ISSN 2347-1921

Bayesian estimation for the birth and death stochastic Markov chain

R. Rezaeyan
Department of Statistic and Mathematics, Islamic Azad University, Nour Branch, Nour, Iran.
*Corresponding Author: r_rezaeyan@iaunour.ac.ir

ABSTRACT

In science and technology, applications of Markov chain models are varied. We consider the stochastic Markov chain by adding a stochastic term to the deterministic Markov chain. In this work, we decide to estimate the parameters of the birth and death stochastic Markov chain by the Bayesian method.

Keywords: Stochastic Differential Equation; Markov Chain; Birth and Death Stochastic Markov Chain; Bayesian Estimation.

Mathematics Subject Classification: Primary 60H10, Secondary 34F05

Council for Innovative Research
Peer Review Research Publishing System
Journal: Journal of Advances in Mathematics

Vol 4, No 2
editor@cirworld.com
www.cirworld.com, member.cirworld.com

1. INTRODUCTION

Randomness is a basic type of object uncertainty and a random variable is a function of sample space to set real number. A stochastic process is set of random variables. A differential equation that contains a random component is known as a Stochastic Differential Equation (SDE). It's solution is known as a random process.

SDEs are used in the modeling of many physical biological and economic systems. Generally, SDEs cannot be solved using traditional mathematics for the steps of the transformation because the Wiener process is non-differentiable instead we need special techniques such as Ito and Stratonovich calculus (Oksendal, 2000). However, there is not always a closed form solution for SDEs, hence researchers have looked for solving them numerically (Kloeden, 1995). The methods based on numerical analysis are reported (Kloeden and Platen, 1995).

One of the most important steps in the modeling process is that of parameter estimation (Timmer, 2000). Timmer discussed three methods to estimate parameters in SDEs, the Maximum Likelihood Estimation (MLE), quasi-maximum likelihood and integration scheme. MLE is a popular method in the case of diffusion process driven by Brownian motions, when the process can be observed continuously (Prakasa Rao, 1983). When a diffusion process is observed at discrete times, in most case the transition density and hence likelihood function of the observations is not explicitly computable.

Bayesian approach for parameters estimation is applicable to a large class of discretely observed process. This method has several advantages over frequents method. One of the ability is to incorporate prior information, if such information are available.

In recent years, Wilson (2005) suggested the some applications of Bayesian statistical inference to SDEs. Panzer (2009) studied nonparametric Bayesian inference for ergodic diffusions. Meulen and Zanten (2011) are showed consistent nonparametric Bayesian estimation for discretely observed scalar diffusions. However, Gugushvili and Spreij (2012) discuss the nonparametric Bayesian drift estimation for SDEs. In this paper, applies Bayesian statistical methods to SDEs used in Markov chain. The main goal of this paper is to parameter estimation of the birth and death stochastic Markov chain.

The structure of this paper is as follows: In next section, we will consider the model stochastic birth and death process. In section 3, we will define the SDEs for the birth and death stochastic Markov chain. The parametric Bayesian estimation for the birth and death stochastic Markov chain is given in section 4.

2. The birth and death stochastic Markov chain

Markov processes represent the simplest generalization of independent processes by permitting the outcome at any instant to depend only on the outcome that precedes it. Thus, in a Markov process $X(t)$, the past has no influence on the future if the present is specified (Doob, 1953). This means that if $t_{n-1} \le t_n$, then

$$P(X(t_n) \le x_n \,|\, X(t),\, t \le t_{n-1}) = P(X(t_n) \le x_n \,|\, X(t_{n-1})). \tag{1}$$

From (1) it follows that if $t_1 \le t_2 \le \cdots \le t_n$, so

$$P(X(t_n) \le x_n \,|\, X(t_{n-1}),\, \cdots,\, X(t_1)) = P(X(t_n) \le x_n \,|\, X(t_{n-1})). \tag{2}$$

A special kind of Markov process is a Markov chain where the system can occupy a finite or countable infinite number of states $e_1, e_2, \cdots, e_j, \cdots$ such that the future evolution of the process once it is in a given state, depends only on the present state and not on how it arrived at that state, depends only on the present state and not on how it arrived at that state. Both Markov chain and Markov process can be discreet-time or continuous-time depending on whether the time index set is discreet or continuous.

