Centrum voor Wiskunde en Informatica

REPORT*RAPPORT*

On the Uniform Convergence of Local Time and the Uniform Consistency of Density Estimators for Ergodic Diffusions

J.H. van Zanten

Probability, Networks and Algorithms (PNA)

PNA-R9909 September 30, 1999

Report PNA-R9909
ISSN 1386-3711

CWI
P.O. Box 94079
1090 GB Amsterdam
The Netherlands

On the Uniform Convergence of Local Time and the Uniform Consistency of Density Estimators for Ergodic Diffusions

J.H. van Zanten
CWI
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
hvz@cwi.nl

ABSTRACT

We prove a theorem on the uniform convergence of the local time of an ergodic diffusion. This result is then used to investigate certain estimators of the invariant density of an ergodic diffusion, including kernel estimators. We show that the pointwise consistency of these estimators can be strengthened to uniform consistency.

1991 Mathematics Subject Classification: 60J55, 60J60, 60H10, 62G07.
Keywords and Phrases: Ergodic diffusions, local time, density estimation.
Note: Work carried out under project PNA3.3 'Stochastic Processes and Applications'.

1 Introduction

This paper is concerned with the local time of a stationary, ergodic diffusion and its relation to the invariant probability measure of such a diffusion. More precisely, we study the semimartingale local time $\{L_t(x) : t \geq 0, x \in \mathbb{R}\}$ of a stationary, ergodic diffusion X that solves the stochastic differential equation (SDE) of the form (1). We prove that the random function

$$x \mapsto \frac{2L_t(x)}{t\sigma^2(x)}$$

converges in probability to the density of the invariant probability measure, uniformly over compact intervals (see theorem 5.1). A number of nonparametric estimators for the density of the invariant probability measure that have been proposed in the literature can be expressed in terms of the local time of the diffusion. As a result, the uniform convergence of the local time allows us to prove uniform consistency results for some density estimators, including kernel estimators and the unbiased density estimators proposed by Kutoyants (1997a, 1997b) (see theorems 6.1 and 6.2).

The connection between the local time of a stationary process and density estimators was already noted by several authors, see e.g. Bosq and Davydov (1999) and the references therein and Kutoyants (1997a). Kutoyants (1997a) is concerned with the same model as we are. He studies the properties of density estimators at a fixed point. Kernel estimators for the invariant density for the model (1) were in fact already studied by Banon (1978) and Nguyen (1979). The relation with local time was however not yet observed in those papers. A result concerning the uniform convergence of the local time of general stationary, ergodic continuous time processes appears in the paper of Bosq and Davydov (1999). Their theorem requires a condition on the modulus of continuity of the local time of the process involved, and is therefore not directly applicable in our situation. The results of the present paper require conditions on the coefficients b and σ of the SDE (1) and the invariant measure of the diffusion, not on the local time itself.

For our asymptotic analysis of local time we use the representation given by the Tanaka-Meyer formula (6). The Brownian integral I_t and the Lebesgue integral J_t appearing in (6) are first studied separately. For the Brownian integral we prove a weak convergence result that is in fact stronger than we need, and that is interesting in its own right (see theorem 3.3). For the Lebesgue integral, we note that a straightforward analogue of a uniform law of large numbers for i.i.d. random variables holds in the context of ergodic diffusions (see lemma 4.1). In the special case of the uniform convergence of the empirical distribution function of an ergodic diffusion this was already noted by Kutoyants (1997c).

This paper is organized as follows. Section 2 gives the general setup of the paper. We introduce some notation and recall a few basic facts concerning ergodic diffusions and semimartingale local time. In sections 3 and 4 we study the components (7) and (8) of the local time, respectively. We obtain a weak convergence result for the first component in section 3 and a Glivenko-Cantelli theorem for the second component in section 4. These results are then combined in section 5 where we prove the main result of the paper concerning uniform convergence of local time. In section 6 we apply the main theorem to prove uniform consistency of some estimators of the invariant probability density of the ergodic diffusion.

