1

Parametric Stochastic Differential Equations

Stochastic differential equations (SDEs) are a natural choice to model the time evolution of dynamic systems which are subject to random influences (cf. Arnold (1974), Van Kampen (1981)). For example, in physics the dynamics of ions in superionic conductors are modelled via Langevin equations (cf. Dieterich *et al.* (1980)), and in engineering the dynamics of mechanical devices are described by differential equations under the influence of process noise as errors of measurement (cf. Gelb (1974)). Other applications are in biology (cf. Jennrich and Bright (1976)), medicine (cf. Jones (1984)), econometrics (cf. Bergstrom (1976, 1988)), finance (cf. Black and Scholes (1973)), geophysics (cf. Arato (1982)) and oceanography (cf. Adler *et al.* (1996)).

It is natural that a model contains unknown parameters. We consider the model as the parametric Itô stochastic differential equation

$$dX_t \;=\; \mu\,(\theta,\ t,\ X_t)\ dt \;+\; \sigma\,(\vartheta,\ t,\ X_t)\ dW_t, \quad t \geq 0, \quad X_0 = \zeta$$

where $\{W_t, t \geq 0\}$ is a standard Wiener process, $\mu : \Theta \times [0,T] \times \mathbb{R} \to \mathbb{R}$, called the *drift coefficient*, and $\sigma : \Xi \times [0,T] \times \mathbb{R} \to \mathbb{R}^+$, called the *diffusion coefficient*, are known functions except the unknown parameters θ and ϑ, $\Theta \subset \mathbb{R}$, $\Xi \subset \mathbb{R}$ and $E(\zeta^2) < \infty$. The drift coefficient is also called the *trend coefficient* or *damping coefficient* or *translation coefficient*. The diffusion coefficient is also called *volatility coefficient*. Under local Lipschitz and the linear growth conditions on the coefficients μ and σ, there exists a unique strong solution of the above Itô SDE, called the *diffusion process* or simply a *diffusion*, which is a continuous strong Markov semimartingale. The drift and the diffusion coefficients are respectively the instantaneous mean and instantaneous standard deviation of the process. Note that the diffusion coefficient is almost surely determined by the process, i.e., it can be estimated without any error if observed continuously throughout a time interval (see Doob (1953), Genon-Catalot and Jacod (1994)). We assume that the unknown parameter in the diffusion coefficient ϑ is known and for simplicity only we shall assume that $\sigma = 1$ and our aim is to estimate the unknown parameter θ.

First we sketch some very popular SDE models.

Bachelier Model

$$dX_t = \beta\ dt + \sigma dW_t$$

Black-Scholes Model

$$dX_t = \beta X_t\ dt + \sigma X_t dW_t$$

Ornstein-Uhlenbeck Model

$$dY_t = \beta X_t\ dt + \sigma dW_t$$

Feller Square root or Cox-Ingersoll-Ross Model

$$dX_t = (\alpha - \beta X_t)dt + \sigma\sqrt{X_t}dW_t$$

Radial Ornstein-Uhlenbeck Process

$$dX_t = (\alpha X_t^{-1} - X_t)\ dt + \sigma dW_t$$

Squared Radial Ornstein-Uhlenbeck Process

$$dX_t = (1 + 2\beta X_t)\ dt + 2\sigma\sqrt{X_t}dW_t$$

Note that X_t the square of the Ornstein-Uhlenbeck process Y_t

$$dY_t = \beta Y_t\ dt + \sigma dW_t$$

Chan-Karloyi-Logstaff-Sanders Model

$$dX_t = \kappa(\theta -\ X_t)\ dt + \sigma\ X_t^{\gamma}\ dW_t$$

Hyperbolic Diffusion

$$dX_t = \alpha\frac{X_t}{\sqrt{1 + X_t^2}}\ dt + \sigma dW_t$$

Gompertz Diffusion

$$dX_t = (\alpha X_t - \beta X_t \log X_t)\ dt + \sigma X_t dW_t$$

Here X_t is the tumor volume which is measured at discrete time, α is the intrinsic growth rate of the tumor, β is the tumor growth acceleration factor, and σ is the diffusion coefficient.

