution

Both Fisher [3] and Haldane [4] used the Poisson distribution as the reproduction law to model reproductive success in their use of the branching process, i.e., the probability of an individual producing 0, 1, ..k, .. offspring is $\boldsymbol{f(k) = e^{-\lambda}}\frac{\lambda^k}{k!}$. This is also used in epidemiological studies. Let X be the number of individuals who would become infectious "descendants" of an infected individual (before the individual either dies of the disease or clears the infectious pathogen). The PGF and the associated mean and variance are

$$\boldsymbol{G(s) = \sum_{k=0}^{Max\,k} f(k)s^k = e^{-\lambda} + e^{-\lambda}\lambda\, s + e^{-\lambda}\frac{\lambda^2}{2!}\, s^2 + e^{-\lambda}\frac{\lambda^3}{3!}\, s^3 + \cdots} \tag{28}$$

$$\boldsymbol{G(s) = e^{-\lambda}\left(1 + \lambda\, s + \frac{(\lambda s)^2}{2!} + \frac{(\lambda s)^3}{3!} + \cdots\right) = e^{-\lambda}e^{\lambda s} = e^{\lambda(s-1)}} \tag{29}$$

$$\boldsymbol{G'(s) = e^{-\lambda}\lambda + 2e^{-\lambda}\frac{\lambda^2}{2!}\, s + 3e^{-\lambda}\frac{\lambda^3}{3!}\, s^2 + \cdots + ne^{-\lambda}\frac{\lambda^n}{n!}\, s^{n-1} + \cdots} \tag{30}$$

$$\boldsymbol{G''(s) = 2e^{-\lambda}\frac{\lambda^2}{2!} + 6e^{-\lambda}\frac{\lambda^3}{3!}s^1 + \cdots + n(n-1)e^{-\lambda}\frac{\lambda^n}{n!}s^{n-2} + \cdots} \tag{31}$$

$$\boldsymbol{E(X) = G'(1) = e^{-\lambda}\lambda\left(1 + \lambda s + \frac{(\lambda s)^2}{2!} + \cdots\right) = \lambda} \tag{32}$$

$$\boldsymbol{G''(1) = e^{-\lambda}\lambda^2\left(1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \cdots\right) = e^{-\lambda}\lambda^2 e^{\lambda} = \lambda^2} \tag{33}$$

$$\boldsymbol{Var(X) = G''(1) + G'(1) - [G'(1)]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda} \tag{34}$$

3.1. Extinction Probability Q

Obviously, the disease propagation depends on $\boldsymbol{\lambda}$. If $\lambda \leq 1$, then disease will die out. If $\lambda > 1$, then the number of descendants will increase. We will first address the same question of extinction probability *Q*. From what we have learned, we need to solve $Q = G(Q)$ and find the smallest non-negative root.

From Equation (29), we have

$$\boldsymbol{Q = G(Q) = e^{\lambda(Q-1)}} \tag{35}$$

In the critical case when $E(X) = \lambda = 1$, then $\boldsymbol{Q = e^{(Q-1)}}$, and the only root is $Q = 1$, i.e., extinction is certain. This also implies that all cases with $\lambda < 1$ should also have $Q = 1$ because extinction would be even more likely with smaller λ. Thus, $Q = 1$ for all $\lambda \leq 1$.

What would be the extinction probability under supercritical conditions with $\lambda > 1$? Haldane [4] and Felsenstein [10] illustrated the solution with an λ only slightly larger than 1 (i.e., a small selection coefficient $s = \lambda - 1$) so that one can expand the right-hand side of Equation (**35**) as a power series and drop high-order terms. This would not work with larger λ. A better way to compute *Q* under supercritical conditions is to use the following equation

$$\boldsymbol{Q = -\frac{1}{\lambda} W_0\left(-\lambda e^{-\lambda}\right)} \tag{36}$$

where W_0 is the principal branch of the Lambert W function [11]. $W_0(z)$ is defined as

$$\boldsymbol{z = W_0(z) e^{W_0(z)}} \tag{37}$$

It can be calculated in many numerical ways [11], e.g., by using the ‘optim’ function in R or the Solver function in EXCEL. Table 2 lists *Q* values given different λ values. In population genetics where λ is the absolute fitness of a mutant, the selection coefficient $s = \lambda - 1$ (against an implicitly assumed stable wile-type population).