2 Setup and preliminaries

Consider the stochastic differential equation

$$dX_t = b(X_t)\,dt + \sigma(X_s)\,dW_t, \tag{1}$$

where W is a standard Brownian motion and b and $\sigma > 0$ are continuous functions that satisfy certain conditions that ensure the existence of a unique strong solution for every initial condition (for example the usual Lipschitz and linear growth conditions, see Karatzas and Shreve (1991)). Recall that the derivative of the scale function associated to the SDE (1) is the function

s defined by

$$s(x) = \exp\left(-2\int_0^x \frac{b(y)}{\sigma^2(y)}\,dy\right). \tag{2}$$

We suppose that

$$\lim_{|x|\to\infty} s(x) = \infty \quad \text{and} \quad D = \int_{\mathbb{R}} \frac{1}{\sigma^2(x)s(x)}\,dx < \infty \tag{3}$$

and we define the probability measure μ on $\mathbb{R}$ by $\mu(dx) = \pi(x)\,dx$, where

$$\pi(x) = \frac{1}{D\sigma^2(x)s(x)}.$$

Note that (2) implies the relation

$$(\sigma^2\pi)' = 2b\pi. \tag{4}$$

It is well-known (see e.g. Gihman and Skorohod (1972)) that the condition (3) implies that the ergodic theorem holds for the strong solutions of (1), i.e. that for every $f \in L^1(\mu)$

$$\frac{1}{t}\int_0^t f(X_s)\,ds \overset{\text{as}}{\to} \int f\,d\mu$$

as $t \to \infty$. Moreover, the strong solution of SDE (1) that satisfies the initial condition $\mathcal{L}(X_0) = \mu$ is stationary (see Gihman and Skorohod (1972), p. 138). Throughout the paper we denote by X this stationary solution.

The process X is a continuous semimartingale, so its semimartingale local time $L = \{L_t(x) : t \geq 0, x \in \mathbb{R}\}$ is well-defined. We will use the fact that the random field L has the following properties (see Karatzas and Shreve (1991), theorem 7.1):

(i) For every nonnegative, measurable function f

$$\int_0^t f(X_s)\sigma^2(X_s)\,ds = 2\int_{\mathbb{R}} f(x)L_t(x)\,dx. \tag{5}$$

(ii) For every $t \geq 0$ and $x \in \mathbb{R}$

$$L_t(x) = (X_t - x)^- - (X_0 - x)^- + I_t(x) + J_t(x), \tag{6}$$

where

$$I_t(x) = \int_0^t 1_{(-\infty,x]}(X_s)\sigma(X_s)\,dW_s \tag{7}$$

and

$$J_t(x) = \int_0^t 1_{(-\infty,x]}(X_s)b(X_s)\,ds. \tag{8}$$

The relation (6) is known as (a version of) the Tanaka-Meyer formula. In the next two sections we study the asymptotics of the random functions I_t and J_t occurring in (6).

3 A weak convergence theorem

By lemma 3.5.7 of Karatzas and Shreve (1991) there exists a jointly continuous modification of the random field $\{I_t(x) : t \geq 0, x \in \mathbb{R}\}$. The symbol $I_t(x)$ will always refer to this continuous version. In particular, the random function $I_t : x \mapsto I_t(x)$ is a continuous function for every $t \geq 0$. So for every $A > 0$, the restriction of the random function $I_t/\sqrt{t}$ to $[-A, A]$ is a random element of the Banach space $C[-A, A]$ of continuous functions on $[-A, A]$ (endowed with the supremum metric). We will prove that under certain integrability conditions, the random functions $I_t/\sqrt{t}$ have a weak limit in $C[-A, A]$. The first step is the following lemma. Note that the proof is similar to that of lemma 3.5.7 of Karatzas and Shreve (1991).