The knowledge of the distribution of the estimator may be applied to evaluate the distribution of other important growing parameters used to access tumor treatment modalities. Some of these parameters are the plateau of the model $X_\infty = \exp(\frac{\alpha}{\beta})$, tumor growth decay, and the first time the growth curve of the model reaches X_∞.

Logistic Diffusion

Consider the stochastic analogue of the logistic growth model

$$dX_t = (\alpha X_t - \beta X_t^2)\ dt + \sigma X_t dW_t$$

This diffusion is useful for modeling the growth of populations.

Kessler-Sørensen Model

$$dX_t = -\theta \tan(X_t)\ dt + \sigma dW_t$$

By applying Itô formula, a diffusion process with some diffusion coefficient can be reduced to one with unit diffusion coefficient.
Following are most popular short term interest rate models.

Vasicek Model

$$dX_t = (\alpha + \beta X_t)dt + \sigma dW_t$$

Cox-Ingersoll-Ross Model

$$dX_t = (\alpha + \beta X_t)dt + \sigma\sqrt{X_t}dW_t$$

Dothan Model

$$dX_t = (\alpha + \beta X_t)dt + \sigma X_t dW_t$$

Black-Derman-Toy Model

$$dX_t = \beta(t)X_t dt + \sigma(t)X_t dW_t$$

Black-Karasinksi Model

$$d(\log X_t) = (\alpha(t) + \beta(t)\log X_t)dt + \sigma_t dW_t$$

Ho-Lee Model

$$dX_t = \alpha(t)dt + \sigma dW_t^H$$

Hull-White (Extended Vasicek) Model

$$dX_t = (\alpha(t) + \beta(t)X_t)dt + \sigma_t dW_t$$

Hull-White (Extended CIR) Model

$$dX_t = (\alpha(t) + \beta(t)X_t)dt + \sigma_t\sqrt{X_t}dW_t$$

Cox-Ingersoll-Ross 1.5 model

$$dX_t = \sigma X_t^{3/2}dW_t$$

Inverse Square Root Model or Ahn-Gao Model

$$dX_t = \beta(\mu - X_t)X_t dt + \sigma X_t^{3/2}dW_t$$

Ait-Sahalia Model

$$dX_t = (\alpha + \beta X_t + \gamma X_t^{-1} + \delta X_t^2)dt + \sigma X_t^{\gamma} dW_t$$

This a nonlinear interest rate model.

For existence and uniqueness of solutions of finite dimensional stochastic differential equations, properties of stochastic integrals, and diffusion and diffusion type processes see e.g., the books by McKean (1969), Gikhman and Skorohod (1972), Itô and McKean (1974), McShane (1974), Arnold (1974), Friedman (1975), Stroock and Varadhan (1979), Elliot (1982), Ikeda and Watanabe (1989), Rogers and Williams (1987), Karatzas and Shreve (1987), Liptser and Shiryayev (1977, 1989), Kunita (1990), Protter (1990), Revuz and Yor (1991), Øksendal (1995), Krylov (1995), Mao (1997). For numerical analysis and approximations of SDEs see the books by Gard (1988), Kloeden, Platen and Schurz (1994), Kloeden and Platen (1995) and Milshtein (1995). For existence, uniqueness of solutions and other properties of infinite dimensional SDEs see the books by Curtain and Pritchard (1978), Metivier and Pellaumail (1980), Itô (1984), Walsh (1986) and Kallianpur and Xiong (1995).

