J. Appl. Math. & Informatics Vol. **28**(2010), No. 3 - 4, pp. 1003 - 1008
Website: http://www.kcam.biz

ON THE MARTINGALE PROBLEM AND SYMMETRIC DIFFUSION IN POPULATION GENETICS

WON CHOI AND YOO JUNG JOUNG

ABSTRACT. In allelic model $X = (x_1, x_2, \cdots, x_d)$,

$$M_f(t) = f(p(t)) - \int_0^t Lf(p(t))ds$$

is a P-martingale for diffusion operator L under the certain conditions. In this note, we define

$$T_t f = E_{P_0}^{P^*}[f((P(t))] \quad for \quad t \geq 0$$

for using a new diffusion operator L^* and we show the diffusion relations between T_t and diffusion operator L^*.

AMS Mathematics Subject Classification: 92D25,60H30,47A05
Key words and phrases: Countable allelic model, martingale problem.

1. Introduction

Consider n locus model

$$X = (x_1, x_2, \cdots, x_d) \in R^d,$$

so we find n genes on a chromosome. A partition X describes a state of a chromosome and X means that there exist d kinds of alleles which occupy x_1 loci, x_2 loci, $\cdots$, x_d loci. If the partition X has α_i parts equal to i, then X describes that there exists α_i kinds of alleles occurring i loci for each i. Let q_{ij} denote "mutation rate" or "gene conversion rate" from a partition X_i to another partition X_j per generation measured on the t time scale and p_i denotes the frequency of chromosome of type X_i.

Let S be a countable set. In population genetics theory we often encounter diffusion process on the domain

$$K = \left\{ p = (p_i)_{i \in S} \, ; \, p_i \geq 0, \, \sum_{i \in S} p_i = 1 \right\}$$

Received November 10, 2009. Revised January 28, 2010. Accepted February 17, 2010.
This research was supported by University of Incheon Research Grant, 2009-2010

We suppose that the vector $p(t) = (p_1, p_2, \cdots)$ of gene frequencies varies with time t.

Let L be a second order differential operator on K

$$L = \sum_{i,j\in S} a_{ij}(p)\frac{\partial^2}{\partial p_i \partial p_j} + \sum_{i\in S} b_i(p)\frac{\partial}{\partial p_i}$$

with domain $C^2(K)$, where $\{a_{ij}\}$ is a real symmetric and non-negative definite matrix defined on K and $\{b_i\}$ is an measurable function defined on K.

We assume that $\{a_{ij}\}$ and $\{b_i\}$ are continuous on K. Let $\Omega = C([0,\infty) : K)$ be the space of all K-valued continuous function defined on $[0,\infty)$. A probability P on $(\Omega, \mathcal{F})$ is called a solution of the (K, L, p)-*martingale problem* if it satisfies the following conditions,

(1) $P(p(0) = p) = 1$.

(2) denoting $M_f(t) = f(p(t)) - \int_0^t Lf(p(t))ds$, $(M_f(t), \mathcal{F}_t)$ is a P-martingale for each $f \in C^2(K)$.

The diffusion operator L was first introduced by Gillespie([6]) in case that the partition consists of two points. In this case, L is a one-dimensional diffusion operator. However, the uniqueness of solutions of the (K, L, p)-martingale problem has not been generally established. For this problem, Either([4]) proved that if $\{a_{ij}(p)\} = \{p_i(\delta_{ij} - p_j)\}$ for Kronecker symbol δ_{ij} and $\{b_i(p)\}$ are C^4-functions satisfying a certain condition, then the uniqueness of the (K, L, p)-martingale problem holds. Also, Okada([3]) showed that the uniqueness holds for a rather general class in two dimension. In case that L reduces to an infinite allelic diffusion model of the Wright-Fisher type, Either([5]) gave a partial result. Choi([1]) defined new symmetric preserving transformation. He proved Uniqueness for martingale problem and symmetric property.

In this note, we define

$$T_t f = E_{P_0}^{P^*}\left[f(P(t))\right] \quad for \quad t \geq 0$$

for using a new diffusion operator L^* and we show the diffusion relations between T_t and diffusion operator L^*.

2. The main results

Definition. A sequence $\{X_1, X_2, \cdots, X_J,\}$ of partitions is called (X_1, X_J)-*chain* if X_{j+1} is a consequent of X_j by mutation or gene conversion for each $j = 1, 2, \cdots, J-1$.