Lemma 3.1. *For $p \geq 2$, suppose that*

$$\int |x|^p \, \mu(dx) < \infty, \qquad \int |b(x)|^p \, \mu(dx) < \infty, \qquad \int |\sigma(x)|^p \, \mu(dx) < \infty.$$

Then there exists a constant $C = C(p, b, \sigma) > 0$ such that

$$\mathrm{E} \left| \frac{1}{\sqrt{t}} \int_0^t 1_{(x,y]}(X_s)\sigma(X_s) \, dW_s \right|^{2p} \leq C|x-y|^p$$

for every $x < y$ and $t \geq 1$.

Proof. Let f be the function given by

$$f(u) = \int_0^u \int_0^v 1_{(x,y]}(z) \, dzdv.$$

Note that $|f'| \leq |x-y|$. Applying the generalized Itô formula to the convex function f of the semimartingale X we get

$$\tfrac{1}{2} \int_0^t 1_{(x,y]}(X_s)\sigma^2(X_s) \, ds = f(X_t) - f(X_0) - \int_0^t f'(X_s)b(X_s) \, ds - \int_0^t f'(X_s)\sigma(X_s) \, dW_s.$$

Using the triangle inequality and the fact that $(u+v+w)^p \leq 3^p(u^p+v^p+w^p)$ for $u, v, w \geq 0$ and $p \geq 1$ it follows that

$$\left| \int_0^t 1_{(x,y]}(X_s)\sigma^2(X_s) \, ds \right|^p \leq 6^p \, |f(X_t) - f(X_0)|^p + 6^p \left| \int_0^t f'(X_s)b(X_s) \, ds \right|^p + 6^p \left| \int_0^t f'(X_s)\sigma(X_s) \, dW_s \right|^p.$$

Let us consider the expectations of the three terms on the right hand side. Since $|f'| \le |x-y|$ we have

$$|f(X_t) - f(X_0)|^p \le |x-y|^p |X_t - X_0|^p \le 2^p |x-y|^p (|X_t|^p + |X_0|^p).$$

So by the stationarity of X

$$\mathrm{E}\ |f(X_t) - f(X_0)|^p \le 2^{p+1} |x-y|^p \int |z|^p\, \mu(dz).$$

For the second term, it holds that

$$\begin{aligned}\mathrm{E}\left|\int_0^t f'(X_s) b(X_s)\, ds\right|^p &\le t^{p-1} |x-y|^p \mathrm{E} \int_0^t |b(X_s)|^p\, ds \\ &= t^p |x-y|^p \int |b(z)|^p\, \mu(dz).\end{aligned}$$

As for the third term, note that by the Burkholder-Davis-Gundy inequalities there exists a constant $D_p > 0$ such that

$$\begin{aligned}\mathrm{E}\left|\int_0^t f'(X_s)\sigma(X_s)\, dW_s\right|^p &\le D_p \mathrm{E}\left|\int_0^t \big(f'(X_s)\sigma(X_s)\big)^2\, ds\right|^{p/2} \\ \le t^{p/2-1} D_p |x-y|^p \mathrm{E} \int_0^t |\sigma(X_s)|^p\, ds &= t^{p/2} D_p |x-y|^p \int |\sigma(z)|^p\, \mu(dz).\end{aligned}$$

Combining these bounds we find that there exists a positive constant $B = B(p, b, \sigma)$ such that

$$\mathrm{E}\left|\int_0^t 1_{(x,y]}(X_s)\sigma^2(X_s)\, ds\right|^p \le t^p B |x-y|^p$$

for every $t \ge 1$. Another application of the Burkholder-Davis-Gundy inequalities gives a constant $A_p > 0$ such that

$$\begin{aligned}\mathrm{E}\left|\int_0^t 1_{(x,y]}(X_s)\sigma(X_s)\, dW_s\right|^{2p} &\le A_p \mathrm{E}\left|\int_0^t 1_{(x,y]}(X_s)\sigma^2(X_s)\, ds\right|^p \\ &\le t^p A_p B |x-y|^p.\end{aligned}$$

for every $t \ge 1$. The proof is concluded by dividing both sides of this inequality by t^p. □

It is well-known that tightness of a collection of random functions in the space $C[-A, A]$ is equivalent to pointwise tightness and asymptotic equicontinuity (see e.g. Billingsley (1968), theorem 8.2). Using lemma 3.1, we get the following result concerning the equicontinuity of the random functions $x \mapsto I_t(x)/\sqrt{t}$.