The asymptotic approach to statistical estimation is frequently adopted because of its general applicability and relative simplicity. In this monograph we study the asymptotic behaviour of several estimators of the unknown parameter θ appearing in the drift coefficient based on observations of the diffusion process $\{X_t, t \geq 0\}$ on a time interval $[0, T]$. Note that the observation of diffusion can be continuous or discrete. Continuous observation of diffusion is a mathematical idealization and has a very rich theory, for example Itô stochastic calculus, stochastic filtering, inference for continuously observed diffusions and much more behind it. But the path of the diffusion process is very kinky and no measuring device can follow a diffusion trajectory continuously. Hence the observation is always discrete in practice. Research on discretely observed diffusions is growing recently with a powerful theory of simulation schemes and numerical analysis of SDEs behind it.

The asymptotic estimation of θ, based on continuous observation of $\{X_t\}$ on $[0, T]$ can be studied by different limits, for example, $T \to \infty$, $\sigma(\vartheta, t, X_t) \to 0$, $\mu(\theta, t, X_t) \to \infty$, or any combination of these conditions that provide the increase of the integrals $\int_0^T [\mu(\theta, t, X_t)\sigma^{-1}(\vartheta, t, X_t)]^2 dt$ and $\int_0^T [\mu'(\theta, t, X_t)\sigma^{-1}(\vartheta, t, X_t)]^2 dt$, where prime denotes derivative with respect to θ. Parameter estimation in SDE was first studied by Arato, Kolmogorov and Sinai (1962) who applied it to a geophysical problem. For long time asymptotics ($T \to \infty$) of parameter estimation in stochastic differential equations see the books by Liptser and Shiryayev (1978), Basawa and Prakasa Rao (1980), Arato (1982), Linkov (1993), Küchler and Sørensen (1997), Prakasa Rao (1999) and Kutoyants (1999). For small noise asymptotics ($\sigma \to 0$) of parameter estimation see the books by Ibragimov and Khasminskii (1981) and Kutoyants (1984a, 1994a).

If $\{X_t\}$ is observed at $0 = t_0 < t_1 < t_2 < ... < t_n = T$ with $\Delta_n = \max_{1 \le i \le n} |t_i - t_{i-1}|$ the asymptotic estimation of θ can be studied by different limits, for example, $\Delta_n \to 0, n \to \infty$ and $T \to \infty$ (or $\sigma \to 0$) or $\Delta_n = \Delta$ remaining fixed and $n \to \infty$. See Genon-Catalot (1987).

In the infinite dimensional diffusion models there are even different asymptotic frameworks. For example, in a stochastic partial differential equation, based on continuous observation, asymptotics can also be obtained when the intensity of noise and the observation time length remain fixed, but the number of Fourier coefficients in the expansion of the solution random field increases to infinity. Based on discrete observations, asymptotics can be obtained by this condition along with some sampling design conditions of discrete observations as in the finite dimensional case.

Our asymptotic framework in this monograph is long time for continuous observation and decreasing lag time along with increasing observation time for discrete observations.

The monograph is broadly divided into two parts. The first part (Chapters 2-6) deals with the estimation of the drift parameter when the diffusion process is observed continuously throughout a time interval. The second part (Chapters 7-10) is concerned with the estimation of the drift parameter when the diffusion process is observed at a set of discrete time points.

Asymptotic properties such as weak or strong consistency, asymptotic normality, asymptotic efficiency etc. of various estimators of drift parameter of Itô SDEs when observed continuously throughout a time interval, has been studied extensively during the last three decades. In linear homogeneous SDEs, maximum likelihood estimation was studied by Taraskin (1974), Brown and Hewitt (1975a), Kulinich (1975), Lee and Kozin (1977), Feigin (1976, 1979), Le Breton (1977), Tsitovich (1977), Arato (1978), Bellach (1980, 1983), Le Breton and Musiela (1984), Musiela (1976, 1984), Sørensen (1992), Küchler and Sørensen (1994a,b), Jankunas and Khasminskii (1997) and Khasminskii *et al.* (1999). In nonlinear homogeneous SDEs maximum likelihood estimation was studied by Kutoyants (1977), Bauer (1980), Prakasa Rao and Rubin (1981), Bose (1983a, 1986b), Bayes estimation was studied by Kutoyants (1977), Bauer (1980), Bose (1983b, 1986b), maximum probability estimation was studied by Prakasa Rao (1982), minimum contrast estimation was studied by Lanska (1979), M-estimation was studied by Yoshida (1988, 1990), minimum distance estimation was studied by Dietz and Kutoyants (1997). In nonlinear nonhomogeneous SDEs maximum likelihood estimation was studied by Kutoyants (1978, 1984a), Borkar and Bagchi (1982), Mishra and Prakasa Rao (1985), Dietz (1989) and Levanony, Shwartz and Zeitouni (1993, 1994), Bayes estimation was studied by Kutoyants (1978, 1984a). For survey of work in continuously observed diffusions, see Bergstrom (1976), Prakasa Rao (1985), Barndorff-Neilson and Sørensen (1994) and Prakasa Rao (1999). The following is a summary of Chapters 1-5.

