

ROMAN RÓŻAŃSKI and ADAM ZAGDAŃSKI (Wrocław)

**ESTIMATION OF THE DRIFT FUNCTION FOR ITO
PROCESSES AND A CLASS OF SEMIMARTINGALES
VIA HISTOGRAM SIEVE**

Abstract. A histogram sieve estimator of the drift function in Ito processes and some semimartingales is constructed. It is proved that the estimator is pointwise and L^1 consistent and its finite-dimensional distributions are asymptotically normal. Our approach extends the results of Leśkow and Róžański (1989a).

1. Introduction. Since Grenander (1981) the method of sieves has turned out to be a very useful approach in nonparametric estimation. Many authors have applied different sieves for estimation of time dependent functions which are functional parameters of stochastic processes such as point processes belonging to a multiplicative intensity model (for results on sieve estimation of the intensity function in the multiplicative intensity model see Karr (1987), Leśkow and Róžański (1989b), Róžański and Zagdański (2001)), diffusion processes, Ito processes and more generally, semimartingale regression models; all these models are known to be widely used for describing the behaviour of dynamical systems.

Geman and Hwang (1982), using a sieve based on an orthonormal system, obtained a consistent estimator of an unknown functional parameter α in L^2 norm in the model

$$dX(t) = \alpha(t)dt + dW(t), \quad t \in [-1/2, 1/2],$$

where $W(t)$ is a Wiener process. This result was generalized by Nguyen and Pham (1982) who considered the model

$$dX(t) = \alpha(t)X(t)dt + dW(t), \quad X(0) = x_0.$$

2000 *Mathematics Subject Classification*: Primary 62M09; Secondary 62G05.

Key words and phrases: sieve method of estimation, Ito processes, semimartingales, consistency, asymptotical normality, nonparametric estimation.

Using a sieve based on Fourier expansion the authors proved L^2 consistency for an estimator of the unknown function α . Further, Leśkow and Róžański (1989a) using a histogram sieve constructed a consistent and asymptotically normal estimator of the unknown function α in the more general diffusion model

$$dX(t) = \alpha(t)a(t, X(t))dt + dW(t).$$

It should also be noted that the method of Fourier expansion in L^2 was used by McKeague (1986) in a more general semimartingale regression model. Further, Stone and Huang (2003) have applied another sieve, polynomial splines, to nonparametric estimation of the drift coefficient $\eta(\cdot, \cdot)$ in a diffusion type process $Y(t)$, where

$$dY(t) = \eta(t, X(t))dt + \sigma(t)dW(t)$$

with known random diffusion coefficient $\sigma(t)$ and observable stochastic covariate process $X(t)$. They have obtained the rates of convergence for spline estimates. All the estimators constructed in the above mentioned papers are based on n independent identically distributed (or with some conditions of mixing) copies of a stochastic process (model), which can be rather restrictive when one needs to take into consideration some dependence structure between observations.

In the present paper, we consider the problem of histogram sieve estimation of the drift function α for Ito processes and more generally for a class of semimartingale models. We prove that the histogram sieve estimator of the function α is consistent, L^1 consistent and asymptotically normal. In contrast to the above mentioned papers, the estimator we construct is based on a sequence of n processes (Ito processes or semimartingales) satisfying Conditions (A1), (A2), (A3), (B1), (A4) defined in Sections 3 and 4. The interpretation and meaning of these conditions, especially of Condition (B1), depends on the model describing a dynamical system. One can easily interpret the conditions in models described by the Langevin equation. It is also worth noting that if Conditions I–III from Leśkow and Róžański (1989a) are satisfied then after some transformation of the processes observed we get a model for which the conditions from our paper hold. The histogram sieve estimator which we construct retains its properties in the semimartingale regression model considered by McKeague (1986), satisfying Condition (A1').

The paper is organized as follows: Section 2 contains a short description of the method of sieves and the histogram sieve, Section 3 presents the general results for Ito processes and their proofs, and Section 4 is devoted to analogous results for some semimartingale models. In Section 5 we present some examples and simulation results.

