Mathematica Aeterna, Vol. 1, 2011, no. 02, 65 - 72

On Probability Distribution of First Absorption Time in Branching Processes

Wai-Yuan Tan

Department of Mathematical Sciences
The University of Memphis, Memphis, TN 38152, USA

Abstract

In this paper, by using diffusion approximation we derive the probability distribution of time to extinction of the Galton-Watson branching process. It is shown that this probability distribution can be expressed in terms of Laguerre polynomials.

Mathematics Subject Classification: 60E99, 60G99, 60J60,60J80,60J85

Keywords: Diffusion approximation, First absorption probability, Galton-Watson branching process, Laguerre polynomial.

1 Introduction

Consider a Galton-Watson branching process $\{X(t), t \in T = (0, 1, \ldots,)\}$ with progeny distribution $\{q_j, j = 0, 1, \ldots, \}$. Let μ and σ^2 denote the mean and the variance of the progeny distribution respectively. Then, the fundamental theorem of the branching process indicates that the probability that the process will eventually be distinct is one if and only if (iff) μ is less than or equal to 1 (for proof, see Karlin and Taylor [1], p. 396). That is, $\lim_{t\to\infty} P\{X(t) = 0|X(0) = i\} = 1$ iff $\mu \leq 1$ for any positive integer i.

Let $\phi(t; i)$ be the probability that the process absorbs into 0 for the first time at time t given $X(0) = i$. Then $\sum_{t=1}^{\infty} \phi(t; i) = 1$ iff $\mu \leq 1$. That is, given that $\mu \leq 1$, $\phi(t; i)$ is the probability density function (pdf) of the first absorption time T_a of the process. In many practical problems, it is of considerable interest to find $\phi(t; i)$. For example, in human beings, it is well documented that many of the inherited diseases are caused by mutation of certain genes; see (Jorde et al. [2], and Scriver et al.[3]). Hence, it is extremely useful to derive the probability distribution of extinction of deleterious mutants.

How to derive $\phi(t;i)$? For general progeny distributions, the problem remains unsolved. In this paper, by using the diffusion approximation, we will provide an approximate solution to this problem. We will derive the probability distribution of first time absorption in branching processes. We will illustrate how to derive the probability of distinction of new mutants in human populations.

2 Preliminary Notes

Suppose that the mean μ of the progeny distribution is $\mu = 1 + \frac{1}{N}\alpha + O(N^{-2})$, where N is very large. Then, $\mu \leq 1$ iff $\alpha \leq 0$. Thus, if $\alpha \leq 0$, $\phi(t;i)$ is the pdf of the first absorption time T_i of the process given $X(0) = i$. Let $Y(t) = \frac{1}{N}X(t)$ and let one time unit dt corresponding to $\frac{1}{N}$. Then it can be shown that to order of $O(N^{-2})$, $\{Y(t), t \geq 0\}$ is a diffusion process with state space $\Omega = [0, \infty)$ and with diffusion coefficients $\{m(y) = \alpha\ y = -\xi y, v(y) = y\ \sigma^2\}$, where $\xi = -\alpha \geq 0$. This result was first proved by Feller [4]. We state this result as in the following theorem.

Theorem 2.1 *To order of* $O(N^{-2})$, $\{Y(t), t \geq 0\}$ *is a diffusion process with state space* $\Omega = [0, \infty)$ *and with diffusion coefficients* $\{m(y) = -y\xi, v(y) = y\sigma^2\}$.

For proof, see Feller[4] or Tan ([6], Chapter 6).

3 Main Result 1. The Transition Probability Distribution of Diffusion Process

Let $f(x, y; t)$ be the conditional pdf of $Y(t)$ given $Y(0) = \frac{i}{N} = x$. By Theorem (2.1), to order of $O(N^{-2})$, $f(x, y; t)$ satisfies the following backward equation with initial condition $f(x, y; 0) = \delta(y - x)$, where $\delta(x)$ is the Dirac's δ function defined by $\int_{-\infty}^{\infty} f(x)\delta(x)dx = f(0)$ for any integrable function $f(x)$ over the real line:

$$\frac{\partial}{\partial t} f(x, y; t) = m(x)\frac{\partial}{\partial x} f(x, y; t) + \frac{1}{2}v(x)\frac{\partial^2}{\partial x^2} f(x, y; t). \tag{1}$$

The following theorem shows that the solution of equation (1) can be expressed in terms of Laguerre polynomials. (For basic properties of Laguerre polynomials, see [6], Chapter 1).