In Chapter 2 we start with the historically oldest example of stochastic differential equation called the Langevin equation and whose solution is called the Ornstein-Uhlenbeck (O-U) process. In this case $\mu(\theta, t, X) = \theta X_t$. The first order theory like consistency, asymptotic normality etc. is well known for this case, see Le Breton (1977), Liptser and Shiryayev (1978). We study the rate of convergence in consistency and asymptotic normality via the large deviations probability bound and the Berry-Esseen bound for the minimum contrast estimator (MCE) of the drift parameter when the process is observed continuously over $[0, T]$. Then we study more general nonlinear ergodic diffusion model and study the Berry-Esseen bound for Bayes estimators. We also posterior large deviations and posterior Berry-Esseen bound. Mishra and Prakasa Rao (1985a) obtained $O(T^{-1/5})$ Berry-Esseen bound and $O(T^{-1/5})$ large deviation probability bound for the MLE for the Ornstein-Uhlenbeck model. For the MLE, Bose (1986a) improved the Berry-Esseen bound to $O(T^{-1/2}(\log T)^2)$. (The main result in Bose (1986a) has a misprint and gives the rate as $O(T^{-1/2})$, but by following the proof given there it is clear that the rate is $O(T^{-1/2}(\log T)^2)$.) Bose (1985) obtained the rate $O(T^{-1/2} \log T)$. Bishwal and Bose (1995) improved this rate to $O(T^{-1/2}(\log T)^{1/2})$. For the MLE, Bishwal (2000a) obtained $O(T^{-1})$ bound on the large deviation probability and the Berry-Esseen bound of the order $O(T^{-1/2})$ using nonrandom norming. This bound is consistent with the classical i.i.d. situation. Next we consider nonlinear diffusion model and obtain exponential rate of concentration of the posterior distribution, suitably normalized and centered at the MLE, around the true value of the parameter and also $O(T^{-1/2})$ rate of convergence of posterior distribution to normal distribution. We then establish $o(T^{-1/2})$ bound on the equivalence of the MLE and the Bayes estimator, thereby improving the $O(T^{-3/20})$ bound in Mishra and Prakasa Rao (1991). We obtain $O(T^{-1/2})$ Berry-Esseen bound and $O(T^{-1})$ bound on the large deviation probability of the BEs. This chapter is adapted from Bishwal (2004a) and Bishwal (2005a).

In Chapter 3 we deal with estimation in nonlinear SDE with the parameter appearing nonlinearly in the drift coefficient, based on continuous observation of the corresponding diffusion process over an interval $[0, T]$. In this case $\mu(\theta, t, X_t) = f(\theta, X_t)$. We obtain exponential bounds on large deviation probability for the MLE and regular BEs. The method of proof is due to Ibragimov and Khasminskii (1981). Some examples are presented. This chapter is adapted from Bishwal (1999a).