2. Sieve maximum likelihood estimation. Let us start from a short description of Grenander's (1981) method of sieves.

A family $S(n)$ of subsets of a space $A \subset L^1$ is called a *sieve* if $S(n)$ is increasing in n and $\bigcup_n S(n)$ is dense in A . We assume that the family $S(n)$ is a *histogram sieve*, that is,

$$(2.1) \quad S(n) = \left\{ \alpha \in A : \alpha(s) = \sum_{l=1}^{m(n)} x_l 1_{B_{l,m(n)}}(s) \text{ for all } s \in [0, 1] \right\},$$

where $1_{B_{l,m(n)}}$ denotes the indicator of

$$B_{l,m(n)} = \left(\frac{l-1}{m(n)}, \frac{l}{m(n)} \right] \quad \text{for } l = 2, \dots, m(n),$$

$$B_{1,m(n)} = \left[0, \frac{1}{m(n)} \right],$$

and $\sum_l x_l^2 > 0$. The sequence $\{m(n)\}$ expresses the speed of growth of the sieve $S(n)$.

The maximum likelihood estimator $\hat{\alpha}_n$ based on the sieve $S(n)$ is defined through the equation

$$(2.2) \quad L_n(\hat{\alpha}_n) = \max_{\alpha \in S(n)} L_n(\alpha).$$

For the histogram sieve we can easily derive an exact expression for $\hat{\alpha}_n$.

In Section 4, we construct a histogram sieve estimator in another way. Namely, instead of the maximum likelihood sieve estimator we derive an estimator which is a solution of an estimating equation in the set of the histogram sieve for a suitably chosen estimating function.

3. Histogram sieve estimator

3.1. Model formulation. In this section we present the general form of the model considered. Let (Ω, \mathcal{F}, P) denote a probability space with filtration $\{\mathcal{F}_{n,t}\}$. We consider an Ito process $X_n(t)$, $t \in [0, 1]$, which is a strong solution of the following stochastic differential equation:

$$(3.1) \quad dX_n(t) = \alpha(t)D_n(t)dt + V_n(t)dW_n(t), \quad X_n(0) = 0.$$

The following assumptions will be imposed:

- (A.1) $W_n(t)$ is the standard Wiener process, adapted to the filtration $\{\mathcal{F}_{n,t}\}$.
- (A.2) $D_n(\cdot)$ and $V_n(\cdot)$ are nonanticipating processes relative to $\{\mathcal{F}_{n,t}\}$.
- (A.3) $P\{\int_0^1 |\alpha(t)D_n(t)| dt < \infty\} = 1$ and $P\{\int_0^1 V_n^2(t) dt < \infty\} = 1$.
- (A.4) $|V_n(t)| > 0$ P -a.s.
- (A.5) $\alpha(t)$ is bounded on $[0, 1]$.

3.2. Histogram estimator for diffusion processes. Let us start by constructing a histogram sieve estimator for a diffusion process which is a special case of model (3.1). In this section we will assume that X_n is a strong solution of the equation

$$(3.2) \quad dX_n(t) = \alpha(t)D_n(t, X_n)dt + V_n(t, X_n)dW_n(t), \quad X_n(0) = 0.$$

Furthermore, we will assume that the following conditions hold:

$$(A.6) \quad (a) \quad |V_n(t, x) - V_n(t, y)|^2 \leq L_1 \int_0^t |x_s - y_s|^2 dK(s) + L_2 |x_t - y_t|^2,$$

$$(b) \quad V_n^2(t, x) \leq L_1 \int_0^t (1 + x_s^2) dK(s) + L_2(1 + x_t^2),$$

where L_1, L_2 are some positive constants, $K(\cdot)$ is nondecreasing and right-continuous, $0 \leq K(s) \leq 1$, and $x, y \in C([0, 1])$.

(A.7) For any $0 \leq t \leq 1$, the equation

$$V_n(t, X_n)C_t(\omega) = \alpha(t)D_n(t, X_n)$$

has a P -a.s. bounded solution (relative to $C_t(\omega)$), which may be written as

$$C_t(\omega) = V_n^+(t, X_n)\alpha(t)D_n(t, X_n),$$

where

$$V_n^+(t, X_n) = \begin{cases} 1/V_n(t, X_n) & \text{for } V_n(t, X_n) \neq 0, \\ 0 & \text{for } V_n(t, X_n) = 0. \end{cases}$$

$$(A.8) \quad P \left\{ \int_0^1 \left(\frac{\alpha(s)D_n(s, X_n)}{V_n(s, X_n)} \right)^2 ds < \infty \right\} = 1.$$

$$(A.9) \quad E \left(\exp \left(- \int_0^1 C_t(\omega) dW_t - \frac{1}{2} \int_0^1 C_t^2(\omega) dt \right) \right) = 1.$$