Table 2. Relationship between λ and Q for $\boldsymbol{\lambda} > 1$. The calculation of W_0 is explained in the text.

λ	$z = -\lambda e^{-\lambda}$	$W_0(z)$	Q
1	-0.36788	-1	1
1.01	-0.36786	-0.99014	0.98034
1.1	-0.36616	-0.90630	0.82391
1.2	-0.36143	-0.82353	0.68627
1.3	-0.35429	-0.75013	0.57702
1.4	-0.34524	-0.68461	0.48901
1.5	-0.33470	-0.62581	0.41720
2	-0.27067	-0.40637	0.20319
3	-0.14936	-0.17856	0.05952
4	-0.07326	-0.07931	0.01983
5	-0.03369	-0.03489	0.00698
10	-0.00045	-0.00045	0.00004

Epidemiologists have documented the basic reproduction number (R_0, which is equivalent to λ in Table 2) for various viral diseases. For COVID-19, R_0 varies from 2 to 7 [12] which corresponds to very small *Q*.

The calculation above is also applicable to many evolutionary scenarios, such as the fate of a mutant in a haploid population or of a parthenogenetic individual arising from a sexual population. Imagine a large, stable and genetically homogenous haploid population in which everyone is expected to produce one offspring to replace itself ($\lambda = 1$). A new mutant arises with $\lambda = 1.01$, i.e., it is fitter than other individuals in the population with a fitness differential $s = 0.01$. Even such a beneficial mutation would have an extinction probability $Q = 0.98034$ (Table 2). Its fixation probability is $(1-Q) = 1-0.98034 \approx 0.02 = 2s$. For $s = 0.001, Q = 0.998002, 1-Q = 0.002 = 2s$. It is generally true that, for a small s, the fixation probability is $2s$ when the reproductive success is specified by a Poisson distribution.

While our illustration here is applicable to a haploid population, the conclusion is also true for diploid populations [4,10]. In a large population of wild-type individuals with AA genotypes, each individual produces 0, 1, 2, … k offspring with corresponding probabilities $p_0, p_1, p_2, \dots, p_k$, respectively. What is Q for a new B allele that creates an AB heterozygote? If all alleles are neutral, then the probability of fixation of a new allele is simply its allele frequency, i.e., $1/(2N)$ for a new allele, where N is the population size of the diploid population. Thus, $Q = 1-1/(2N)$. What would be the extinction probability and extinction time when the B allele is advantageous?

The two previous examples revealed that, if an advantageous individual or allele is to go extinct because of bad luck, the extinction happens quickly. If the new beneficial allele B overcomes the bad luck and becomes frequent, then it tends to achieve fixation. This allows us to assume that, during the process of the allele going extinct, its frequencies are low, and it therefore exists in heterozygotes instead of BB homozygotes. For convenience, we may also assume that the B allele is dominant or codominant so that its beneficial effect will manifest in heterozygotes. Given these conditions, the extinction probability can be written in the form that we are already familiar with:

$$\boldsymbol{Q = p_0 + p_1 Q + p_2 Q^2 + p_3 Q^3 + \cdots + p_k Q^k + \cdots} \tag{38}$$

If we further assume that p_i follows the Poisson distribution, then $Q = e^{\lambda(Q-1)}$ as we have derived before in Equation (**35**). The solution for Q and numeric estimates for T and $Var(T)$ have already been derived above. The assumption of independent propagation of the new allele B in the population is reflected in those terms with $Q^2, Q^3, Q^4, \dots$ In other words, all B alleles exist in AB heterozygotes and independently have the same extinction probability as the first AB heterozygote created by the mutation. With Equation (**36**) we can readily compute Q, T, and $Var(T)$ for any λ.

The calculation is also applicable to the extinction probability and extinction time for a new parthenogenetic individual in a diploid sexual population. Such a parthenogenetic individual is supposed to have a two-fold advantage in fitness [13,14], i.e., with $\lambda = 2$ given that an average sexual individual's reproductive output is $\lambda = 1$. The extinction probability for this new parthenogenetic individual is $Q = 0.20319$ (Table 2).