In Chapter 4 we study the asymptotic properties of various estimators of the parameter appearing nonlinearly in the nonhomogeneous drift coefficient of a functional stochastic differential equation when the corresponding solution process, called the diffusion type process, is observed over a continuous time interval $[0, T]$. We show that the maximum likelihood estimator, maximum probability estimator and regular Bayes estimators are strongly consistent and when suitably normalised, converge to a mixture of normal distribution and are locally asymptotically minimax in the Hajek-Le Cam sense as $T \to \infty$

under some regularity conditions. Also we show that posterior distributions, suitably normalised and centered at the maximum likelihood estimator, converge to a mixture of normal distribution. Further, the maximum likelihood estimator and the regular Bayes estimators are asymptotically equivalent as $T \to \infty$. We illustrate the results through the exponential memory Ornstein-Uhlenbeck process, the nonhomogeneous Ornstein-Uhlenbeck process and the Kalman-Bucy filter model where the limit distribution of the above estimators and the posteriors is shown to be Cauchy. This chapter is adapted from Bishwal (2004b).

In Chapter 5 we study estimation of a real valued parameter in infinite dimensional SDEs based on continuous observation of the diffusion. This area is relatively young and in our opinion is exciting and difficult. A few contributions in the existing literature are devoted to parameter (finite or infinite dimensional) estimation in infinite dimensional SDEs see, e.g., Aihara (1992, 1994, 1995), Aihara and Bagchi (1988, 1989, 1991), Bagchi and Borkar (1984), Loges (1984), Koski and Loges (1985, 1986), Huebner, Khasminskii, Rozovskii (1992), Huebner (1993), Huebner and Rozovskii (1995), Kim (1996). We consider the drift coefficient as $\theta A X_t$ with A being the infinitesimal generator of a strongly continuous semigroup acting on a real separable Hilbert space H and θ real valued. We obtain the Bernstein-von Mises theorem concerning the normal convergence of the posterior distributions and and strong consistency and asymptotic normality of the BEs of a parameter appearing linearly in the drift coefficient of Hilbert space valued SDE when the solution is observed continuously throughout a time interval $[0, T]$ and $T \to \infty$. It is also shown that BEs, for smooth priors and loss functions, are asymptotically equivalent to the MLE as $T \to \infty$. Finally, the properties of sequential maximum likelihood estimate of θ are studied when the corresponding diffusion process is observed until the observed Fisher information of the process exceeds a predetermined level of precision. In particular, it is shown that the estimate is unbiased, uniformly normally distributed and efficient. This chapter is adapted from Bishwal (1999b) and Bishwal (2002a).

In Chapter 6 we consider non-Markovian non-semimartingale models. Recently long memory processes or stochastic models having long range dependence phenomena have been paid a lot of attention in view of their applications in finance, hydrology and computer networks (see Beran (1994), Mandelbrot (1997), Shiryaev (1999), Rogers (1997), Dhehiche and Eddahbi (1999)). While parameter estimation in discrete time models having long-range dependence like the autoregressive fractionally integrated moving average (ARFIMA) models have already been paid a lot of attention, this problem for continuous time models is not well settled. Here we study estimation problem for continuous time long memory processes. This chapter is adapted from Bishwal (2003a).

Parameter estimation in diffusion processes based on observations at discrete time points is of much more practical importance due to the impossibility of observing diffusions continuously throughout a time interval. Note that diffusion process can be observed either at deterministic or at random sampling instants. For random sampling, e.g., from a point process, the sampling process may be independent of the observation process or may be dependent on it. Also in random sampling scheme, e.g., Poisson sampling (see Duffie and Glynn (2004)) the inverse estimation problem arises i.e., the estimation of the parameters of the sampling process when the parameters of the observation process is known. For a survey of earlier works on inference in continuous time processes based on observations at random sampling schemes see Stoyanov (1984). Jacod (1993) studied random sampling from a process with independent increments. Later on this scheme was used by Genon-Catalot and Jacod (1994) for the estimation of the parameter of the diffusion coefficient of a diffusion process. Duffie and Glynn (2004) studied the asymptotics of generalized method of moments estimators for a continuous time Markov process from observations at random sampling instants. In this monograph we will only be dealing with deterministic sampling scheme. We assume that the diffusion process $\{X(t)\}$ is observed at $\{0 = t_0 < t_1 < \ldots < t_n = T\}$ with $t_i - t_{i-1} = \frac{T}{n} = h \to 0, i = 1, 2, \ldots, n$.