Condition (A.6) guarantees the existence and uniqueness of strong solution of the stochastic differential equation

$$dY_n(t) = V_n(t, Y_n)dW_n(t).$$

It is known (Liptser and Shiriyayev (1981), Theorem 7.18) that under assumptions (A.1)–(A.9) the measure μ_α^n generated by the process X_n is dominated by the measure μ_{Y_n} corresponding to the process Y_n defined by $dY_n(t) = V_n(t, Y_n)dW_n(t)$. The density of μ_α^n with respect to μ_{Y_n} is

$$(3.3) \quad \frac{d\mu_\alpha^n}{d\mu_{Y_n}} = \exp \left(\int_0^1 \alpha(t) \frac{D_n(t, X_n)}{V_n^2(t, X_n)} dX_n(t) - \frac{1}{2} \int_0^1 \alpha^2(t) \frac{D_n^2(t, X_n)}{V_n^2(t, X_n)} dt \right).$$

Denote the logarithm of the above density by $L_n(\alpha)$.

For simplicity of notation, we will write $D_n(t), V_n(t)$ instead of $D_n(t, X_n), V_n(t, X_n)$. Our aim is to construct an estimator for the unknown functional drift parameter $\alpha(t)$. This problem is addressed via Grenander's (1981) method of sieves applied for the histogram sieve as defined in Section 2. Standard computations for the likelihood equation (2.2) with the density of the form (3.3) yield the exact expression for $\hat{\alpha}_n$:

LEMMA 3.1. *The maximum likelihood estimator $\hat{\alpha}_n$ based on the histogram sieve $S(n)$ is*

$$(3.4) \quad \hat{\alpha}_n(s) = \sum_{l=1}^{m(n)} \frac{\int_{B_{l,m(n)}} \frac{D_n(t)}{V_n^2(t)} dX_n(t)}{\int_{B_{l,m(n)}} \frac{D_n^2(t)}{V_n^2(t)} dt} 1_{B_{l,m(n)}}(s) 1_{C_{l,m(n)}}(s),$$

where

$$C_{l,m(n)} = \left\{ \int_{B_{l,m(n)}} \frac{D_n^2(t)}{V_n^2(t)} dt > 0 \right\}, \quad l = 1, \dots, m(n).$$

Let $s \in [0, 1]$ be fixed and choose $l(n, s) \in \{1, \dots, m(n)\}$ such that $s \in B_{l(n,s), m(n)}$. If we put $B_{m(n)}(s) = B_{l(n,s), m(n)}$ then the estimator $\hat{\alpha}_n(s)$ may be rewritten in the following compact form:

$$(3.5) \quad \hat{\alpha}_n(s) = \frac{\int_{B_{m(n)}(s)} \frac{D_n(t)}{V_n^2(t)} dX_n(t)}{\int_{B_{m(n)}(s)} \frac{D_n^2(t)}{V_n^2(t)} dt} 1_{C_{m(n)}(s)}.$$

3.2.1. Consistency and asymptotic normality. Consider the following conditions:

(B.1) There exists a positive and continuous function $y : [0, 1] \rightarrow \mathbb{R}_+$ such that

$$\sup_{t \in [0, 1]} \left| \frac{D_n^2(t)}{V_n^2(t)} \frac{1}{n} - y(t) \right| \xrightarrow[n \rightarrow \infty]{P} 0.$$

($\xrightarrow[n \rightarrow \infty]{P}$ denotes as usual convergence in probability).

(B.2) The function $\alpha(s)$ is continuous on $[0, 1]$.

(B.3) The speed of growth of the sieve is $m(n) = n^{1/2}$.

(B.4) There exists $\beta > 1/2$ and a positive constant $C(\alpha)$ such that

$$\forall_{s, t \in [0, 1]} |\alpha(s) - \alpha(t)| \leq C(\alpha) |t - s|^\beta,$$

in other words the Hölder condition holds.

Under assumptions (B.1)–(B.4) we prove the following theorems on consistency and asymptotic distribution of the histogram sieve estimator $\hat{\alpha}_n(s)$.

THEOREM 3.1. *If conditions (B.1) to (B.3) hold then the maximum likelihood estimator $\hat{\alpha}_n(s)$ defined in (3.4) and (3.5) is pointwise consistent for each $s \in [0, 1]$, that is, $\hat{\alpha}_n(s)$ converges to $\alpha(s)$ in probability for any $s \in [0, 1]$ as $n \rightarrow \infty$.*

THEOREM 3.2. *Let $\{s_1, \dots, s_p\}$ be an arbitrary finite collection of points from the interval $[0, 1]$. If conditions (B.1) to (B.4) are fulfilled then the sequence of random vectors*

$$n^{1/4}(\widehat{\alpha}_n(s_1) - \alpha(s_1), \dots, \widehat{\alpha}_n(s_p) - \alpha(s_p))$$

converges in distribution, as $n \rightarrow \infty$, to the p -dimensional normal distribution with zero expectation and diagonal covariance matrix with diagonal entries $\sigma_i = 1/y(s_i)$, $i = 1, \dots, p$.

