Abdul-Rehim EA

Department of Mathematics and Computer Science, Suez Canal University, Egypt

Submission: March 10, 2017; Published: May 08, 2017

*Corresponding author: Abdul-Rehim EA, Department of Mathematics and Computer Science, Suez Canal University, Egypt, Email: entsarabdolrehim@gmail.com

Abstract

In this review article, I prove mathematically that the diffusion process in genetics is a stochastic process and is a reversible process. MSC 2010: Primary 26A33, Secondary 45K05, 60J60, 44A10, 42A38, 65N06, 47G30, 80-99 Key Words and Phrases: exponential diffusion in genetic, stochastic process, Markov chain, reversibility property.

Introduction

Feller [1] was the first scientist introducing the still open problem “diffusion process in genetics”. R. A. Fisher and Sewall Wright studied many problems in this field. Most of these problems were based on the simple branching process; see [2] and [3]. R. A. Fisher used this branching process to describe the simplest possible populations in which there is no attraction among the possible generations. There are many difficulties arise when there is interaction between individuals or the population consists of different types of individuals being provided by various breeding. These simplest assumptions lead to Markov process.

In this review article, I am interested in give a survey about the relation between the diffusion in genetic process and Markov chain. I am interested also in proving that this stochastic process is reversible.

The Genetic Diffusion Processes

The branching process is a class of Markov chains. In this process, one considers a population which consists of individuals being able to produce offsprings of the same kind. By supposing that each individuals will be by the end of his lifetime produce j new offspring with 0,1,2,…, with probabilities $p_0, p_1, p_2, \ldots$ independently of all the other offsprings. Naturally $\sum_k p_k = 1$. Let $Z^{(n)}$ be a random variable of a Markov chain and suppose that $p_{jk}$ be the transition probabilities to the next generation $Z^{(n+1)} = k$. It is not surprise that it was difficult to find an explicit expression for $p_{jk}$. It definitely that these transition probabilities satisfy the Chapman-Kolmogrov [4] equation

$$p^{(n+1)}_{jk} - \sum_{\nu \neq j} p^{(n)}_{\nu k} = 0$$

The simple random mating is a good example to this approach in which one has amount of 2N genes. This number of genes is formed in 2N independent trials. For example suppose you have a pair of genes a and A. Each individual belongs to one of the three genotypes (a,a), (a,A) and (A,A). If the parent population consists of j of a- genes then there are 2N- j for A-genes. At each trials the genes a or A happen with the probabilities

$$p_j = \left( \begin{array}{c} 2N \\k\ \end{array} \right) p^j q^{2N-k}$$

That leads to the a transition probability having the binomial distribution

This type of treatment is similar to transition of the weight and black balls between the two urns of the Ehrenfest model.

The Derivation of the Genetic Diffusion Equation

Let $Y(t)$ is a random variable such as gene frequency or population size depending on the time variable t. As described above $Y(t)$ is a stochastic process of a Markov type. In the case of a finite chain, the process depends on its initial value at $t = 0$. Define the function $p(t; y, x)$ as the conditional probability density that $Y(t + t_0) = x$ given that $Y(t_0) = y$ at the fixed time $t_0$. Then the expected value of $Y(t)$ at a later time $t_0 + t$ is

$$E[Y(t_0 + t)|Y(t_0) = x] = \int_{Y(t_0) = x}^{} p(t; y, x) dx$$

The average value of the increase (the so called mean of displacement) is

$$E[\Delta Y(t) = \int_{t_0}^{t} p(T; y, x) dx]$$

and its variance is

$$Var[\Delta Y(t_0)] = \int_{t_0}^{t} p(T; y, x) dx$$
Taking the limit as $\tau \to 0$, one gets

\[
\lim_{\tau \to 0} \frac{\lambda_{\tau}(x)}{\tau} = 2b(x) \quad \text{and} \quad \lim_{\tau \to 0} \frac{p_{\tau}(x,y)}{\tau} = a(x).
\]

Since the stochastic variable is of Markov type then where $\beta$ is the rate of increasing. In other words the population increases exponentially. It was shown by Kolmogorov that the probability density $u(x, t)$ of the stochastic variable $Y(t)$ satisfies the diffusion equation

\[
\frac{\partial u(x,t)}{\partial t} + a(x) \frac{\partial u(x,t)}{\partial x} + \frac{\partial^2 u(x,t)}{\partial x^2} = 0, \quad x \to h, \quad x \to 0, \quad t \to 0, \quad (1)
\]

where the coefficients $a(x)$ and $b(x)$ depend on the choice of the genetic model. It worth to say that this equation is considered as a special form of the Fokker-Planck equation which has a huge applications on all various scientific fields.