Note that when one observes the process continuously throughout a time interval the diffusion coefficient is almost surely determined by the process. But when one has discrete observations, the problem of estimation of the diffusion coefficient also arises. Estimation of the parameter in the diffusion coefficient from discrete observations has been studied by Penev (1985), Dohnal (1987), Genon-Catalot and Jacod (1993), Florens-Zmirou (1993), and others. However, we will not deal with estimation of the diffusion coefficient in this monograph.

Drift parameter estimation in diffusion processes based on discrete observations has been studied by many authors. Le Breton (1976) and Dorogovcev (1976) appear to be the first persons to study estimation in discretely observed diffusions. While Le Breton (1976) used approximate maximum likelihood estimation, Dorogovcev (1976) used conditional least squares estimation. Robinson (1977) studied exact maximum likelihood estimation in discretely observed Ornstein-Uhlenbeck process. Other works on approximate maximum likelihood estimation (where the continuous likelihood is approximated), also called the maximum contrast estimation, are Bellach (1983), Genon-Catalot (1987, 1990), Yoshida (1992), Bishwal and Mishra (1995), Harison (1996), Clement (1993, 1995, 1997a,b) and Kessler (1997). Dacunha-Castelle and Florens-Zmirou (1986) studied consistency and asymptotic normality of MLE by using an expansion of the transition density of an ergodic diffusion. While ideally one should of course use maximum likelihood, in practice it is difficult because only in a few cases the transition densities are available. Pedersen (1995a,b) used numerical approximations based on iterations of the Gaussian transition densities emanating from the Euler-Maruyama scheme and studied approximate maximum likelihood estimation. Ait-Sahalia (2002) used Hermite function expansion of the transition density giving an accurate theoretical approximation and studied approximate maximum likelihood estimation.

In the conditional least squares estimation method (see Hall and Heyde (1981)) one minimizes the quadratic

$$Q_n(\theta) = \sum_{i=1}^{n} \frac{\left[X_{t_i} - X_{t_{i-1}} - \mu(\theta, t_{i-1}, X_{t_{i-1}})\Delta t_i\right]^2}{\sigma^2(t_{i-1}, X_{t_{i-1}})\Delta t_i}.$$

For equally spaced partition $0 = t_0 < t_1 < \ldots < t_n = T$ with $\Delta t_i = t_i - t_{i-1} = \frac{T}{n}, i = 1, 2, \ldots, n$, for the homogeneous stationary ergodic diffusion, Dorogovcev (1976) proved the weak consistency of the CLSE $\hat{\theta}_{n,T}$ as $T \to \infty$ and $\frac{T}{n} \to 0$ which we call the *slowly increasing experimental design* (SIED) condition. Kasonga (1988) proved the strong consistency of $\hat{\theta}_{n,T}$ as $T \to \infty$ and $\frac{T}{n} \to 0$. Prakasa Rao (1983) proved the asymptotic normality and asymptotic efficiency of $\hat{\theta}_{n,T}$ as $T \to \infty$ and $\frac{T}{\sqrt{n}} \to 0$, called the *rapidly increasing experimental design* (RIED) condition (see, Prakasa Rao (1988b)). Penev (1985) studied the consistency and asymptotic normality of a multidimensional parameter extending Dorogovcev (1976) and Prakasa Rao (1983). Florens-Zmirou (1989) proved the weak consistency of the minimum contrast estimator as $T \to \infty$ and $\frac{T}{n} \to 0$. He proved the asymptotic normality of this estimator as $T \to \infty$ and $\frac{T}{n^{2/3}} \to 0$ which we call the *moderately increasing experimental design* (MIED) condition. The properties of AMLE based on Itô type approximation of the Girsanov density are studied by Yoshida (1992) by studying the weak convergence of the approximate likelihood ratio random field when $T \to \infty$ and $\frac{T}{n^{2/3}} \to 0$. Genon-Catalot (1990) studied the asymptotic properties of maximum contrast estimators (using contrast function related to an approximate likelihood) in the nonlinear SDE as the intensity of noise becomes small and $T \to \infty$. Genon-Catalot *et al.* (1998a,b) studied the asymptotic properties of minimum contrast estimator in a stochastic volatility model that is a partially observed diffusion process. Note that the estimation methods discussed in this paragraph through different approaches are equivalent.

