Martingales and diffusions, limit theory and statistical inference
van Zanten, J.H.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Introduction

In this thesis we study diffusion processes and martingales. Diffusions are used in many branches of applied stochastics. The basic example is the Brownian motion process, which models the irregular movements of pollen suspended in water. Other classical examples are the Ornstein-Uhlenbeck process and the Wright-Fisher model that are used in physics and genetics, respectively. Since Black and Scholes (1973) and Merton (1973) did their important work on the pricing and hedging of options, diffusion processes and stochastic differential equations also play an important role in mathematical finance. They are used to model all kinds of financial phenomena, such as stock prices, exchange rates and interest rates.

There exists already a vast amount of literature on statistical inference for diffusion processes. It is a field of research that still expands rapidly, for a large part motivated by applications in mathematical finance. A partial overview of recently developed methods can be found in the paper Dzhaparidze et al. (2000). Compared to other branches of statistical theory, statistics for diffusion processes is still a relatively young area. As a result, there are a lot of questions that still have to be answered. In this thesis we focus in particular on problems related to the asymptotic behaviour of statistical procedures for diffusions. We will develop new limit theory for diffusion processes and apply this in the asymptotic analysis of various parametric and nonparametric estimators.

Martingales play an important role in the theory of diffusions and stochastic differential equations. To prove asymptotic statistical results for diffusion processes one has to rely heavily on martingale limit theory. Although this area of probability theory is already very well developed, improvements and extensions are sometimes necessary. A substantial part of this dissertation is therefore devoted to the study of martingales. This work is motivated by the statistical questions sketched above, but the area of applications of martingale theory is of course much wider.
Introduction

The thesis is divided into two parts. Part I deals with continuous martingales and diffusions, the second part is concerned with càdlàg martingales (martingales with sample paths that are right continuous and that have limits from the left). The parts can be read independently. The remainder of this introduction gives a brief overview of the presented results.

Overview of Part I

The first chapter deals with continuous local martingales. The material in this chapter is applied throughout part I of the thesis. We start by recalling a fundamental result in the theory of continuous martingales: the Dambis-Dubins-Schwarz theorem. This theorem states that each continuous local martingale is in fact a time-changed Brownian motion. The result can be used to reduce all kinds of questions regarding continuous local martingales to questions concerning Brownian motion. We use the time-change theorem to improve a number of well-known results. We first provide an extension of the classical Bernstein inequality for continuous martingales. Later in the chapter, this inequality will be used as the starting point in our treatment of entropy methods for continuous martingales. The Dambis-Dubins-Schwarz theorem is also used to relax the conditions of the central limit theorem for multivariate, normalized martingales given by Küchler and Sørensen (1999).

An important step in the proof of this central limit theorem is a result on so-called nested Brownian motions that is of independent interest. If $W$ is a standard Brownian motion and $a_n$ is a sequence of positive numbers that increase to infinity, then the scaling property of Brownian motion implies that the processes $W^n$ defined by $W^n_t = W_{a_n t} / \sqrt{a_n}$ are again Brownian motions. The sequence $W^n$ is a typical example of the nested sequences that we will consider. It turns out that such nested Brownian motions are asymptotically independent of any other random element on the same probability space. This result will allow us to prove the central limit theorem for normalized continuous martingales, by embedding them in nested Brownian motions.

Chapter 1 is closed with a treatment of entropy methods for continuous martingales. In particular, we give a bound on the modulus of continuity of an indexed family of continuous martingales in terms of the entropy of the index set. In chapter 2 we use this result to prove the uniform weak convergence of certain families of martingales. Entropy methods for con-
tinuous martingales have recently been studied systematically by Nishiyama (1997, 1999, 2000). We give new proofs of his results, relying on general results on sub-Gaussian random maps taken from Van der Vaart and Wellner (1996). This approach is simpler and therefore more transparent than the direct arguments of Nishiyama.