In a discreet-time Markov chain X_n with a finite or infinite set of states $e_1, e_2, \cdots, e_j, \cdots$, let $X_n = X(t_n)$ represents state of the system at $t = t_n$. The numbers $p_{ij}(m,n)$,

$$p_{ij}(m,n) = P(X_n = e_j \,|\, X_m = e_i)$$

represent the transition probabilities of the Markov chain from state e_i at t_m to e_j at t_n.

P = (P_{ij}) called one-step probability transition matrix whose entries are all nonnegative, and elements in each row add to unity. Theoretical results for Markov chains with known transition probability matrix are extensive. Discrete-time Markov chain models, continuous-time Markov chain models and stochastic differential equation models are three types of stochastic models commonly used in population biology. When population sizes are large, SDEs are used to approximate the discrete-time Markov chain models or continuous-time Markov chain models (Dennis, 2002).

Let $X(t)$ denote the random variable for the total population size at time t and assume that the birth and death rates $\mu(X)$ and $\lambda(X)$ satisfy assumptions,

$i)$ $\mu(0)=\lambda(0)=0$ and $\mu(x)=0$ for $X \geq N.$

$ii)$ $\mu(X) \succ 0$ for $X \in (0,N)$ and $\lambda(X) \succ 0$ for $X \in (0,N).$

$iii)$ $\mu(X) \succ \lambda(X)$ for $X \in (0,K).$

$iv)$ $\mu(X) \prec \lambda(X)$ for $X \in (K,N).$

K *and* N are numbers such that $0 \prec K \prec N.$ In the discreet-time Markov chain model, we have

$$p_{ij}(\Delta t) = P(X(t+\Delta t) = j \mid X(t) = i),$$

where,

$$p_{ij}(\Delta t) = \begin{cases} \mu(i).\Delta t, & if \;\; j=i+1, \;\; j \in \{0,1,\cdots,N-1\} \\ \lambda(i).\Delta t, & if \;\; j=i-1, \;\; j \in \{1,2,3,\cdots,N\} \\ 1-[\mu(i)+\lambda(i)].\Delta t, & if \;\; j=i, \;\; j \in \{0,1,2,\cdots,N\} \\ 0, & otherwise \end{cases} \tag{3}$$

and $P = (\, p_{ij}(\Delta t))$ is the transition matrix (Allen, 2003)

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ \lambda(1)\Delta t & 1-[\lambda(1)+\mu(1)]\Delta t & \mu(1)\Delta t & 0 & \cdots & 0 \\ 0 & \lambda(2)\Delta t & 1-[\lambda(2)+\mu(2)]\Delta t & \mu(2)\Delta t & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1-\lambda(N)\Delta t \end{pmatrix}. \tag{4}$$

However, for the continuous-time Markov chain model, $t \in [0,\infty)$, $X(t) = 0,1,\cdots,N$ and Δt sufficiently small. The transition probabilities for the continuous-time Markov chain model assume:

$$p_{ij}(\Delta t) = \begin{cases} \mu(i).\Delta t + O(\Delta t), & if \;\; j=i+1, \;\; j \in \{0,1,\cdots,N-1\} \\ \lambda(i).\Delta t + O(\Delta t), & if \;\; j=i-1, \;\; j \in \{1,2,3,\cdots,N\} \\ 1-[\mu(i)+\lambda(i)].\Delta t + O(\Delta t), & if \;\; j=i, \;\; j \in \{0,1,2,\cdots,N\} \\ 0, & otherwise. \end{cases} \tag{5}$$

This can be written in matrix form as:

$$\frac{dP}{dt} = QP, \;\; p_{i0}(0) = 1,$$

where $Q = (\, q_{ij})$ is the infinitesimal generator matrix,

$$Q=\begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ \lambda(1) & 1-[\lambda(1)+\mu(1)] & \mu(1) & 0 & \cdots & 0 \\ 0 & \lambda(2) & 1-[\lambda(2)+\mu(2)] & \mu(2) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1-\lambda(N) \end{pmatrix}. \qquad (6)$$

3. Stochastic differential equation model

Fluctuations in statistical mechanics are usually modeled by adding a stochastic term to the deterministic differential equation. By doing this one obtains what is called SDEs, and the term stochastic called noise (Gard, 1988). Then, a SDE is a differential equation in which one or more of the terms are stochastic process, and resulting in a solution which is itself a stochastic process. In recent decades the mathematics of SDEs has played an important role in many application areas including biology, chemistry, environmental modeling and engineering (Farnoosh and Rezaeyan, 2011).