### The exponential genetic diffusion equation

According to W. Feller [1] the general exponential equation reads

\[
\frac{d\lambda(x)}{dx} - \frac{d^2\lambda(x)}{dx^2} = 0, \quad x \to 1, \quad x \to 0, \quad (2)
\]

where $u(x, t)$ is the probability density of the gene frequency at time $t$ with the initial condition $u(x, 0) = x$ and is subject to the boundary condition $u(0, t) = u(1, t) = 0$ if $T_n$ is the time of the $n$-th generation. Here $\alpha$ and $\beta$ are constants. $\alpha$ is the diffusion constants and $\beta$ is drift constant and takes the values $\{1,0,-1\}$ respectively. Suppose here $\alpha = 1$ and $\beta = 1$ and apply the common finite difference rule to descretize equation (2). To do so, let $t = n$ with $0 < \tau << 1$ and $x = j\tau$ with $0 < h << 1, 0 < j < N$ where $N$ is the number of diffused genes. The discrete difference scheme reads

\[
j_{j+1}^{(n+1)} - 2j_{j}^{(n)} + j_{j-1}^{(n)} = \frac{2\beta_{j}^{(n)}}{h^2}, \quad (3)
\]

Define the scaling relation as $\frac{\tau}{\tau}$, put $j_{j}^{(n)} = \frac{1}{\tau}$ and solve for $j_{j}^{(n+1)}$, to get

\[
j_{j}^{(n+1)} = (1 - 2\mu)j_{j}^{(n)} + \mu (j+1)l_{j}^{(n+1)} + \mu (j-1)l_{j-1}^{(n)}, \quad (4)
\]

the only condition imposed on $\mu$ in this scheme is in order to have a stable scheme. This equation can be written in the following form

\[
y_{j}^{(n+1)} = p_{j}^{(n)} y_{j+1}^{(n+1)} + p_{j+1}^{(n)} y_{j+1}^{(n+1)} + p_{j}^{(n)} y_{j-1}^{(n+1)} + p_{j-1}^{(n)} y_{j-1}^{(n+1)}, \quad (5)
\]

where $p_j$, $P_{n+1}$, $P_{n}$, are the transition probabilities from the point $x(t_n)$ to the points $x(t_n + 1), x_n - 1(t_n + 1)$, and $x_n + 1(t_n + 1)$ respectively. As discrete transition probabilities $P_{n+1} P_{n} P_{n} = 1$.

Constitute the column vector $y^{(n)} = (y_{j}^{(n)})_{j=1}^{N}$ the equation (4) can be written in matrix form

\[
y^{(n+1)} = P_{n+1} y^{(n)}, \quad (6)
\]

the elements of the matrix $P$ are the above transition probabilities. Then all the rows of the matrix $P$ are summed to one and it really represents the transition matrix of a Markov chain. Then the matrix $P$ is a stochastic matrix. According to theory of stochastic process, especially Ehrenfest model, this matrix has the stationary probability distribution vector $\pi$ whose its elements are the binomial $\sum \binom{N}{j} \pi_{j} = 0, 1, 2, \ldots, N$. In what follows, I prove that the diffusion in genetics is reversible. Firstly, the definition of the reversible process as is stated at the book of Kelly [5] is as follows: A stochastic process $Y(t)$ is said to be reversible if $Y(t_1), Y(t_2), \ldots, Y(t_n)$ has the same distribution as $Y(t_n - t_1), Y(t_n - t_2), \ldots, Y(t_n - t_n)$, for all $t_1, t_2, \ldots, t_n$, see [5]. The balance condition reads

\[
\sum_{j} \pi_{j} p_{j,k} = \sum_{k} \pi_{k} p_{k,j} = 1, j, k \in N,
\]

where $\pi = P_{n} \pi$. Our Matrix $P$ and its stationary solution $\pi$ satisfy this balance equation, then the studied process is reversible.

### References

1. Feller W Diffusion Process in Genetics. Naval Research at Cornell University for developing probability theory, pp. 227-246.
2. Fisher RA (1930) The Genetical Theory of Natural Selection. Oxford University Press, UK, p. 308.
3. Wright S (1939) Statistical genetics in relation to evolution, Scientific and Industrial News, Hermann & Cie, Paris, France.
4. Kolmogorov (1931) On the analytical methods of the probability calculation. Math Analen 94: 397-381.
5. Kelly FP (1979) Reversibility and Stochastic Networks, Wiley series in Probability and Mathematical Statistics. In: John Wiley & Sons (Ed.), Chichester, New York, USA.