Theorem 3.1 *If $\xi > 0$, then the solution $f(x,y;t)$ of equation (1) under the initial condition $f(x,y;0) = \delta(y-x)$ is given by:*

$$f(x,y;t) \quad = \quad \beta(\beta\, x)e^{-\beta\, y}\sum_{k=1}^{\infty}\frac{1}{k}e^{-k\xi\, t}L_{k-1}^{(2)}(\beta\, x)L_{k-1}^{(2)}(\beta\, y), \tag{2}$$

where $\beta = 2\xi/\sigma^2$ and $L_{k-1}^{(2)}(x)$ the $(k-1)-th$ degree Laguerre polynomial with parameter 2.

Proof

To prove Theorem (3.1), write $f(x,y;t) = f(x,t)$ by suppressing y and let $f(x,t) = h(x)e^{-\lambda\, t}$, where λ is a constant. Then $h(x)$ satisfies the equation:

$$x\sigma^2\frac{d^2}{dx^2}h(x) - 2\xi\, x\frac{d}{dx}h(x) + 2\lambda\, h(x) = 0. \tag{3}$$

In the above equation, the λ's (say λ_j) satisfying the above equation are the eigenvalues of the operator $S = \frac{\sigma^2}{2}x\frac{d^2}{dx^2} - \xi\, x\frac{d}{dx}$ and the solution $h_j(x)$ with $\lambda = \lambda_j$ in equation (3) is an eigenfunction corresponding to the eigenvalue λ_j. The general solution $f(x,y;t) = f(x,t)$ of equation (2) is given by:

$$f(x,y;t) = f(x,t) = \sum_j C_j h_j(x)e^{-\lambda_j\, t}, \tag{4}$$

where the C_j's are constants to be determined by the initial condition $f(x,y;0) = f(x,0) = \delta(y-x)$.

To find these eigenvalues and the associated eigenfunctions, consider the series solution $h(x) = \sum_{i=0}^{\infty} a_i x^i$. On substituting this series solution into equation (3) and equating to zero the coefficient of x^k for $k = 0, 1, \ldots$, we obtain $a_0 = 0$ and for $k = 1, \ldots, \infty$,

$$a_{k+1} = \frac{2}{(k+1)k\sigma^2}(k\xi - \lambda)a_k.$$

In order that the solution is finite for finite x and not identically 0, we must require that $a_1 \neq 0$ and $\lambda - \xi\, k = 0$, or $\lambda_k = \xi\, k$, for $k = 1, \ldots, \infty$.

Now, given $\lambda = k\xi$ we have $a_{k+j} = 0$ for all $j = 1, \ldots,$ and for $j = 2, \ldots, k$:

$$\begin{aligned} a_j &= \frac{2\xi(j-1-k)}{j(j-1)\sigma^2}a_{j-1} \\ &= \frac{1}{j!(j-1)!}(2\xi/\sigma^2)^{j-1}(j-1-k)\ldots(1-k)a_1 \end{aligned}$$

$$
\begin{aligned}
&= (-\beta)^{j-1}\frac{1}{j!}\binom{k-1}{j-1}a_1 \\
&= (-\beta)^{j-1}\frac{1}{\Gamma(j+1)}\binom{k-1}{j-1}a_1,
\end{aligned}
$$

where $\beta = 2\xi/\sigma^2$.