There are many other approaches to drift estimation for discretely observed diffusions. Prakasa Rao (1988b, 1999) gave a survey of estimation in discretely observed stochastic processes. Because of the difficulty in performing accurate maximum likelihood, much research has focussed on finding alternatives in the form of various estimating functions. Bibby and Sørensen (1995a, b) allowed $T \to \infty$ and $n \to \infty$ letting $t_i - t_{i-1} = \Delta (i = 1, 2, \ldots, n)$ fixed and found approximate martingale estimating functions based on approximate log-likelihood function and showed that the estimators based on these estimating functions are consistent and asymptotically normal as $n \to \infty$. Other estimating functions have also been proposed in the literature, e.g., estimating functions based on eigenvalues (see Kessler and Sørensen (1999)) and simple, explicit estimating functions (see Kessler (2000)). Sørensen, H. (2001, 2004), Sørensen, M. (1997, 1999) and Jacobsen (2002) also studied discretely observed diffusions through estimating functions. Bibby and Sørensen

(1995b) gave a review of martingale estimating functions based on discretely observed diffusions. Pedersen (1994) used quasi-likelihood approach for martingale estimating functions. McLeish and Kolkiewicz (1997) proposed method of estimation based on higher order Itô-Taylor expansion.

Lo (1988) studied the maximum likelihood estimation (both based on exact likelihood function based on transition densities and also on approximate discretized model) of a jump-diffusion process. Laredo (1990) studied the asymptotic sufficiency property of incomplete observations of a diffusion process which include discrete observations and studied consistency and asymptotic normality of the minimum contrast estimator. Sørensen (1992) studied the properties of estimates based on discrete observations from a linear SDE by embedding the discrete process into the continuous one. Kloeden *et al.* (1992) studied the effect of discretization on drift estimation. Gouriéroux *et al.* (1993), Gouriéroux and Monfort (1994) and Broze *et al.* (1998) studied the properties of estimates by indirect inference method. Overbeck and Ryden (1997) studied the asymptotics of the MLE and the CLSE in the Cox-Ingersoll-Ross model whose solution is a Bessel process. Gallant and Long (1997) showed the asymptotic properties of minimum chi-squared estimator as the moment function entering the chi-squared criterion and the number of past observations entering each moment function increase. Elerian *et al.* (2001) studied Bayesian estimation in nonlinear SDEs through a MCMC method using the Euler-Maruyama discretization scheme. The following is a summary of Chapters 7-10.

In Chapter 7 we assume $\mu(\theta, t, X_t) = \theta f(X_t)$. If complete (continuous) observation were available, a desirable estimate would be the maximum likelihood estimate. Based on discrete observations, one approach of finding a good estimate would to try to find an estimate which is as close as possible to the continuous MLE. For further refinement, one should measure the loss of information due to discretization. For the Ornstein-Uhlenbeck process, Le Breton (1976) proposed an approximate maximum likelihood estimator (AMLE) and showed that the difference between his AMLE and the continuous MLE is of the order $O_P((\frac{T^2}{n})^{1/2})$. We obtain several AMLEs with faster rates of convergence than that of Le Breton (1976). For this purpose first we obtain several higher order discrete approximations of the Fisk-Stratonovich integral. We use these approximations and the rectangular rule ordinary integral approximation to obtain different AMLEs. *Interalia* we introduce a new stochastic integral which will be of independent interest.