Lemma 3.2. *Suppose that*

$$\int |x|^2\,\mu(dx) < \infty, \qquad \int |b(x)|^2\,\mu(dx) < \infty, \qquad \int |\sigma(x)|^2\,\mu(dx) < \infty.$$

Then for every $A > 0$ and every $\varepsilon, \eta > 0$ there exists a $\delta > 0$ such that

$$\mathrm{P}\left(\sup_{\substack{x,y\in[-A,A]\\ |x-y|<\delta}} \tfrac{1}{\sqrt{t}} |I_t(x) - I_t(y)| \geq \varepsilon \right) \leq \eta$$

for every $t \geq 1$.

Proof. By lemma 3.1 there exists a constant $C > 0$ such that

$$\left\| \tfrac{1}{\sqrt{t}} (I_t(x) - I_t(y)) \right\|_{L^4(\mathrm{P})} \leq C\sqrt{|x-y|}$$

for every $t \geq 1$ and $x, y \in \mathbb{R}$. It then follows from theorem 2.2.4 of Van der Vaart and Wellner (1996) that there exists a constant $K = K(A) > 0$ such that

$$\mathrm{E} \sup_{\substack{x,y\in[-A,A]\\ |x-y|<\delta}} \tfrac{1}{\sqrt{t}} |I_t(x) - I_t(y)| \leq K\left(\eta^{3/4} + \delta^2 \eta^{-1/2}\right)$$

for every $t \geq 1$ and $\delta, \eta > 0$. Now use Markov's inequality to complete the proof. □

We arrive at the following theorem concerning the weak convergence of the random maps $I_t/\sqrt{t}$.

Theorem 3.3. *Suppose that*

$$\int |x|^2\,\mu(dx) < \infty, \qquad \int |b(x)|^2\,\mu(dx) < \infty, \qquad \int |\sigma(x)|^2\,\mu(dx) < \infty.$$

Then for every $A > 0$ the restrictions of the random functions $I_t/\sqrt{t}$ to $[-A, A]$ converge weakly in $C[-A, A]$ to a zero mean Gaussian random function I with covariance function

$$\mathrm{E}\, I(x)I(y) = \int_{-\infty}^{x\wedge y} \sigma^2(z)\,\mu(dz).$$

Proof. To establish finite dimensional convergence, take $x_1, \ldots, x_d \in \mathbb{R}$. The process $(I(x_1), \ldots, I(x_d))$ is a d-dimensional martingale and by the ergodic theorem we have

$$\frac{1}{t}\,\langle I(x_i), I(x_j)\rangle_t = \frac{1}{t}\int_0^t 1_{(-\infty, x_i\wedge x_j]}(X_s)\sigma^2(X_s)\,ds \overset{\text{as}}{\to} \int_{-\infty}^{x_i\wedge x_j} \sigma^2(x)\,\mu(dx).$$

So the central limit theorem for continuous local martingales (see for instance Van Zanten (1998), theorem 5.1) confirms that the marginals of $I_t/\sqrt{t}$ converge weakly to the marginals of I. By lemma 3.2 the restrictions of the random functions $I_t/\sqrt{t}$ to $[-A, A]$ are asymptotically equicontinuous and therefore tight in $C[-A, A]$ (see Billingsley (1968), theorem 8.2). It then follows from Prohorov's theorem that we have the desired weak convergence. $\square$

4 A uniform ergodic theorem

Let us recall the definition of the bracketing numbers of a normed function space. Given two functions l and u, the bracket $[l, u]$ is the set of all functions f with $l \leq f \leq u$. An ε-bracket is a bracket $[l, u]$ such that $\|u - l\| < \varepsilon$. The bracketing number $N_{[\,]}(\varepsilon, \mathcal{F}, \|\cdot\|)$ is the minimum number of ε-brackets needed to cover $\mathcal{F}$.