Hence, given $\lambda = \lambda_k = k\xi$, the solution of equation (3) is

$$
\begin{aligned}
h_k(x) &= \sum_{j=1}^{k} a_j x^j = a_1 \sum_{j=1}^{k}(-\beta)^{j-1}\binom{k-1}{j-1}x^j\frac{1}{\Gamma(j+1)} \\
&= a_1 x \sum_{j=1}^{k}(-1)^{j-1}\frac{1}{\Gamma(j+1)}\binom{k-1}{j-1}(\beta\, x)^{j-1} \\
&= \frac{a_1}{k}x\frac{1}{(k-1)!}\sum_{j=0}^{k-1}(-1)^j\binom{k-1}{j}(\beta\, x)^j\frac{\Gamma(k-1+2)}{\Gamma(j+2)} \\
&= \frac{a_1}{k}xL_{k-1}^{(2)}(\beta x),
\end{aligned}
$$

where $L_k^{(\omega)}(x) = \frac{1}{(k)!}\sum_{j=0}^{k}(-1)^j\binom{k}{j}(x)^j\frac{\Gamma(k+\omega)}{\Gamma(j+\omega)}$ is the kth degree Laguerre polynomial with parameter ω; see Tan and Tiku ([6], Chapter 1).

Using these results, the general solution of equation (1) is

$$
f(x,y;t) = x\sum_{k=1}^{\infty} C_k e^{-\xi\, kt}L_{k-1}^{(2)}(\beta x), \tag{5}
$$

where the C_k's are constants to be determined by the conditions $f(x,y;0) = \delta(y-x)$ for all $x > 0$.

Now by the basic property of Laguerre polynomials as given in Tan and Tiku ([6], p 6), we have:

$$
\begin{aligned}
\int_0^\infty L_k^{(2)}(x)L_j^{(2)}(x)xe^{-x}dx &= 0 \text{ if } j \neq k \\
&= \binom{k+2-1}{k} = k+1 \text{ if } j = k.
\end{aligned}
$$

Multiplying both sides of equation (4) by $e^{-\beta x}L_{j-1}^{(2)}(\beta x)$, putting $t = 0$ and integrating the function from 0 to ∞, we obtain:

$$
\begin{aligned}
e^{-\beta\, y}L_{j-1}^{(2)}(y) &= C_j\int_0^\infty x[L_{j-1}^{(2)}(\beta\, x)]^2 e^{-\beta\, x}dx \\
&= C_j\beta^{-2}\int_0^\infty [L_{j-1}^{(2)}(x)]^2 xe^{-x}dx \\
&= C_j\beta^{-2}\binom{j-1+2-1}{j-1} = jC_j\beta^{-2}.
\end{aligned}
$$

Thus, $C_j = \frac{1}{j}\beta^2 L_{j-1}^{(2)}(\beta\, y)e^{-\beta\, y}$ for $j = 1, \ldots, \infty$ so that,

$$f(x, y; t) = \beta(\beta\, x)e^{-\beta\, y} \sum_{k=1}^{\infty} \frac{1}{k} e^{-k\xi t} L_{k-1}^{(2)}(\beta\, x) L_{k-1}^{(2)}(\beta\, y). \tag{6}$$

Q.E.D.

Let $Q(t; x)$ be the probability of $0 < Y(t)$ given $Y(0) = \frac{i}{N} = x$. Then

$$Q(t; x) = \int_0^{\infty} f(x, y; t) dy. \tag{7}$$

To derive $Q(t; x)$, notice that

$$\begin{aligned}
\beta \int_0^{\infty} e^{-\beta\, y} L_{k-1}^{(2)}(\beta y) dy &= \int_0^{\infty} e^{-y} L_{k-1}^{(2)}(y) dy \\
&= \frac{\Gamma(k-1+2)}{(k-1)!} \sum_{j=0}^{k-1} (-1)^j \binom{k-1}{j} \frac{1}{\Gamma(j+2)} \int_0^{\infty} e^{-z} z^j dz \\
&= k \sum_{j=0}^{k-1} (-1)^j \binom{k-1}{j} \frac{1}{(j+1)} \\
&= k \int_0^1 \{\sum_{j=0}^{k-1} (-1)^j \binom{k-1}{j} x^j\} dx \\
&= k \int_0^1 (1-x)^{k-1} dx = 1.
\end{aligned}$$