Every unwanted signal that adds to the information is called noise. A noise in dynamical system is usually considered a nuisance. However in certain nonlinear systems, including electronic circuits and biological sensory systems, the presence of noise can enhance the detection of weak signals. Noise has the most important role in SDE.

Since the path of a Wiener process is nowhere differentiable, a white noise cannot be considered a stochastic process in the usual way but it can be approximate by conventional stochastic processes with wide spectral bands which are commonly known as color noise process. The most famous example of this noise, is the Ornestein-Uhlenbeck process.
Let $X(t)$ denote the random variable for the total population size at time t and assume that both time and state are continuous variables, $t\in[0,\infty)$ and $X(t)\in[0,N]$ where N represents the maximum population size. Gardiner (1985), Allen (1999) showed that sample paths $X(t)$ of the stochastic process satisfy in the following Ito stochastic integral equation:

$$X(t)=X(0)+\int_0^t[\mu(X(u))-\lambda(X(u))]du+\int_0^t\sqrt{[\mu(X(u))+\lambda(X(u))]}\,dW(u). \qquad (7)$$

In (7), W is the standard Wiener process, where $\Delta W(t)=W(t+\Delta t)-W(t)$ has a normal distribution, $N(0,\Delta t)$. The first integral in (7) is a Riemann integral, but the second integral is an Ito stochastic integral (Gard 1988, Oksendal 2000). For notational convenience, the stochastic integral equation (7) is often expressed as the SDE:

$$\frac{dX(t)}{dt}=\mu(X(t))-\lambda(X(t))+\sqrt{\mu(X(t))+\lambda(X(t))}\,\frac{dW(t)}{dt}, \qquad X(0)=x_0\succ 0. \qquad (8)$$

Then

$$dX=(\mu-\lambda)X\,dt+\sqrt{(\mu+\lambda)\,X}\,dW, \qquad X(0)=x_0.$$

This equation no explicit solution and

$$E(X)=x_0\,e^{(\mu-\lambda)t},$$

$$Var(X)=x_0\,\frac{\mu+\lambda}{\mu-\lambda}e^{(\mu+\lambda)t}\,[e^{(\mu+\lambda)t}-1].$$

4. Parameters estimation

For simplicity in calculations, we consider in (8),

$$\mu(X(t))-\lambda(X(t))=\alpha\,X(t),$$

and

$$\mu(X(t))+\lambda(X(t))=\sigma^2\,X^2(t), \qquad (9)$$

or,

$$\mu(X(t))+\lambda(X(t))=\sigma\,X(t). \qquad (10)$$

So, from equation (8) we obtain

$$\frac{dX(t)}{X(t)} = \alpha\, dt + \sigma\, dW(t), \qquad (11)$$

or, from (10) we get

$$\frac{dX(t)}{dt} = \alpha\, X(t) + \sqrt{\sigma\, X(t)}\, \frac{dW(t)}{dt}. \qquad (12)$$

For a given positive integer n, let $\Delta t = \frac{T}{n}$ and consider the following partitions

$$\pi_n = \{0, \Delta t, 2\Delta t, \cdots, (n-1)\Delta t, T\},$$

of $[0,T]$, with a simple forward Euler discretization of the equation (11), we derive the following stochastic difference equation:

$$X_{t+\Delta t} - X_t = \alpha\, X_t\, \Delta t + \sigma\, X_t \sqrt{\Delta t}\, W_{t+\Delta t}, \qquad t = (i-1)\Delta t, \quad i = 1, 2, \cdots, n \qquad (13)$$

where $W_{t+\Delta t}$ is independent, identically distributed with the standard normal distribution.