The following lemma is the analogue of a classical uniform law of large numbers for i.i.d. random variables that can be found for instance in Van de Geer (1999). It gives a condition in terms of bracketing numbers under which the ergodic theorem holds uniformly over a class of functions. The lemma can be proved by a straightforward modification of the proof of lemma 3.1 of Van de Geer (1999).

Lemma 4.1. *Let $\mathcal{F}$ be a class of measurable functions and suppose that $N_{[\,]}(\varepsilon, \mathcal{F}, \|\cdot\|_{L^1(\mu)}) < \infty$ for every $\varepsilon > 0$. Then*

$$\sup_{f \in \mathcal{F}} \left| \frac{1}{t} \int_0^t f(X_s)\, ds - \int f\, d\mu \right| \stackrel{\text{as}}{\to} 0.$$

In general, the supremum occurring in the lemma does not have to be measurable. In the non-measurable case the lemma also holds, but almost sure convergence should then be understood in the sense of Van der Vaart and Wellner (1996), definition 1.9.1. In this paper we only use the lemma in the proof of theorem 4.3, which involves a measurable supremum. Indeed, since the random functions J_t defined by (8) are càdlàg (see Karatzas and Shreve (1991), p. 222) the supremum in (9) does not change when we take it over a countable dense subset of $\mathbb{R}$. This implies that the supremum is in fact measurable. We will use lemma 4.1 in conjunction with the following simple result concerning the bracketing entropy of a certain class of functions.

Lemma 4.2. *Let $f \in L^1(\mu)$ be a nonnegative function. Then the class $\mathcal{F} = \{1_{(-\infty,x]} f : x \in \mathbb{R}\}$ satisfies $N_{[\,]}(\varepsilon, \mathcal{F}, \|\cdot\|_{L^1(\mu)}) < \infty$ for every $\varepsilon > 0$.*

Proof. For ease of notation, define the function $f_x = 1_{(-\infty,x]} f$ for every $x \in \mathbb{R}$. Moreover, we define the finite measure ν on $\mathbb{R}$ by putting $d\nu = f\, d\mu$.

Now fix $\varepsilon > 0$. The fact that ν is finite implies that we can find a finite number of points $-\infty = x_0 < x_1 \cdots < x_n = \infty$ such that $\nu(x_i, x_{i+1}] < \varepsilon$ for every i. Since $x \le y$ implies $f_x \le f_y$, the brackets $[f_{x_i}, f_{x_{i+1}}]$ cover $\mathcal{F}$. By construction the bracket $[f_{x_i}, f_{x_{i+1}}]$ is an ε-bracket, since

$$\int (f_{x_{i+1}} - f_{x_i})\, d\mu = \int 1_{(x_i, x_{i+1}]} f\, d\mu = \nu(x_i, x_{i+1}] < \varepsilon.$$

This conclude the proof of the lemma. □

Lemma's 4.1 and 4.2 give us the following result concerning the random maps J_t.

Theorem 4.3. *Suppose that* $\int |b(x)|\, \mu(dx) < \infty$. *Then*

$$\sup_{x\in\mathbb{R}} \left| \tfrac{2}{t} J_t(x) - \sigma^2(x)\pi(x) \right| \overset{\text{as}}{\to} 0. \tag{9}$$