In Chapter 8 we return to the Ornstein-Uhlenbeck process, i.e., $\mu(\theta, t, X_t) = \theta X_t$ and we investigate the asymptotic properties of the conditional least squares estimator (CLSE) (see Hall and Heyde (1981)) as $T \to \infty$ and $\frac{T}{n} \to 0$. For the homogeneous nonlinear stationary ergodic case, i.e., $\mu(\theta, t, X) = f(\theta, X_t)$ under some regularity conditions Dorogovcev (1976) showed that the weak consistency and Kasonga (1988) showed the strong consistency of the CLSE as $T \to \infty$ and Prakasa Rao (1983) showed that one needs $T \to \infty$ and $\frac{T}{\sqrt{n}} \to 0$ as $n \to \infty$. This means that one needs larger number of observations to obtain asymptotic normality of the CLSE than for consistency. Till date no approximation results are known for this estimator. We obtain Berry-Esseen bound of the order $O(\max(T^{-1/2}(\log T)^{1/2}, \frac{T^2}{n}(\log T)^{-1}, \frac{T^2}{n}(\log T)^{-1}))$ for this estimator using nonrandom and parameter free random nomings. Using parameter dependent random norming, we obtain the rate $O(T^{-1/2} \bigvee (\frac{T^2}{n})^{1/3})$. We also obtain large deviation probability bound for the CLSE. Its rate of convergence to the continuous MLE for fixed T, is of the order $O_P(\frac{T^2}{n})^{1/2}$. We study another approximate MLE here whose Berry-Esseen bound is of the order $O(T^{-1/2}(\log T)^{1/2} \bigvee \frac{T^4}{n^2}(\log T)^{-1})$ using nonrandom and sample dependent random normings. With a random norming which is parameter dependent the Berry-Esseen bound is shown to be of the order $O(T^{-1/2} \bigvee (\frac{T}{n})^{2/3}))$ and its rate of convergence to the continuous MLE is of the order $O_P(\frac{T^2}{n})$. From the above result it is clear that one needs $T \to \infty$ and $\frac{T}{n^{2/3}} \to 0$ as $n \to \infty$ for the asymptotic normality of the AMLE and the AMLE has a faster rate of convergence than CLSE. This chapter is adapted from Bishwal and Bose (2001) and Bishwal (2006a).

In Chapter 9 we consider discretely observed SDE with homogeneous stationary ergodic solution where the parameter and the process appear nonlinearly in the drift coefficient, i.e., $\mu(\theta, t, X_t) = f(\theta, X_t)$. Asymptotic normality of approximate MLEs, approximate Bayes estimators and approximate maximum probability estimators of the drift parameter based on two different approximate likelihoods are obtained via the study of weak convergence of the approximate likelihood ratio random fields. Also the *Bernstein-von Mises type theorems* with the two approximate likelihoods are studied. Asymptotic properties of conditional least squares estimator are studied via the weak convergence of the least squares random field. We relax to some extent the regularity conditions and the RIED condition used by Prakasa Rao (1983) who obtained asymptotic normality through the usual normal equations and Cramer's approach. Instead we use the *moderately increasing experimental design* (MIED) condition, i.e., $T \to \infty$ and $\frac{T}{n^{2/3}} \to 0$ as $n \to \infty$. This chapter is adapted from Bishwal (1999c) and Bishwal (2005b).

In Chapter 10 we show that discretization after the application of Itô formula in the Girsanov likelihood produces estimators of the drift which have faster rates of convergence than the Euler estimator for stationary ergodic diffusions and is free of approximating the stochastic integral. The discretization schemes are related to the Hausdorff moment problem. We show strong consistency, asymptotic normality and a Berry-Esseen bound for the corresponding approximate maximum likelihood estimators of the drift parameter from high frequency data observed over a long time period. This chapter is adapted from Bishwal (2007a).