Denote $X_{i\Delta t}$ by X_i. The Least Square Estimator (LSE) of α for equation (13) is to minimize the following contrast function:

$$\phi_{n,\varepsilon} = \phi_n(\{X_i\}_{i=0}^{n}) = \sum_{i=1}^{n} \left|(X_i - X_{i-1}) - \alpha X_{i-1}\, \Delta t\right|^2.$$

Then the LSE $\hat{\alpha}_{n,\varepsilon}$ is defined as:

$$\hat{\alpha}_{n,\varepsilon} = \arg\min_{\alpha} \phi_{n,\varepsilon}(\alpha),$$

$$\frac{\partial \phi_{n,\varepsilon}}{\partial \alpha} = 0 \Rightarrow 2\sum_{i=1}^{n} [(X_i - X_{i-1}) - \alpha\, X_{i-1}\, \Delta t]\, X_{i-1}\, \Delta t = 0,$$

$$\hat{\alpha}_{n,\varepsilon} = \frac{\sum_{i=1}^{n} X_i X_{i-1} - \sum_{i=1}^{n} X_{i-1}^{2}}{\sum_{i=1}^{n} X_i\, X_{i-1}}.$$

In a similar manner, we get from SDE (12),

$$\hat{\alpha}_{n,\varepsilon} = \frac{\sum_{i=1}^{n} X_i X_{i-1} - \sum_{i=1}^{n} X_{i-1}^{2}}{\sum_{i=1}^{n} X_i\, X_{i-1}}.$$

The SDE (12), has no explicit solution, but the equation (11) can be solved by making the transformation $Y(t) = \ln X(t)$ and using Ito's lemma to get (Kloeden, 1995),

$$dY(t) = \alpha\, dt + \sigma\, dW(t). \qquad (14)$$

If we consider (11) as a parametric of (α, σ), then these parameters are estimated via Bayesian paradigm. Bayesian estimation does not assume that α and σ have fixed values, instead it assumes that these are random variables with probability distribution.

So, this method incorporates knowledge about a particular parameter as the prior distribution.
From this a partial differential equation is derived and solved for the transition probability $P(Y_f, t_f | Y_i, t_i)$ yielding:

$$P(Y_f, t_f | Y_i, t_i) \propto \exp\left(-\frac{(Y_f - Y_i - \alpha(t_f - t_i)^2)}{2\sigma^2 (t_f - t_i)}\right). \tag{15}$$

The sampling distribution for data $Y_i = Y(t_i)$ and discrete set of times $t_0 \prec t_1 \prec \cdots \prec t_n$ is:

$$P(Y_n, \cdots, Y_0 | \alpha, \sigma) \propto \left(\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2 \Delta t_i}}\right) \exp\left(-\sum_{i=1}^{n} \frac{(\Delta Y_i - \alpha \Delta t_i)^2}{2\sigma^2 \Delta t_i}\right), \tag{16}$$

and the posterior probability for α and σ is:

$$P(\alpha, \sigma | S_1, \cdots, S_n) \propto P(S_n, S_{n-1}, \cdots, S_1 | \alpha, \sigma)\, \pi(\alpha, \sigma), \tag{17}$$

where, $S_i = \Delta Y_i = \ln\left(\frac{X_i}{X_{i-1}}\right)$.

The values of α and σ which maximum the posterior distribution for uniform $\pi(\alpha, \sigma)$ are just the maximum likelihood estimates:

$$\hat{\alpha}_{ML} = \frac{1}{t_n - t_0} \sum_{i=1}^{n} S_i,$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{S_i}{\Delta t_i} - \hat{\alpha}_{ML}\right)^2 \Delta t_i.$$

(Timmer, 2000).

With using of the standard Jeffery's prior $\pi(\alpha, \sigma) = \frac{1}{\sigma}$, the posterior distribution for σ is:

$$P(\sigma | S_1, \cdots, S_n) \propto \int d\alpha\, P(S_n, S_{n-1}, \cdots, S_1 | \alpha, \sigma)\, \pi(\alpha, \sigma), \tag{18}$$

Or with using of the prior $\pi(\alpha, \sigma) = \delta(\alpha - \hat{\alpha}_{ML})$, the posterior distribution for σ is:

$$P(\sigma | S_n, S_{n-1}, \cdots, S_1) = P(\sigma | \hat{\alpha}_{ML}) = \frac{(n\hat{\sigma}^2_{ML})^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2})\, 2^{\frac{n-3}{2}}} \left(\frac{1}{\sigma^n}\right) \exp\left(\frac{-n\hat{\sigma}^2_{ML}}{2\sigma^2}\right). \tag{19}$$

Thus,

$$\frac{1}{\sigma^2} \approx \frac{1}{n\hat{\sigma}^2_{ML}} \chi^2(n-1). \tag{20}$$

