J. Appl. Math. & Informatics Vol. 31(2013), No. 1 - 2, pp. 241 - 246
Website: http://www.kcam.biz

ON THE MARTINGALE PROPERTY OF LIMITING DIFFUSION IN SPECIAL DIPLOID MODEL†

WON CHOI

Abstract. Choi [1] identified and characterized the limiting diffusion of this diploid model by defining discrete generator for the rescaled Markov chain. In this note, we define the operator of projection S_t on limiting diffusion and new measure $dQ = S_t dP$. We show the martingale property on this operator and measure. Also we conclude that the martingale problem for diffusion operator of projection is well-posed.

AMS Mathematics Subject Classification : 92D10, 60H30, 60G44.
Key words and phrases : diploid model, limiting diffusion, martingale problem.

1. Introduction

We begin by formulating a Wright-Fisher model that is general enough to include the normal selection model. Let E(a locally compact separable metric space) be the set of all possible allels and ν_0(in $\mathcal{P}(E)$, the set of Borel probability measures on E) the distribution of the type of a new mutant. Suppose that N(a positive integer) is the diploid population size and $s(\mathbf{x})$ is the selection coefficient of allele $\mathbf{x}$.

The Wright-Fisher model is a Markov chain describing the evolution of the composition of the population of gameters $(x_1, x_2, \cdots, x_{2N})$ in E^{2N} or, since the order of the gameter is unimportant,

$$\frac{1}{2N}\sum_{i=1}^{2N}\delta_{x_i} \in \mathcal{P}(E)$$

(Here $\delta_x \in \mathcal{P}(E)$ is the unit mass at $x \in E$).

We now consider the normal-selection model in this note. The type space E is unspecified. However, ν_0 and the function s must jointly satisfy the following

Received August 21, 2012. Accepted October 16, 2012.
†This research was supported by University of Incheon Research Grant, 2012-2013

condition; If X is a random variable with distribution ν_0, then $s(X)$ has the normal distribution with mean 0 and variance σ^2. Furthermore, $\sigma = \sigma_0/2N$ for an appropriate constant σ_0. There are therefore a number of possible choice for E, ν_0, and s, including;

$$E = (0,1),\ \nu_0 = U(0,1),\ s(\mathbf{x}) = \sigma\mathbf{\Phi}^{-1}(\mathbf{x}),$$

where Φ is the standard normal distribution function,

$$E = R,\ \nu_0 = N(0,\sigma^2),\ s(\mathbf{x}) = \mathbf{x},$$

and

$$E = R,\ \nu_0 = N(0,\sigma_0^2),\ s(\mathbf{x}) = \mathbf{x}/2N.$$

For each positive integer M, let ω_M be a positive, symmetric, bounded, Borel function on E^2, let $R_M((p,q), dx \times dy)$ be a one-step transition function on $E^2 \times \mathcal{B}(E^2)$ satisfying

$$R_M((p,q), dx \times dy) = R_M((q,p), dy \times dx),$$

and $Q_M(p, dx)$ be a one-step transition function on $E \times \mathcal{B}(E)$.

Let N be the diploid population size. We consider $M = 2N$ gametes and the mapping $\eta_M\ :\ E^M \to \mathcal{P}(E)$ by letting

$$\eta_M(p_1, p_2, \cdots, p_M) = \frac{1}{M}(\delta_{p_1} + \delta_{p_2} + \cdots + \delta_{p_M}).$$

Here $\delta_p \in \mathcal{P}(S)$ denotes the unit mass at $p \in S$. The state space for this model is

$$\mathcal{K}_M(E) = \eta_M(E^M).$$

Given $\mu \in \mathcal{P}(E)$, we define $\mu_1 \in \mathcal{P}(E^2)$ and $\mu_2, \mu_3 \in \mathcal{P}(E)$ by

$$\mu_1(dp \times dq) = \omega_M(p,q)\mu^2(dp \times dq)/\langle \omega_M, \mu^2\rangle, \tag{1}$$

$$\mu_2(dx) = \int_{E^2} R_M((p,q), dx \times E)\mu_1(dp \times dq), \tag{2}$$

$$\mu_3(dx) = \int_E Q_M(p, dx)\mu_2(dp). \tag{3}$$

The Markov chain has one-step transition function $P_M(\mu, d\theta)$ on $\mathcal{K}_M(E) \times (\mathcal{K}_M(E))$ defined by

$$P_M(\mu, \cdot) = \int_{E^M} (\mu_3)^M (dp_1 \times dp_2 \times \cdots \times dp_M)\delta_{\eta_M(p_1,p_2,\cdots,p_M)}(\cdot).$$

Choi [1] identified and characterized the limiting diffusion of this diploid model by defining discrete generator for the rescaled Markov chain. In this note, we define the operator of projection S_t on limiting diffusion and new measure $dQ = S_t dP$. We show the martingale property on this operator and measure. Also we conclude that the martingale problem for diffusion operator of projection is well-posed.