Proof. Denote by b^+ and b^- the positive and negative parts of b, so $b = b^+ - b^-$. It clearly holds that

$$\begin{aligned}\sup_{x\in\mathbb{R}} \left| \frac{1}{t}\int_0^t 1_{(-\infty,x]}(X_s) b(X_s)\, ds - \int_{-\infty}^x b(y)\, \mu(dy) \right| \le \\ \sup_{x\in\mathbb{R}} \left| \frac{1}{t}\int_0^t 1_{(-\infty,x]}(X_s) b^+(X_s)\, ds - \int_{-\infty}^x b^+(y)\, \mu(dy) \right| \\ + \sup_{x\in\mathbb{R}} \left| \frac{1}{t}\int_0^t 1_{(-\infty,x]}(X_s) b^-(X_s)\, ds - \int_{-\infty}^x b^-(y)\, \mu(dy) \right|\end{aligned} \tag{10}$$

By lemma's 4.1 and 4.2, the two terms on the right hand side of (10) converge almost surely to 0. It thus remains to show that

$$2\int_{-\infty}^x b(y)\, \mu(dy) = \sigma^2(x)\pi(x).$$

But this follows immediately from the relation (4). □

5 Convergence of local time

Let B be a Borel subset of $\mathbb{R}$. Taking $f = 1_B/\sigma^2$ in (5) we see that for every $t > 0$

$$\frac{1}{t}\int_0^t 1_B(X_s)\, ds = \int_B \frac{2L_t(x)}{t\sigma^2(x)}\, dx.$$

In other words, the random function

$$x \mapsto \pi_t(x) = \frac{2L_t(x)}{t\sigma^2(x)}$$

is a density of the empirical measure

$$B \mapsto \frac{1}{t}\int_0^t 1_B(X_s)\,ds$$

with respect to the Lebesgue measure on $\mathbb{R}$.

By the ergodic theorem and (5) we have the convergence

$$\int_{\mathbb{R}} f(x)\pi_t(x)\,dx \stackrel{\text{as}}{\to} \int_{\mathbb{R}} f(x)\pi(x)\,dx$$

for every $f \in L^1(\mu)$. It follows that almost surely, we have the weak convergence $\pi_t \rightsquigarrow \pi$.

More can be said about the convergence of π_t to π. Recall the representation of the local time given by the Tanaka-Meyer formula (6). It follows from the law of large numbers for martingales, the ergodic theorem and the relation (4) that for every $x \in \mathbb{R}$ we have the convergence

$$\frac{2L_t(x)}{t} \stackrel{\text{as}}{\to} \int_{-\infty}^x 2b(y)\pi(y)\,dy = \sigma^2(x)\pi(x).$$

Since the local time $L_t(x)$ is càdlàg in x and σ is continuous, the random density π_t is also càdlàg and it follows that almost surely, we have pointwise convergence of π_t to π. By Scheffé's theorem it then follows that we also almost surely have the convergence of π_t to π in $L^1(\mu)$. The following theorem is a consequence of theorems 3.3 and 4.3 and concerns the uniform convergence of local time.

Theorem 5.1. *Suppose that*

$$\int |x|^2\,\mu(dx) < \infty, \qquad \int |b(x)|^2\,\mu(dx) < \infty, \qquad \int |\sigma(x)|^2\,\mu(dx) < \infty.$$

Then

$$\sup_{|x|\leq A} \left|\tfrac{2}{t}L_t(x) - \sigma^2(x)\pi(x)\right| \stackrel{\text{P}}{\to} 0 \quad \textit{and} \quad \sup_{|x|\leq A} |\pi_t(x) - \pi(x)| \stackrel{\text{P}}{\to} 0$$

for every $A > 0$.