Chapter 2 is devoted to limit theory for one-dimensional, regular diffusion processes. We consider a regular diffusion $X$ whose speed measure $m$ has finite total mass on the state space of the diffusion. It is well-known that for such processes, the empirical measures $\mu_t$, defined by

$$\mu_t(B) = \frac{1}{t} \int_0^t 1_B(X_u) \, du,$$

are absolutely continuous with respect to the normalized speed measure $\mu$. The random density $\rho_t = d\mu_t/d\mu$ is the main object of investigation. To study the asymptotic properties of this density we rely again on time-change arguments. Indeed, a central result in diffusion theory states that a regular diffusion $X$ in natural scale is a time-changed Brownian motion. It holds that $X_t = W_{\tau_t}$, where $W$ is a Brownian motion and the time-change $\tau_t$ is determined by the speed measure. This result is closely related to the Dambis-Dubins-Schwarz theorem for continuous martingales.

The first new theorem states that $\tau_t/t^2$ converges in distribution as $t \to \infty$. Once this has been established, various 'laws of large numbers' for the random density $\rho_t$ can be proved. In particular, we prove a weak law of large numbers, assuming nothing more than the finiteness of the speed measure. An interesting consequence of this result is the fact that a regular diffusion with finite speed measure is necessarily recurrent. It seems that this basic fact has not been explicitly noted before in the literature, at least not for general regular diffusions.

If $\nu$ is a signed measure with finite total variation on the state space, we prove that

$$\int (\rho_t - 1) \, d\nu \to 0$$

in probability. The attention is then turned to the corresponding uniform central limit problem. We consider a (possibly infinite) collection $\Lambda$ of signed measures on the state space and give conditions that imply weak convergence of the random maps

$$\nu \mapsto \sqrt{t} \int (\rho_t - 1) \, d\nu$$
in the space $\ell^{\infty}(\Lambda)$ of bounded functions on $\Lambda$. The most important requirement is a metric entropy condition. This is a mathematical formulation of the intuitively obvious fact that the collection $\Lambda$ should not be 'too large' if we want to have uniform convergence on $\Lambda$. Nishiyama's entropy methods for continuous martingales play a key role in the proof of our central limit theorem.

Chapter 2 is closed with a section that gives some interesting consequences and special cases of the general results mentioned above. It turns out for instance that a uniform law of large numbers is valid under a very mild condition. If $\mathcal{F}$ is a class of functions on the state space and $\mathcal{F}$ has a $\mu$-integrable envelope, then it holds that

$$\sup_{f \in \mathcal{F}} \left| \frac{1}{t} \int_0^t f(X_u) \, du - \int f \, d\mu \right| \to 0$$

in (outer) probability. Note that the stated requirement on the class $\mathcal{F}$ is much weaker than the conditions that are needed for the i.i.d. counterpart of this result (see for instance Van der Vaart and Wellner (1996)). For certain special classes of diffusions, a number of uniform limit theorems was recently given by Kutoyants (1998) and Negri (1998). We extend these results to the more general setting of regular diffusions.

In chapter 3 we use the limit theory of the first two chapters to study nonparametric estimators for regular diffusions. We consider a stationary, regular diffusion $X$ with finite speed measure. The normalized speed measure $\mu$ is then the marginal distribution of the process. It is assumed that $\mu$ has a density $f$ and the objective is to estimate $f$ and its derivatives $f^{(m)}$ (if they exist), based on the observation of a sample path $(X_u)_{u \leq t}$. The standard kernel estimator for $f$ is defined by

$$\hat{f}_{t,h}(x) = \frac{1}{ht} \int_0^t K \left( \frac{x - X_u}{h} \right) \, du,$$

where $K$ is some appropriate kernel function and the parameter $h > 0$ is the bandwidth. Obvious estimators for the derivatives $f^{(m)}$ of $f$ are obtained by differentiating this expression. If the kernel $K$ has an $m$-th derivative $K^{(m)}$, we put

$$\hat{f}^{(m)}_{t,h}(x) = \frac{1}{h^{m+1}t} \int_0^t K^{(m)} \left( \frac{x - X_u}{h} \right) \, du.$$ 

These kernel estimators have been studied by several authors. The cases $m = 0$ and $m = 1$ were first considered by Banon (1978). In Banon's
paper, conditions are given for pointwise mean-square consistency. For \( m = 0 \), uniform consistency was proved by Nguyen (1979). Kutoyants (1998) recently studied the kernel estimator for the density \( f \) for a special class of diffusions generated by stochastic differential equations. He proved pointwise consistency, asymptotic normality with rate \( \sqrt{t} \) and asymptotic efficiency. Note that the rate is independent of the bandwidth, so the situation is quite different from the i.i.d. setting.