The value

$$\left(\frac{q_{12}}{q_{21}}\right)\left(\frac{q_{23}}{q_{32}}\right)\cdots\left(\frac{q_{J-1\ J}}{q_{J\ J-1}}\right)\cdots$$

does not depend on the choice of (X_1, X_K)-chain.

Let X be any partition of n and let $\{X_1, X_2, \cdots, X_i, \cdots\}$ be a $((n), X_i)$-chain. Put

$$P_i = \prod_{j=1}^{J-1}\left(\frac{q_{j\ j+1}}{q_{j+1\ j}}\right), \quad P_{(n)} = 1.$$

Let

$$K^* = \left\{P = (P_i)_{i\in S} \ :\ \sum_{i\in S} P_i < +\infty\right\}.$$

Lemma 1. *Let L^* be a second order differential operator on K^**

$$L^* = \sum_{i,j\in S} \tilde{a}_{ij}(P)\frac{\partial^2}{\partial P_i \partial P_j} + \sum_{i\in S} \tilde{b}_i(P)\frac{\partial}{\partial P_i}$$

where

$$\tilde{a}_{ij} = \begin{cases} (\textit{number of elements } S) \times \sqrt{\beta_i\beta_j P_i(t)P_j(t)} & \textit{if } S \textit{ is finite} \\ 0 & \textit{if } S \textit{ is infinite.}\end{cases}$$

Then the uniqueness of solution for the (K^*, L^*, P_0)-martingale problem holds.

Proof. It is well-known that to show the existence and uniqueness of solutions for the (K^*, L^*, P_0)-martingale problem is equivalent to show that the stochastic differential equation has a unique solution. Therefore this result follows from Choi([1]). □

We say that the probability measure P^* is a minimal solution to the martingale problem for L^* starting from P_0 (abbreviated by $P^* \sim L^*$ at P_0)

Suppose that there is a unique $P^*_{P_0} \sim L^*$ at P_0. Then the family $P^*_{p_0}$ is measurable, where $P^*_\triangle$ is determined by the equation

$$P^*_\triangle\Big(p(t) = \triangle, t \geq 0\Big) = 1.$$

Define

$$T_t f = E^{P^*}_{P_0}\Big[f(P(t))\Big] \quad \textit{for} \quad t \geq 0$$

Then we have :

Theorem 2. *Suppose $P^*_{P_0} \sim L^*$. Then for each $1 \leq p < \infty$,*

$$\|T_t f\|_{L^p(m)} \leq \|f\|_{L^p(m)} \quad \textit{and} \quad \int fT_t g dm = \int gT_t f dm. \tag{1}$$

In order to prove Theorem1. we can call on the machinery on Fukushima([2]) to construct a strong Markov family Q_x of probability measure such that

(1) $Q_x(P(0) = P_0) = 1$ and $Q_\triangle(P(t) = \triangle) = 1$

(2) $Q_x(P(t) \neq \triangle) = 0$.

(3) if $Q_x f(p(t)) = E^{Q_x}[f(P(t))]$ for $t \geq 0$

then $\Big\|Q_t f\Big\|_{L^2(m)} \leq \Big\|f\Big\|_{L^2(m)}$ and $\int fQ_t g dm = \int gQ_t f dm$

$$(4) \quad \int gQ_t f dm - \int gf dm = \int_0^t \Big(\int gQ_s L^* f dm \Big) ds.$$

In order for us to use the existence of Q_x to gain information about P^*, we need to connect Q_x with the martingale problem for L^*.

Theorem 3. *Let Q_x satisfy condition (1) $\sim$ (4) above and let g be nonnegative. Set $Q_g = \int Q_x g(P(t))m(dt)$. Then $M_f^*(t)$ is a Q_g -martingale. In particular, under the hypothesis of Theorem 2,*

$$Q_g = \int P^*_{P_0} g(P(t))m(dt).$$

Proof. Suppose Φ is a bounded nonnegative measurable function satisfying

$$E^{P^*_g}(\Phi) = 1.$$

Set $\mu(\Gamma) = E^{Q_g}[\Phi, P(s) \in \Gamma]$. Then $\mu \ll m$ *and* $h = \dfrac{d\mu}{dm}$ is nonnegative. To see this, note that in general

$$\left| \int \psi(x)\mu(dx) \right| = \left| E^{Q_g}[\Phi\psi(x(s))] \right| \leq \| \Phi \| \left| E^{Q_g}[\psi(x(s))] \right| = \| \Phi \| \left| \int gQ_s \psi dm \right|$$

$$= \| \Phi \| \left| \int \psi Q_s g dm \right| \quad \leq \| \Phi \| \| g \| \| \psi \|_{L^1(m)}$$