3.2. Extinction Time

Let us first consider the subcritical condition with $\lambda = 0.7$. This way, we can compare the results to those in the first example where $E(X) = 0.7, Var(X) = 0.61$. For a Poisson distribution, $E(X) = Var(X) = \lambda = 0.7$. Thus, we expect the mean and variance of the extinction time to be similar to those in our first example. Again, here are the first few $\boldsymbol{q_t}$ values.

$$\boldsymbol{q_1 = G(q_0) = G(0) = e^{-\lambda} \approx 0.4966} \tag{39}$$

$$\boldsymbol{q_2 = G(q_1) = G\left(e^{-\lambda}\right) \approx 0.7030} \tag{40}$$

$$\boldsymbol{q_3 = G(q_2) \approx 0.8123} \tag{41}$$

$$\boldsymbol{q_4 = G(q_3) \approx 0.8769} \tag{42}$$

In all subcritical cases, q_t will increase monotonously towards 1. The probability that extinction happens exactly at time t, i.e., $\mathbf{Pr}(\boldsymbol{T = t})$ can also be calculated as before:

$$\mathbf{Pr}(\boldsymbol{T} = \mathbf{1}) = \boldsymbol{q_1 - q_0 = 0.4966} \tag{43}$$

$$\mathbf{Pr}(\boldsymbol{T} = \mathbf{2}) = \boldsymbol{q_2 - q_1 = 0.2064} \tag{44}$$

$$\mathbf{Pr}(\boldsymbol{T} = \mathbf{3}) = \boldsymbol{q_3 - q_2 = 0.808 - 0.7 = 0.1093} \tag{45}$$

doi:10.20944/preprints202511.1105.v1

$$\Pr(T = 4) = q_4 - q_3 = 0.876 - 0.808 = 0.0646 \tag{46}$$

Given the distribution, we can calculate *T* with the two formulae that we have used before for subcritical conditions:

$$T = \sum_{t=1}^{\infty} t \Pr(T = t) = 1 \times 0.4966 + 2 \times 0.2064 + \cdots \approx 2.3762 \tag{47}$$

$$T = \sum_{t=0}^{\infty} (1 - q_t) = (1 - 0) + (1 - 0.4966) + (1 - 0.7030) + \cdots \approx 2.3762 \tag{48}$$

What would be the extinction time *T* and its *Var*(*T*) under supercritical conditions? *T* can be calculated the same way by using Equation (**21**). *Var*(*T*) can also be calculated with Equation (27). For example, if $\lambda = 1.1$, then $T = 4.102, Var(T) = 33.550$. Table 3 lists the q_t and ($q_t - q_{t-1}$) values needed to computing *T* and *Var*(*T*) for $\lambda = 1.1$. We already have listed $Q = 0.82391$ for $\lambda = 1.1$ in Table 2, which is consistent with the q_t column in Table 3 that comes very close to this number when *t* reaches 1000 generations.

Table 3. The q_t and $q_t - q_{t-1}$ for computing extinction time *T* and its variance *Var*(*T*) for $\lambda = 1.1$.

T	q_t	$q_t - q_{t-1}$
0	0.0000000000	
1	0.3328710837	0.332871084
2	0.4800611401	0.147190056
3	0.5644334773	0.084372337
4	0.6193261945	0.054892717
5	0.6578744403	0.038548246
6	0.6863702213	0.028495781
7	0.7082254834	0.021855262
8	0.7254580952	0.017232612
…	…	…
999	0.8238658564	0
1000	0.8238658564	0

In our numeric illustration of the supercritical condition with $\lambda = 1.1$, *T* is longer than the subcritical condition with $\lambda = 0.7$. However, it is possible for a smaller *T* under supercritical conditions. For example, if $\lambda = 2$, then $T = 1.545, Var(T) = 0.953$. This is because the extinction time is conditional on extinction. An individual that reproduces with $\lambda = 2$ will either have bad luck and go extinct quickly (i.e., a small *T*), or increase in population size rapidly and never go extinct.