2. Main Results

In order to consider a limiting diffusion, we define the discrete generator $\mathcal{L}_M$ for the M-the rescaled Markov chain and canonical coordinate process $\{\rho_t, t \geq 0\}$:

$$(\mathcal{L}_M\phi)(\rho_t) = M \int_{\mathcal{P}_M} (\phi(\nu_t) - \phi(\rho_t)) P_M(\rho_t, \nu_t)$$

where $\mathcal{P}_M$ is given in the diploid models as described above.

We restrict our attention to test functions θ of the form

$$\theta(\nu_t) = \beta_1\langle f_1, \nu_t\rangle \cdots \beta_k\langle f_k, \nu_t\rangle, \ \ \theta(\rho_t) = \langle f_1, \rho_t\rangle \cdots \langle f_k, \rho_t\rangle$$

where $f_1, \cdots, f_k \in \mathcal{B}(E)$ and $\{\beta_i\}$ is non-negative constant satisfying that $\sup_i \beta_i < +\infty$. Assume that "mutation or gene conversion rate" is

$$\sum_{k\in S} \beta_k\langle f_i, \rho_t\rangle - \beta_i - \beta_j \ \text{ for every } \ i < j,$$

in the diploid models as described above. This means that mutations or gene conversions occur with particular rate in case of $i < j$ (see Choi [2]).

Choi [1] proved that there exist $a_{f_i,f_j}, \ b_{f_i} \in \mathcal{B}(\mathcal{P}(S))$ such that

$$(\mathcal{L}_M\theta)(\rho_t) \ \rightarrow \ (\mathcal{L}_\pi\theta)(\rho_t) \ \text{ as } \ M \rightarrow \infty$$

uniformly in $\rho_t \in \mathcal{K}_M(S)$, where

$$(\mathcal{L}_\pi\theta)(\rho_t) = \sum_{1\leq i<j\leq k} a_{f_i,f_j} \prod_{l \,:\, l\neq i,j} \langle f_l, \rho_t\rangle + \sum_{i=1}^{k} b_{f_i} \prod_{l \,:\, l\neq i} \langle f_l, \rho_t\rangle.$$

We start with;

Lemma 1. Suppose the conditions (1), (2) and (3) are satisfied and θ have the form

$$\theta(\rho_t) = F(\langle f_1, \rho_t\rangle, \langle f_2, \rho_t\rangle, \cdots, \langle f_k, \rho_t\rangle) = F(\langle \mathbf{f}, \rho_t\rangle)$$

where $F \in C^2(\mathbf{R}^k)$. Then there exist $a_{f_i,f_j}, \ b_{f_i} \in \mathcal{B}(\mathcal{P}(E))$ such that

$$\lim_{M \,\rightarrow\, \infty} (\mathcal{L}_M\theta)(\rho_t) = \sum_{1\leq i<j\leq k} a_{f_i,f_j} F_{z_i z_j}(\langle \mathbf{f}, \rho_t\rangle) + \sum_{i=1}^{k} b_{f_i} F_{z_i}(\langle \mathbf{f}, \rho_t\rangle)$$

uniformly in $\rho_t \in \mathcal{K}_M(E)$, where F_{z_i} and $F_{z_i z_j}$ mean the partial derivative with respect to i and i, j, respectively.