Proof. The second statement follows from the first one upon noting that the continuous positive function σ^2 is bounded away from zero on the compact interval $[-A, A]$. So it remains to prove the first statement. By the Tanaka-Meyer formula (6) we have

$$\begin{aligned}\tfrac{2}{t}L_t(x) - \sigma^2(x)\pi(x) &= \\ &\tfrac{2}{t}(X_t - x)^- - \tfrac{2}{t}(X_0 - x)^- + \tfrac{2}{t}I_t(x) + \tfrac{2}{t}J_t(x) - \sigma^2(x)\pi(x).\end{aligned}$$

for every $x \in \mathbb{R}$. So for $|x| \leq A$ it holds that

$$
\begin{aligned}
&|\tfrac{2}{t} L_t(x) - \sigma^2(x)\pi(x)| \leq \\
&\qquad \tfrac{2}{t}\left(|X_t| + |X_0| + 2A\right) + \tfrac{2}{t}|I_t(x)| + \left|\tfrac{2}{t} J_t(x) - \sigma^2(x)\pi(x)\right|.
\end{aligned} \tag{11}
$$

Let us now consider the three terms on the right hand side of (11). By the stationarity of X, the expectation of the first term is bounded by $4(\int |x|\,\mu(dx) + A)/t$, so the first term converges to 0 in mean, hence also in probability. It follows from theorem 3.3 that the restrictions of the random functions $I_t/\sqrt{t}$ to $[-A, A]$ have a weak limit in $C[-A, A]$. It follows that the restrictions of the random functions $2I_t/t$ to $[-A, A]$ converge weakly to 0 in $C[-A, A]$. This is equivalent to

$$
\sup_{|x|\leq A} \tfrac{2}{t}|I_t(x)| \xrightarrow{\mathrm{P}} 0,
$$

which covers the second term. The convergence of the third term in (11) follows from theorem 4.3. □

6 Applications

6.1 Kernel estimators

Let K be a probability density on $\mathbb{R}$ with compact support. We investigate the uniform consistency of the kernel estimator $\hat{\pi}_t$ of the invariant density defined by

$$
\hat{\pi}_t(x) = \frac{1}{th_t} \int_0^t K\left(\frac{X_s - x}{h_t}\right) ds,
$$

where the h_t are positive numbers such that

$$
h_t \downarrow 0 \quad \text{as} \quad t \to \infty. \tag{12}
$$

Theorem 6.1. *Under the conditions of theorem 5.1 and (12) we have*

$$
\sup_{|x|\leq A} |\hat{\pi}_t(x) - \pi(x)| \xrightarrow{\mathrm{P}} 0
$$

for every $A > 0$.

Proof. Clearly we have the relation

$$
|\hat{\pi}_t(x) - \pi(x)| \leq |\hat{\pi}_t(x) - \mathrm{E}\,\hat{\pi}_t(x)| + |\mathrm{E}\,\hat{\pi}_t(x) - \pi(x)| \tag{13}
$$

We will show that the supremum for $|x| \leq A$ of both of the terms on the right hand side of (13) converges to 0 in probability. For the first term, note that by relation (5) and the stationarity of X

$$
\begin{aligned}
|\hat{\pi}_t(x) - \mathrm{E}\,\hat{\pi}_t(x)| &\leq \frac{1}{h_t}\int_{\mathbb{R}} K\left(\frac{y-x}{h_t}\right)|\pi_t(y) - \pi(y)|\,dy \\
&= \int_{\mathbb{R}} K(u)|\pi_t(x+h_t u) - \pi(x+h_t u)|\,du.
\end{aligned}
$$

Now say that $\mathrm{supp}(K) \subseteq [-B, B]$ for $B > 0$. Then it follows that for every $x \in \mathbb{R}$ and $t > 0$

$$
|\hat{\pi}_t(x) - \mathrm{E}\,\hat{\pi}_t(x)| \leq \sup_{|u|\leq B} |\pi_t(x+h_t u) - \pi(x+h_t u)|\,.
$$