4. Come Back to the Galton Questions

Galton asked two questions. N adult men each have a unique surname. In each generation, each man produces *k* adult sons with probability a_k. Galton limited *k* to a maximum of five. The PGF for addressing Galton's two questions is similar to our first example:

$$\begin{aligned}
G(s) &= a_0 + a_1 s + a_2 s^2 + a_3 s^3 + a_4 s^4 + a_5 s^5 \\
G'(s) &= a_1 + 2a_2 s + 3a_3 s^2 + 4a_4 s^3 + 5a_5 s^4 \\
G''(s) &= 2a_2 + 6a_3 s + 12a_4 s^2 + 20a_5 s^4 \\
E(X) &= G'(1) = a_1 + 2a_2 + 3a_3 + 4a_4 + 5a_5 \\
Var(X) &= G''(1) + G'(1) - [G'(1)]^2
\end{aligned} \tag{49}$$

Suppose we have $a_0 = 0.1, a_1 = 0.3, a_2 = 0.25, a_3 = 0.15, a_4 = 0.1, a_5 = 0.1$. The mean number of sons produced by a man (*n*) and the variance of *n* are

$$E(X) = G'(1) = 2.15 \tag{50}$$

$$\boldsymbol{Var(X) = G''(1) + G'(1) - [G'(1)]^2 = 2.1275}$$

This is a supercritical condition, so the population size will more than double itself each generation. The extinction probability $Q < 1$. Solving the equation $Q = G(Q)$ gives us $Q = 0.1519416038$. The probability of extinction by generation t is q_t which we have calculated recursively in previous examples. In mathematics, this q_t is calculated as the t-fold composition of G or the tth iterate of G;

$$\boldsymbol{q_t = G^{(t)}(0)} \tag{51}$$

This function does not have a closed form, which is why we resorted to computing q_t recursively in previous examples. For the first few generations:

$$\begin{aligned} &\boldsymbol{q_0 = 0} \\ &\boldsymbol{q_1 = G(q_0) = G(0) = a_0 = 0.1} \\ &\boldsymbol{q_2 = G(q_1) = G(0.1) = 0.132661} \\ &\boldsymbol{q_3 = G(q_2) = .1445833203} \end{aligned} \tag{52}$$

Thus, after one generation, 10% of the *N* surnames are lost. At generation 23, the proportion of surnames lost has already reached *Q* (=0.1519416038). After that, all remaining surnames would already have escaped bad luck and will survive forever, if the earth were infinitely large. The extinction time *T* is only 1.547654, with *Var*(*T*)=0.916671. This means that ~15.2% of lost surnames must have been lost in early generations, otherwise a few generations of rapid increase in population size (more than doubling in each generation) would secure their permanent existence.

Thus, to address Galton's first question, as long as a_i values and *t* are given, it is easy to recursively compute q_t, which is the proportion of surnames that have gone extinct by generation *t* (Table 4). For example, about 14.9% of the surnames are expected to perish by generation 4 under the extremely favorable condition with $\mu = 2.15$.

Table 4. q_t values for $a_0 = 0.1, a_1 = 0.3, a_2 = 0.25, a_3 = 0.15, a_4 = 0.1, a_5 = 0.1$.

T	q_t	$q_t - q_{t-1}$
0	0	
1	0.1000000000	0.1000000000
2	0.1326610000	0.0326610000
3	0.1445833203	0.0119223203
4	0.1491044603	0.0045211400
5	0.1508434063	0.0017389461
…	…	…
23	0.1519416037	0.0000000001
24	0.1519416038	0.0000000000
25	0.1519416038	0.0000000000

Galton's second question is trickier. At generation *t*, how many of those starting *N* surnames are carried by 0, 1, 2, …, men? Our illustrative example has $a_0 = 0.1, a_1 = 0.3, a_2 = 0.25, a_3 = 0.15, a_4 = 0.1, a_5 = 0.1$. Suppose we have *N* = 10000 men each bearing a unique surname at time *t* = 0. At time *t* = 1, we expect 1000 ($= N \cdot a_0$) surnames to perish. We expect 3000, 2500, 1500, 1000 and 1000 surnames, respectively, to have 1, 2, 3, 4 and 5 bearers, respectively (Figure 2A). At time *t* = 2, the proportion of perished surnames is expected to increase from 10% to 13.277% (Figure 2B). The distribution is also wider (Figure 2B), because there are some lucky surnames that might have as many as 25 (= 5 × 5) bearers after the second generation. For example, the man with a surname Clark might produce 5 sons in the first generation, and these sons each produce 5 sons in the second generation. Now the surname Clark has 25 bearers. However, the probability of having such good luck is small, i.e., the product of a_5 from the first generation multiplied by a_5^5 in the second generation (i.e., each of the five sons independently produces five sons in the second generation.