Proof. If we let

$$\begin{aligned}
\theta(\rho_t) &= F(\langle f_1, \rho_t\rangle, \langle f_2, \rho_t\rangle, \cdots, \langle f_k, \rho_t\rangle) = F(\langle \mathbf{f}, \rho_t\rangle) \\
a_{f_i,f_j} &= \beta_i\langle f_i f_j, \rho_t\rangle - \langle f_i, \rho_t\rangle\langle f_j, \rho_t\rangle(\sum_{k\in S} \beta_k\langle f_i, \rho_t\rangle - \beta_i - \beta_j) \\
b_{f_i} &= \langle A f_i, \rho_t\rangle + \langle B f_i, \rho_t^2\rangle + \langle (f_i \circ \pi)\sigma, \rho_t^2\rangle - \langle f_i, \rho_t\rangle\langle \sigma, \rho_t^2\rangle,
\end{aligned}$$

for selection function σ on E^2 and π is the projection of E^2, we have

$$(\mathcal{L}_M\theta)(\rho_t) = (\mathcal{L}_\pi\theta)(\rho_t) + o(1)$$

uniformly and this result is immediate from a second order Taylor expansion with the result of Choi [1] □

Lemma 2. Let X be a progressively measurable process and f, g, c be bounded Borel functions. If

$$f(X(t)) - \int_0^t g(X(s))ds$$

is $\{\mathcal{F}_t^X\}$-martingale, then

$$f(X(t))e^{-\int_0^t c(X(s))ds} - \int_0^t \{g(X(s)) - c(X(s))f(X(s))\}e^{-\int_0^s c(X(r))dr}ds$$

is $\{\mathcal{F}_t^X\}$-martingale.

Proof. See Ethier and Kurtz [3] □

Let π_x, π_y be the projection of E^2 on x and y-coordinate, respectively and $P \in \mathcal{P}(E)$ be a solution of the martingale problem for $\mathcal{L}_{\pi_x}$. Define

$$S_t = \exp\{\langle \pi_y, \rho_t\rangle - \langle \pi_y, \rho_0\rangle - \int_0^t e^{-\langle \pi_y . \rho_s\rangle}\mathcal{L}_{\pi_x}e^{\langle \pi_y, \rho_s\rangle}ds\}$$

and measure Q by

$$dQ = S_t dP.$$

Then S_t is a mean one martingale and we have;

Theorem 3.

$$E^Q\left[\left(\theta(\rho_t) - \theta(\rho_s) - \int_s^t (\mathcal{L}_{\pi_x+\pi_y}\theta)(\rho_r)dr\right)H_s\right] = 0$$

for $\theta \in \mathcal{D}(\mathcal{L}_{\pi_x+\pi_y})$ and bounded measurable function H_s.

Proof. Since

$$E^Q\left[\left(\theta(\rho_t) - \theta(\rho_s) - \int_s^t (\mathcal{L}_{\pi_x+\pi_y}\theta)(\rho_r)dr\right)H_s\right]$$
$$= E^p[\theta(\rho_t)S_tH_s] - E^P[\theta(\rho_s)S_sH_s] - \int_s^t E^P[(\mathcal{L}_{\pi_x+\pi_y}\theta)(\rho_r)S_rH_s]dr,$$

we have

$$E^Q\left[\left(\theta(\rho_t) - \theta(\rho_s) - \int_s^t (\mathcal{L}_{\pi_x+\pi_y}\theta)(\rho_r)dr\right)H_s\right]$$
$$= E^P[e^{-\langle \pi_y, \rho_0\rangle}\{\theta(\rho_t)e^{\langle \pi_y, \rho_t\rangle}\exp\left(-\int_0^t e^{-\langle \pi_y . \rho_r\rangle}\mathcal{L}_{\pi_x}e^{\langle \pi_y, \rho_r\rangle}dr\right)$$

$$-\theta(\rho_s)e^{\langle\pi_y,\rho_s\rangle}\exp\left(-\int_0^s e^{-\langle\pi_y.\rho_r\rangle}\mathcal{L}_{\pi_x}e^{\langle\pi_y,\rho_r\rangle}dr\right)$$

$$-\int_s^t(\mathcal{L}_{\pi_x+\pi_y}\theta)(\rho_r)e^{\langle\pi_y,\rho_r\rangle}\exp\left(-\int_0^r e^{-\langle\pi_y.\rho_x\rangle}\mathcal{L}_{\pi_x}e^{\langle\pi_y,\rho_x\rangle}dx\}\right)drH_s].$$