So for t large enough to ensure that $h_t \leq 1$ we have

$$
\begin{aligned}
\sup_{|x|\leq A} |\hat{\pi}_t(x) - \mathrm{E}\,\hat{\pi}_t(x)| &\leq \sup_{|x|\leq A, |u|\leq B} |\pi_t(x+h_t u) - \pi(x+h_t u)| \\
&\leq \sup_{|x|\leq A+B} |\pi_t(x) - \pi(x)|,
\end{aligned}
$$

which converges to 0 in probability by theorem 5.1. It remains to show that the second term on the right hand side of (13) converges to 0, uniformly for $|x| \leq A$. Using the same arguments that we used to bound the first term, we see that it is enough to show that

$$
\sup_{|x|\leq A, |u|\leq B} |\pi(x+h_t u) - \pi(x)| \to 0.
$$

But this follows readily from (12) and the fact that π is continuous and therefore uniformly continuous on compact intervals. □

6.2 Other density estimators

Under the assumption that the diffusion coefficient σ is known, some unbiased estimators for the invariant density π were proposed and studied by Kutoyants (1997a, 1997b). The most basic one is the estimator $\tilde{\pi}_t$ defined by

$$
\tilde{\pi}_t(x) = \frac{2}{t\sigma^2(x)}\int_0^t 1_{(-\infty,x]}(X_s)\,dX_s.
$$

We will prove the uniform consistency of this estimator. The uniform consistency of the other estimators in Kutoyants (1997b) can be shown in a similar manner.

Theorem 6.2. *Under the conditions of theorem 5.1 we have*

$$\sup_{|x|\le A} |\tilde{\pi}_t(x) - \pi(x)| \xrightarrow{\mathrm{P}} 0$$

for every $A > 0$.

Proof. By (1) and the Tanaka-Meyer formula (6) we have

$$\int_0^t 1_{(-\infty,x]}(X_s)\,dX_s = L_t(x) - (X_t - x)^- + (X_0 - x)^-.$$

It follows that

$$\tilde{\pi}_t(x) = \pi_t(x) - \frac{2}{\sigma^2(x)}\frac{1}{t}\left((X_t - x)^- - (X_0 - x)^-\right),$$

which implies the relation

$$|\tilde{\pi}_t(x) - \pi(x)| \le |\pi_t(x) - \pi(x)| + \frac{2}{\sigma^2(x)}\frac{1}{t}\left(|X_t| + |X_0| + 2x\right).$$

Theorem 5.1 gives the desired convergence of the first term. For the convergence of the second term, use the fact that σ^2 is bounded away from 0 on a compact interval and the stationarity of X. □

References

[1] G. Banon (1978), Nonparametric Identification for Diffusion Processes, *Siam J. Contr. Opt.* **16**, no. 3, 380–395.

[2] P. Billingsley (1968), *Convergence of Probability Measures*, Wiley.

[3] D. Bosq and Yu. Davydov (1999), Local Time and Density Estimation in Continuous Time, *Math. Meth. Stat.* **8**, no. 1, 22–45.

[4] S.A. van de Geer (1999), *Applications of Empirical Process Theory*, to appear.

[5] I.I. Gihman and A.V. Skorohod (1972), *Stochastic Differential Equations*, Springer.

[6] I. Karatzas and S.E. Shreve (1991), *Brownian Motion and Stochastic Calculus*, 2nd edition, Springer.

[7] Yu.A. Kutoyants (1997a), Some problems of nonparametric estimation by observations of ergodic diffusion process, *Stat. Prob. Lett.* **32**, 311–320.

[8] Yu.A. Kutoyants (1997b), On unbiased density estimation for ergodic diffusion, *Stat. Prob. Lett.* **34**, 133–140.

[9] Yu.A. Kutoyants (1997c), Efficiency of the empirical distribution for ergodic diffusion, *Bernoulli* **3**(4), 445–456.

[10] H.T. Nguyen (1979), Density Estimation in a Continuous-Time Stationary Markov Process, *Ann. Stat.* **7**, 341–348.

[11] A.W. van der Vaart and J.A. Wellner (1996), *Weak Convergence and Empirical Processes, With Applications to Statistics*, Springer.

[12] J.H. van Zanten (1998), *A multivariate central limit theorem for continuous local martingales*, CWI-Report PNA-R9814.