There are about 1159 surnames with two bearers after the second generation (Figure 2B). After the third generation, the largest possible number of bearers of a surname is 125 (=5^3) (Figure 2C). However, the probability for such an event to occur is 10^{-31}. The total probability of having a surname with 50 or more bearers is only 3.24^{-4}. Note that this is under extremely favorable conditions with $\mu = 2.15$. In a stable population with $\mu = 1$ (e.g., when $a_0 = 0.45, a_1 = 0.35, a_2 = 0.15, a_3 = 0.06, a_4 = 0.03, a_5 = 0.01$), then the surnames lost would be 45% after the first generation, 64% after the second generation, and 78% after the third generation. The total probability for surnames with 10 or more bearers is only 2.97^{-3} after three generations.

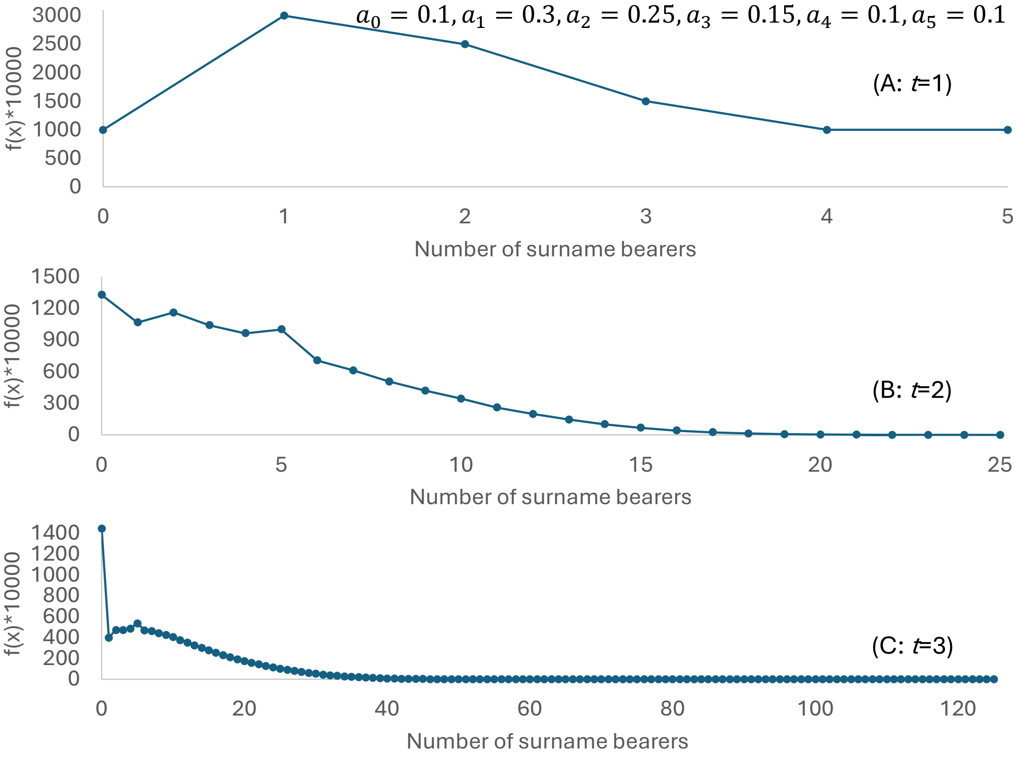


Figure 2. Probability distribution of surnames with 0, 1, 2, … bearers after the first three generations, multiplied by 10000. Reproduction is specified by a_k values, with $\mu = 2.15$.

5. Recent Extensions and Applications

The largest number of illustrations and potential biological application of the branching process is available in a book [15]. Recent studies have extended the classical branching process in several ways. The first is parameter estimation. Equation (1) listed three parameter values, 0.5, 0.3 and 0.2 that defines the reproduction law needed for the branching process. The λ parameter in the Poisson distribution is also a parameter that can be estimated from empirical data. Such estimation can be done either nonparametrically or by maximum likelihood or Bayesian approaches [16], the last being particularly relevant for small populations with limited empirical data.