Hence $\mu \ll m$ *and if* $h = \dfrac{d\mu}{dm}$, *then*

$$\| h \| \leq \| \Phi \| \| g \| .$$

Using the Markov property, we now see that for Φ 's of the sort described above,

$$E^{Q_g}\Big[f(P(t+s)) - f(P(s))\Phi \Big] = E^{Q_h}\Big[f(P(t)) - f(P(0)) \Big]$$

and

$$E^{Q_g}\left[\Phi \int_s^{s+t} L^* f(P(u)) d\mu \right] = E^{Q_g}\left[\Phi E^{Q_{x(s)}} \int_0^t L^* f(P(u)) d\mu \right]$$

$$= E^{Q_h}\left[\int_0^t L^* f(P(u)) d\mu \right]$$

On the other hand

$$E^{Q_h}[f(P(t)) - f(P(0))] = \int hQ_t f dm - \int hf dm = \int_0^t \Big(\int hQ_u L^* f dm \Big) d\mu$$

$$= E^{Q_h}\left[\int_0^t L^* f(P(u)) d\mu \right]$$

Thus,

$$E^{Q_g}[f(P(t+s)) - f(P(s))\Phi] = E^{Q_g}\left[\Big(\int_s^{t+s} L^* f(P(u)) d\mu \Big) \Phi \right]$$

for Φ 's of the special sort with which we have been dealing. It is now an easy step to derive

$$E^{Q_g}\Big[f(P(t+s)) - f(P(s)) \mid \sigma(P(s):0\leq s\leq t)\Big]$$

$$= E^{Q_g}\left[\int_s^{s+t} L^* f(P(u))d\mu \mid \sigma(P(s):0\leq s\leq t)\right]$$

and therefore $M_f^*(t)$ is a Q_g -martingale. To complete the theorem, note that by , if $\Gamma = \Big\{P_0; M_f^*(t) \text{ is a } Q_g\text{-martingale for all } f\Big\}$, then

$$\int_{E|\Gamma} g(x)m(dx) = 0.$$

But then, under the hypothesis of Theorem 2,

$$Q_x = P_{P_0}^* \quad for \quad P_0 \in P.$$

Hence

$$Q_g = \int P_{P_0}^* g(P(t))m(dt)$$

□

We prepare to prove Theorem 2;

(*Proof of Theorem 2*)

Given nonnegative f, g , we have

$$\int gT_t f dm = \int E^{P_x}[f(P(t))]g(P(t))m(dt) = E^{Q_g}[f(P(t))] = \int gQ_t f dm$$

$$= \int fQ_t g dm = E^{Q_g}[g(P(t))] = E^{P_{P_0}^*}[g(P(t))]f(P)m(dt)$$

$$= \int fT_t g dm.$$

Thus (1) holds for all f, g. From (1), we see that

$$\left|\int gT_t f dm\right| = \left|\int fT_t g dm\right| \leq \| f \|_{L^1(m)} \| g \|$$

for all f, g . Hence $T_t f \in L^1(m) and \quad \| T_t f \|_{L^1(m)} \leq \| f \|_{L^1(m)}$ for all f. Since $\| T_t f \| \leq \| f \|$, it follows by interpolation that $\| T_t f \|_{L^{P(m)}} \leq \| f \|_{L^{P(m)}}$.

REFERENCES

1. Won Choi, *Some symmetry preserving transformation in population genetics*, J. of Appl. Math. amd info., **27** (2009), 757-762.
2. Fukushima, M, *Dirichlet forms and Markov processes*, Math. Libr. Ser. No., **23**, North-Holland Publ. Amsterdam (1980).
3. N.Okada, *On the uniqueness problem of two dimensional diffusion processes occurring in population genetics*, Z.Wahr.Verw.Geb., **56** (1981), 63-74.
4. S.N.Either, *A class of degenerate diffusion processes occurring in population genetics*, Comm. Pure Appl. Math., **29** (1976), 483-493.
5. S.N.Either, *A class of infinite dimensional diffusions occurring in population genetics*, Indiana Univ. J., **30** (1981), 925-935.
6. J.H.Gillespie, *Natural selection for within-generation variance in offspring number,* Genetics, **76** (1974), 601-606.

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

Yoo Jung Joung

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