But

$$\begin{aligned}&\mathcal{L}_{\pi_x}[\theta(\rho)e^{\langle\pi_y,\rho\rangle}]-\theta(\rho)\mathcal{L}_{\pi_x}e^{\langle\pi_y,\rho\rangle}\\ &\quad=(\mathcal{L}_{\pi_x}\theta)(\rho)e^{\langle\pi_y,\rho\rangle}+\sum_{i=1}^k(\langle f_i\pi_y,\rho\rangle-\langle f_i,\rho\rangle\langle\pi_y,\rho\rangle)F_{z_i}(\langle\mathbf{f},\rho\rangle)e^{\langle\pi_y,\rho\rangle}\\ &\quad=(\mathcal{L}_{\pi_x+\pi_y}\theta)(\rho)e^{\langle\pi_y,\rho\rangle}.\end{aligned}$$

Therefore, from Lemma 2, we have

$$\begin{aligned}&E^Q\left[\left(\theta(\rho_t)-\theta(\rho_s)-\int_s^t(\mathcal{L}_{\pi_x+\pi_y}\theta)(\rho_r)dr\right)H_s\right]\\ &\quad=E^P[e^{-\langle\pi_y,\rho_0\rangle}\{\theta(\rho_t)e^{\langle\pi_y,\rho_t\rangle}\exp\left(-\int_0^t e^{-\langle\pi_y.\rho_r\rangle}\mathcal{L}_{\pi_x}e^{\langle\pi_y,\rho_r\rangle}dr\right)\\ &\qquad-\theta(\rho_s)e^{\langle\pi_y,\rho_s\rangle}\exp\left(-\int_0^s e^{-\langle\pi_y.\rho_r\rangle}\mathcal{L}_{\pi_x}e^{\langle\pi_y,\rho_r\rangle}dr\right)\\ &\qquad-\int_s^t(\mathcal{L}_{\pi_x}[\theta(\rho_r)e^{\langle\pi_y,\rho_r\rangle}]-\theta(\rho_r)\mathcal{L}_{\pi_x}e^{\langle\pi_y,\rho_r\rangle})\\ &\qquad\exp\left(-\int_0^r e^{-\langle\pi_y.\rho_x\rangle}\mathcal{L}_{\pi_x}e^{\langle\pi_y,\rho_x\rangle}dx\right)dr\}H_s]=0.\end{aligned}$$

□

Theorem 3 allows us to define Q by

$$dQ=S_tdP.$$

We show that Q solve the martingale problem for $\mathcal{L}_{\pi_x+\pi_y}$.

Corollary 4. The measure Q is a solution of martingale problem for $\mathcal{L}_{\pi_x+\pi_y}$.

Proof. Since S_t is mean-one martingale and

$$E^Q\left[\left(\theta(\rho_t)-\theta(\rho_s)-\int_s^t(\mathcal{L}_{\pi_x+\pi_y}\theta)(\rho_r)dr\right)H_s\right]=0$$

for $\theta\in\mathcal{D}(\mathcal{L}_{\pi_x+\pi_y})$ and bounded measurable function H_s, this result follows directly from the definition of martingale problem. □

We conclude with;

Corollary 5. The martingale problems for $\mathcal{L}_{\pi_x}$ and $\mathcal{L}_{\pi_y}$ are well-posed.

Proof. Apply Corollary 4 with $\pi_x=0$ or $\pi_y=0$ using the fact that existence and uniqueness are known for $\mathcal{L}_0$. □

Remark. Let P_ρ be the unique solution of the martingale problem for $\mathcal{L}_0$ starting at ρ and $\pi_1, \pi_2, \cdots, \pi_n$ the projection of E^n on $x_1, x_2, \cdots, x_n$. Then we can define

$$dQ_\rho = S_t dP_\rho$$

and show that Q_ρ solves the martingale problem for $\mathcal{L}_\pi$ starting at ρ.

References

1. W.Choi, *On the limiting diffusion of special diploid model in population genetics*, Bulletin of the Korean Mathematical Society, Vol 42, No 2 (2005), 397-404.
2. W.Choi and D.H.Choi *The application of stochastic differential equation to population genetics*, Bulletin of the Korean Mathematical Society, Vol 40, No 4 (2003), 677-683.
3. S.N.Ethier and T.G.Kurtz, *Markov processes: Characterization and Convergence*, Wiley, New York (1993).

Won Choi received his Ph.D at Sung Kyun Kwan University under the direction of Yeong Don Kim. Since 1993 he has been at the University of Incheon. In 1996 and 2003, he was in Steklov Mathematical Institute of Russia and University of Iowa, respectively. His research interests center on stochastic processes and Bio-mathematics.

Department of Mathematics, University of Incheon, Incheon 406-772, Korea
e-mail:choiwon@incheon.ac.kr