Second, the branching process itself can be generalized. In our illustration, the same reproduction law is applied to all individuals. However, parental reproductive success could be correlated with offspring reproductive success. Such a scenario implies that some lineages are less likely to go extinct than others. For example, successful mothers tend to have successful daughters and so on. One can modify the reproduction law to accommodate this parent-offspring correlation in reproductive success [17]. Also, my illustrative examples are for populations with discrete generations. Under certain circumstances, the branching process can also accommodate populations with overlapping generations [17].

The theory of multi-type branching processes has been developed a long time ago [5,6], but only recently been applied to solve practical biological problems on the consequences of dormancy on population demography. A haploid population was implicitly assumed. The empirical challenge is to determine what environmental factors trigger the switch between the active and dormant state and whether the probability of switching between the two states depends quantitatively on the degree of environmental harshness.

Conclusions

This paper aims to popularize the branching process with the objective of highlighting its potential application. The illustrative examples can be easily extended to more complicated scenarios. For example, the reproduction law was illustrated with the empirically determined distribution and with the Poisson distribution, but one can replace the reproduction law with a geometric distribution or a negative binomial distribution to model more unequal reproductive success among individuals, e.g., most individuals fail to reproduce but a few produce many offspring.

Funding: This research was funded by a Discovery Grant from the Natural Science and Engineering Research Council (NSERC, RGPIN-2024-05641) of Canada. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results

Acknowledgements: I thank S. Lion and members of XiaLab for comments and discussion.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Feller, W., *An Introduction to Probability Theory and Its Applications*. Wiley series in probability and mathematical statistics, ed. R.A. Bradley, et al. Vol. 1. 1968, New York: Wiley. 509.
2. Galton, F. and H.W. Watson, *On the probability of extinction of families.* Journal of the Anthropological Institute of Great Britain and Ireland, 1874. **4**: p. 138-144.
3. Fisher, R.A., *On the dominance ratio.* Proc. Royal. Soc. Edin., 1922. **42**: p. 321-431.
4. Haldane, J.B.S., *A mathematical theory of natural and artificial selection. Part V.* Proceedings of the Cambridge Philosophical Society, 1927. **23**: p. 838-844.
5. Harris, T.E., *The Theory of Branching Processes*. 1963: Springer Berlin Heidelberg.
6. Athreya, K.B., P.E. Ney, and P.E. Ney, *Branching Processes*. 2004, Mineola: Dover Publications. 287.
7. Freeman, A. and X. Xia, *Phylogeographic Reconstruction to Trace the Source Population of Asian Giant Hornet Caught in Nanaimo in Canada and Blaine in the USA.* Life, 2024. **14**(3): p. 283.
8. Xia, X., *Phylogeographic Analysis for Understanding Origin, Speciation, and Biogeographic Expansion of Invasive Asian Hornet, Vespa velutina Lepeletier, 1836 (Hymenoptera, Vespidae).* Life, 2024. **14**(10): p. 1293.
9. Azevedo, R.B.R. and P. Olofsson, *A branching process model of evolutionary rescue.* Math Biosci, 2021. **341**: p. 108708.
10. Felsenstein, J., *Theoretical evolutionary genetics*. 2019, Seattle, Washington: Self-published.
11. Corless, R.M., et al., *On the Lambert W function.* Advances in Computational Mathematics, 1996. **5**: p. 329-359.
12. Viceconte, G. and N. Petrosillo, *COVID-19 R0: Magic number or conundrum?* Infect Dis Rep, 2020. **12**(1): p. 8516.
13. Williams, G.C., *Sex and evolution*. 1975, Princeton, New Jersey: Princeton University Press.
14. Maynard Smith, J., *The Evolution of Sex*. 1978, Cambridge: Cambridge University Press.
15. Haccou, P., P. Jagers, and V.A. Vatutin, *Branching Processes: Variation, Growth, and Extinction of Populations*. Cambridge Studies in Adaptive Dynamics. 2005, Cambridge: Cambridge University Press.
16. Cloez, B., et al., *A bayesian version of Galton–Watson for population growth and its use in the management of small population.* MathematicS In Action, Maths Bio, 2023. **12**(1): p. 49-64.
17. Sagitov, S., *Critical Galton–Watson Processes with Overlapping Generations.* 2021. **36**(2): p. 87-110.