Estimators are chosen to minimize the expected loss,

$$\int d\sigma\, P(\sigma | \hat{\sigma}_{ML})\, L(\hat{\sigma}, \sigma), \tag{21}$$

for various loss functions $L(\hat{\sigma}, \sigma)$. Set $L(\hat{\sigma}, \sigma) = |\hat{\sigma} - \sigma|$, then the resulting estimators the median is defined by:

$$\int_0^{\hat{\sigma}_{ML}} d\sigma\, P(\sigma | \hat{\sigma}_{ML}) = \frac{1}{2}. \tag{22}$$

If we put $L(\hat{\sigma},\sigma)=(\hat{\sigma}-\sigma)^2$, then the resulting estimator is the expected value

$$\hat{\sigma}_{EXP} = \int_0^\infty d\sigma\, P(\sigma|\,\hat{\sigma}_{ML}) = \frac{\Gamma(\frac{n-2}{2})}{\Gamma(\frac{n-1}{2})}(\frac{n}{2})^{\frac{1}{2}}\,\hat{\sigma}_{ML}. \qquad (23)$$

5. CONCLUSION

We converted the birth and death Markov chain to the stochastic Markov chain. The purpose of this paper is the LSE for the parameter α of the birth and death stochastic Markov chain when σ is given. Also, we estimated the parameters α and σ of the stochastic model by using the Bayesian method. We used the various loss functions in order to minimize the expected loss.

REFERENCES

[1] Alcock J, Burrage K (2004). A genetic estimation algorithm for parameters of stochastic differential equations. Computational Statistic and Data Analysis**.** 47: 255-275.

[2] Allen LJS, Allen EJ (2003). A comparison of three different stochastic population models with regard to persistence time, Theoretical Population Biology. 64**:** 439-449.

[3] Allen EJ (1999). Stochastic differential equations and persistence time for two interesting Populations. Dynam. Contin. Disc. and Impul. Systems. 5: 271-281.

[4] Dennis B (2002). All effects in stochastic populations. Oikos. 96: 389-401.

[5] Doob JL (1953). Stochastic processes, John Wiley and Sons. New York.

[6] Farnoosh R, Rezaeyan R (2011). A stochastic perspective of RL electrical circuit using different noise terms. The International Journal for Computation and Mathematics in Electrical and Electronic Engineering. 30 (2): 812-822**.**

[7] Gard TC (1988). Introduction to stochastic differential equations. Marcel dekker, Inc, New York and Basel.

[8] Gardiner CW (1985). Handbook of stochastic methods for physics, chemistry and the natural Science. Springer.

[9] Gugushvili S, Spreij P (2012). Non-parametric Bayesian drift estimation for SDEs. ArXiv. Math. St.

[10] Kloeden PE, Platen E (1995). Numerical solution of stochastic differential equations. Springer, Berlin.

[11] Meulem V, Zanten HV (2011). Consistent nonparametric Bayesian estimation for discretely observed scalar diffusions. To appear in Bernoulli.

[12] Oksendal B (2000). Stochastic differential equations an introduction with applications. 5th Edition, Springer, Berlin.

[13] Panzer L, Zanten HV (2009). Nonparametric Bayesian inference for ergodic diffusions, J. Statis. Plann Inference, 139: 4193-4199.

[14] Parakasa Rao BLS (1983). Asympotic theory for nonlinear least square estimator for diffusion processes. Math. Operat. Forsch. 14: 195-209.

[15] Rezaeyan R, Baloui EJ (2011). The Effect of Different Noise Perturbations on the Parameter of the RC Circuit use in SDE. World Applied Sciences Journal.13 (10): 2198-2202.

[16] Rezaeyan R, Farnoosh R (2010). On the support of an Ito diffusion process in the SDEs. The 23th International Conference of Jangjeon Mathematical Society. P.107-109.

[17] Rezaeyan R, Farnoosh R , Baloui EJ (2011). Application of the Kalman Bucy Filter in the SDEs for the Modeling of RL Circuit. Int. J. Nonlinear Anal. Appl. 2(1): 35-41.

[18] Timmer J (2000). Parameter estimation in nonlinear stochastic differential equations, Chaos Solitions and Fractals. 11: 2571-2578.

[19] Wilson JA (2005). Some applications of Bayesian statistical inference to SDEs, A thesis for B. S., M. S., The Ohio state university