We further investigate the asymptotic properties of the kernel estimators and obtain a number of new results. Firstly, the asymptotic normality for \( m = 0 \) turns out to hold uniformly in \( x \). More precisely, for every compact subinterval \( J \) of the state space we have weak convergence

\[
\sqrt{t} (\hat{f}_{t,h} - f) \Rightarrow H
\]

in the space \( C(J) \) of continuous functions on \( J \), where \( H \) is a Gaussian random map in \( C(J) \). The estimator of the \( m \)-th derivative \( f^{(m)} \) of \( f \) is also uniformly consistent, but its asymptotic distributional properties are quite different. Under some conditions on the bandwidths, we show that

\[
\sup_{x \in J} \left| \frac{\hat{f}^{(m)}_{t,h_t}(x)}{h_t^{m+1}} - f^{(m)}(x) \right| = o_P \left( \frac{1}{h_t^m \sqrt{t}} \right)
\]

for every compact subinterval \( J \) of the state space. Clearly, this does not yet give the exact rate of the estimator, it only implies that it must be faster than \( h_t^m \sqrt{t} \). The exact rate turns out to be \( \sqrt{t} h_t^{2m-1} \), but the convergence is not uniform at this rate. Instead we prove that for different \( x \), the normalized differences

\[
\sqrt{t} h_t^{2m-1} (\hat{f}^{(m)}_{t,h_t}(x) - f^{(m)}(x))
\]

converge weakly to independent Gaussian random variables.

Solutions of stochastic differential equation are probably the most important examples of diffusions. In the chapter 4 we study the problem of drift estimation for such processes. We first consider the parametric model

\[
dX_t = b_\theta(X_t) \, dt + \sigma(X_t) \, dW_t,
\]

where \( \theta \) ranges over a subset \( \Theta \) of Euclidean space. The classical estimator for the true parameter in this model is the maximum likelihood estimator. The maximum likelihood method for nonlinear diffusion models has been studied by many authors, see for instance Kutoyants (1977), Lánská (1979), Prakasa Rao and Rubin (1981), Basu (1983), Kutoyants (1984), Yoshida
(1990) and Prakasa Rao (1999). It turned out to be quite difficult to derive the asymptotic properties of the maximum likelihood estimator under reasonable regularity conditions. In particular, most authors needed conditions that depend on the dimension of the parameter space $\Theta$. In the first part of this chapter we show that this is not necessary. Using uniform limit theorems from chapters 1 and 2, the continuity of the likelihood and consistency of the estimator are proved under a Hölder condition that is independent of the dimension of the parameter space. In other branches of statistics the application of entropy methods lead to a better understanding of asymptotic behaviour, see for instance Van de Geer (1995) and Nishiyama (1999). Our results in the first part of chapter 4 show that this approach is also very useful in asymptotic statistics for diffusions.

In the second part of the chapter 4 we study a nonparametric drift estimator that was proposed by Banon (1978). Under some technical conditions, the drift function $b$ and the diffusion function $\sigma$ are related to the density $f$ of the speed measure and its derivative $f'$ by the equation

$$b(x) = \frac{1}{2} \sigma^2(x) \frac{f'(x)}{f(x)} + \sigma(x)\sigma'(x).$$

The drift estimator $\hat{b}_{t,h}$ that Banon proposed is obtained by simply plugging in the kernel estimators for $f$ and $f'$ that are studied in chapter 3. Banon only proved that this estimator is pointwise consistent. Using the delta-method it is relatively easy to derive additional asymptotic properties from the results of chapter 3. We prove that the estimator is uniformly consistent, that the exact rate of convergence is $\sqrt{th_t}$ and that for different $x$, the normalized differences

$$\sqrt{th_t}(\hat{b}_{t,h_t}(x) - b(x))$$

converge weakly to independent Gaussian random variables.