Thus,

$$\begin{aligned}
Q(t; x) &= \int_0^{\infty} f(x, y; t) dy = (\beta\, x) \sum_{k=1}^{\infty} \frac{1}{k} e^{-k\xi t} L_{k-1}^{(2)}(\beta x) \beta \int_0^{\infty} e^{-\beta\, y} L_{k-1}^{(2)}(\beta y) dy \\
&= (\beta\, x) \sum_{k=1}^{\infty} \frac{1}{k} e^{-k\xi t} L_{k-1}^{(2)}(\beta x). \qquad (8)
\end{aligned}$$

4 Main Result 2. The Probability Distribution of First Absorption Time

To derive the pdf $g(t; x)$ of the first absorption time into the state 0 given $Y(0) = x$, denote by $G(t; x)$ the absorption probability into 0 at or before time t given $Y(0) = x$. (Note that $Y(t) = \frac{1}{N} X(t)$.) Since 0 is the only absorbing state (persistent state),

$$G(t; x) = 1 - Q(t; x) = 1 - (\beta\, x) \sum_{k=1}^{\infty} \frac{1}{k} e^{-k\xi t} L_{k-1}^{(2)}(\beta x). \tag{9}$$

It follows that

$$
\begin{aligned}
g(t;x) &= \frac{\partial}{\partial t}G(t;x) = \frac{\partial}{\partial t}\{1-Q(t;x)\} \\
&= (\beta\, x)\xi \sum_{k=1}^{\infty} e^{-k\xi t} L_{k-1}^{(2)}(\beta x). \qquad (10)
\end{aligned}
$$

Notice that $\lim_{t\to\infty} Q(t;x) = 0$ for all $x > 0$ so that $\lim_{t\to\infty} G(t;x) = 1$ for all $x > 0$. This is equivalent to stating that with probability one the process will eventually be absorbed into the state 0 starting with $X(0) = i > 0$. The next theorem shows that $\int_0^\infty g(t;x)dt = 1$ for all $x > 0$ so that $g(t;x)$ is indeed a pdf for all given $x > 0$. For proving the next theorem, we first prove the following lemma.

Lemma 4.1 *If $\beta > 0$, then*

$$\psi_r(x) = (\beta\, x)\sum_{k=1}^{\infty}\frac{1}{k^r}L_{k-1}^{(2)}(\beta x) = 1 \textit{ for all } x > 0 \textit{ and for all } r \geq 1.$$

Proof

By the unicity theorem of Laplace transform (see Widder [7]), it suffices to show that

$$\int_0^\infty e^{-\beta x}\{\psi_r(x) - 1\}dx = 0.$$

Now

$$\int_0^\infty e^{-\beta x}dx = \beta^{-1},$$

and

$$
\begin{aligned}
\int_0^\infty e^{-x} x L_{k-1}^{(2)}(x)dx &= \frac{\Gamma(k-1+2)}{(k-1)!}\sum_{j=0}^{k-1}(-1)^j\binom{k-1}{j}\frac{1}{\Gamma(j+2)}\int_0^\infty x^{j+1}e^{-x}dx \\
&= k\sum_{j=0}^{k-1}(-1)^j\binom{k-1}{j} = \delta_{1k}.
\end{aligned}
$$

It follows that

$$
\begin{aligned}
\int_0^\infty e^{-\beta x}\psi_r(x)dx &= \sum_{k=1}^{\infty}\frac{1}{k^r}\int_0^\infty e^{-\beta\, x}(\beta\, x)L_{k-1}^{(2)}(\beta x)dx \\
&= \beta^{-1}\sum_{k=1}^{\infty}\frac{1}{k^r}\int_0^\infty e^{-x}x\, L_{k-1}^{(2)}(x)dx = \beta^{-1}\sum_{k=1}^{\infty}\frac{1}{k^r}\delta_{1k} = \beta^{-1}.
\end{aligned}
$$

Thus,

$$\int_0^\infty e^{-\beta x}\{\psi_r(x) - 1\}dx = 0 \text{ so that } \psi_r(x) = 1 \text{ for all } x > 0.\mathbf{Q.E.D}.$$