**Overview of Part II**

In the chapter 5 we study Bernstein-type inequalities for martingales with jumps. Such inequalities have many applications in probability and statistics. See for instance Shorack and Wellner (1986) and Van de Geer (1995) for applications in empirical process theory for i.i.d. and dependent data. Other
applications include the study of the rate of convergence in the functional central limit theorem for martingales (see Coquet et al. (1994) and Courboit (1998)) and the theory of decoupling (see De la Peña and Giné (1999) and De la Peña (1999)).

The work presented in chapter 5 unifies a number of different known inequalities. The Bernstein inequality for martingales with bounded jumps states that if the jumps of the local martingale $M$ are bounded in absolute value by a constant $a > 0$, then for every finite stopping time $\tau$ it holds that

$$P \left( \sup_{t \leq \tau} |M_t| \geq z, \langle M \rangle_\tau \leq L \right) \leq 2e^{-\frac{1}{2} L + az^3}$$

for every $z, L > 0$. On the other hand, Barlow et al. (1986) proved that if $M$ is locally square integrable (not necessarily with bounded jumps) then

$$P \left( \sup_{t \leq \tau} |M_t| \geq z, \langle M \rangle_\tau + \sum_{t \leq \tau} (\Delta M_t)^2 \leq L \right) \leq 2e^{-\frac{1}{2} \frac{z^2}{L}}$$

for every $z, L > 0$. Both inequalities may be applied if $M$ is a local martingale with bounded jumps, but the results are clearly different.

It turns out however that the two inequalities can be viewed as extreme cases of a more general result. If $M$ is locally square integrable martingale and $a > 0$, we can consider the process

$$H_t^a = \sum_{s \leq t} (\Delta M_s)^2 1_{\{\Delta M_s \leq a\}} + \langle M \rangle_t.$$

We prove in chapter 5 that for every finite stopping time $\tau$ it holds that

$$P \left( \sup_{t \geq \tau} |M_t| \geq z, H_\tau^a \leq L \right) \leq 2e^{-\frac{1}{2} L + az^3}$$

for every $a, z, L > 0$. The inequality for martingales with bounded jumps is then obtained if we choose $a$ such that $|\Delta M| \leq a$. If we take $a = 0$ the result reduces to the inequality of Barlow et al. (1986). The method we use in the proof of our general inequality can be used to prove several other useful exponential inequalities for general martingales. To illustrate this, we extend results of Van de Geer (1995) and Nishiyama (2000) for martingales that satisfy an ‘exponential moment condition’ and a result of De la Peña (1999) for conditionally symmetric martingales.
In the final chapter 6, a central limit theorem for martingales with jumps is presented. In chapter 1 the central limit theorem for continuous martingales is proved by using the Dambis-Dubins-Schwarz time-change theorem. If the martingales in question have jumps this method cannot be used. In chapter 6, we show that it is possible to use another time-change device in this case. Instead of the Dambis-Dubins-Schwarz theorem, we use a general version of Skorohod's embedding theorem to prove a central limit theorem for normalized martingales with jumps. The usual approach relies on characteristic function-type methods or rather the 'method of stochastic exponentials' (see Hall and Heyde (1980), Liptser and Shiryayev (1989), Feigin (1985), Jacod and Shiryaev (1987)). In the literature on central limit theory, the Skorohod embedding theorem has primarily been used to investigate the rate of convergence of the (functional) central limit theorem for martingales (see Hall and Heyde (1980), Kubilius (1985), Coquet et al. (1994) and Courbot (1998)). The embedding technique does not seem suitable to prove central limit theorems of the same generality as can be done with the characteristic function methods. However, in the particular setup of chapter 6 it yields a short and elegant proof of the central limit theorem.