Theorem 4.2 *If $x > 0$, then*

$$\int_0^\infty g(t;x)dt = 1.$$

Proof

To prove Theorem (4.2), denote by $\int_0^\infty g(t;x)dt = \psi(x)$. Then

$$\psi(x) = (\beta\, x)\sum_{k=1}^{\infty} \frac{1}{k} L_{k-1}^{(2)}(\beta x).$$

From Lemma (4.1), $\psi(x) = 1$ for all $x > 0$. **Q.E.D.**

Using the result from Lemma (4.1), it is easy to show that the mean value of first absorption time given $Y(0) = x$ is $(\beta x)\xi^{-1}$ for all $x > 0$. Similarly, the variance of first absorption time given $Y(0) = x$ is $(\beta x)(1 - \beta x)\xi^{-2}$ for all $x > 0$.

5 Some Applications.

To illustrate, consider a large haploid population. Suppose that at the t_0-th generation, a mutant gene is introduced into the population; with no loss of generality we let $t_0 = 0$. Suppose further that each mutant gene produces j mutant genes with probability p_j $(j = 0, 1, \ldots, \infty)$ in the next generation independently of other genes. Let $X(t)$ be the number of the mutant gene at generation t. Then $X(t)$ is a branching process with progeny distribution $\{p_j, j = 0, 1, \ldots, \infty\}$.

To specify p_j, let the fitness (i.e. average number of progenies per generation) of the wild gene and the mutant be given by μ and $\mu(1 + v)$ $(\mu > 0)$ respectively. Let N be the population size. Then in the 0th generation, the frequency of the mutant is

$$\frac{(1+v)\mu}{(N-1+1+v)\mu} = \frac{1}{N}(1+v) + o((N)^{-1}) = p + o((N)^{-1})$$

for finite v, where $p = \frac{1}{N}(1+v)$. When N is sufficiently large, and if the mating is random, then to order of $o(N^{-1})$, the probability that there are j mutants in the next generation is

$$p_j = \binom{N}{j} p^j (1-p)^{N-j}.$$

Since $\lambda = Np = (1+v) + (N)o(N^{-1}) \to (1+v)$ as $N \to \infty$, when N is sufficiently large, $1+v$ is the average number of progenies of the a allele and

$$p_j \sim e^{-(1+v)} \frac{(1+v)^j}{j!}, j = 0, 1, 2, \ldots.$$

Since the new mutants are usually slightly disadvantageous when comparing with the wild genes, one may assume that $-v = \frac{\alpha}{N} + o(N^{-1}) > 0$ for some constant α. Then, by the fundamental theorem of branching process, the probability is one that the new mutant will eventually be extinct. The pdf of the probability distribution of first time absorption is then given by equation (10) with $\alpha = \xi$ in the previous section.

References

[1] Karlin, S. and Taylor, H.M. : *A First Course in Stochastic Processes. Second edition. New York: Academic Press, 1975.*

[2] Jorde, L.B., Carey, J.C., Bamshad, M.J. and White, R.J.: *Medical genetics. Second Edition. New York: Moshy Inc., 1999*

[3] Scriver, C.R., Beaudet, A.I., Sly, W.S. and Valle, D. (eds.) *The Metabolic and Molecular Basis of Inherited Diseases. Vol. 3. New York: McGraw-Hill, 1995.*

[4] Feller, W.: *Diffusion process in genetics. Proc. 2nd Berkeley Symposium Math. Statist. Probab.,University of California Press, Berkeley, CA. USA.*, 1951, p227-246.

[5] Tan, W.Y.: *Stochastic Models With Applications to Genetics, Cancers, AIDS and Other Biomedical Systems. Singapore and River Edge, New Jersey: World Scientific, 2002.*

[6] Tan, W.Y. and Tiku, M.: *Sampling Distributions in Terms of Laguerre Polynomials and Applications. New Delhi, India: New Age International Publisher,1999.*

[7] Widder, P.: *The Laplace Transform. Princeton: Princeton University Press, 1946.*

Received: March, 2011