REMARK. It is worth pointing out that assumption (B.3) can be varied. Namely, we can replace the sequence $m(n) = \sqrt{n}$, which determines the speed of growth of the sieve, by any sequence $m(n)$ tending to infinity more slowly than \sqrt{n} . This assumption still ensures the consistency of the estimator $\widehat{\alpha}_n(t)$.

In order to prove results on asymptotic distribution we need to impose some additional assumption on the smoothness of the estimated function $\alpha(t)$ and use the normalizing sequence $(n/m(n))^{1/2}$. For instance, assuming that α is differentiable at s and $n^{1/2}/m(n) \rightarrow \infty$, $n/m(n)^3 \rightarrow 0$ we deduce that $\widehat{\alpha}_n(s)$ is asymptotically normal $\mathcal{N}(\alpha(s), m(n)^2/ny(s))$.

Moreover, if we impose the additional conditions (B.5)–(B.6) below, it is possible to obtain L^1 consistency of the derived estimator.

(B.5) The function $\left| \frac{D_n^2(t)}{V_n^2(t)} \frac{1}{n} \right|$ is bounded away from zero, i.e. there exists

$\sigma > 0$ such that $\inf_{t \in [0,1]} \left| \frac{D_n^2(t)}{V_n^2(t)} \frac{1}{n} \right| > \sigma$ P -almost everywhere.

(B.6) The function $y(t)$ defined in condition (B.1) satisfies

$$E \left| \frac{D_n^2(t)}{V_n^2(t)} \frac{1}{n} - y(t) \right| \xrightarrow{n \rightarrow \infty} 0 \quad \forall t \in [0,1],$$

and $E \frac{D_n^2(t)}{V_n^2(t)}$ is continuous in t for every n .

THEOREM 3.3. *Assume that conditions (B.1)–(B.3), (B.5) and (B.6) are satisfied. Then the histogram sieve estimator $\widehat{\alpha}_n(s)$ is L^1 consistent, i.e.*

$$\int_0^1 |\widehat{\alpha}_n(s) - \alpha(s)| ds \xrightarrow[n \rightarrow \infty]{P} 0.$$

3.2.2. Proofs

Proof of Theorem 3.1. To prove the consistency of $\widehat{\alpha}_n(s)$ write

$$(3.6) \quad \widehat{\alpha}_n(s) - \alpha(s) = \frac{\frac{1}{\sqrt{n}} \int_{B_{m(n)}(s)} \frac{D_n^2(t)}{V_n^2(t)} (\alpha(t) - \alpha(s)) dt + \frac{1}{\sqrt{n}} \int_{B_{m(n)}(s)} \frac{D_n(t)}{V_n(t)} dW_n(t)}{\frac{1}{\sqrt{n}} \int_{B_{m(n)}(s)} \frac{D_n^2(t)}{V_n^2(t)} dt}.$$

Consider the sequence of martingales

$$M_n(t) = \frac{1}{\sqrt{n}} \int_0^t \frac{D_n(u)}{V_n(u)} dW_n(u),$$

which is also a sequence of random elements of $D([0, 1])$ (the space of right-continuous functions having left sided limits with Skorokhod topology). By Rebolledo's theorem (Rebolledo (1980)) the sequence $M_n(\cdot)$ is convergent in distribution in $D([0, 1])$ to an element $\widetilde{M}(\cdot)$ which is a continuous Gaussian martingale with independent increments for which $E\widetilde{M}^2(t) = \int_0^t y(u) du$.

Further, using Theorem 5.5 from Billingsley (1968) on weak convergence of a sequence of continuous mappings of random elements we obtain

$$M_n(B_{m(n)}(s)) = M_n(t_{m(n)+1}(s)) - M_n(t_{m(n)}(s)) \xrightarrow[n \rightarrow \infty]{D} 0,$$

where

$$B_{m(n)}(s) = (t_{m(n)}(s), t_{m(n)+1}(s)], \quad t_{m(n)}(s) \nearrow s, \quad t_{m(n)+1}(s) \searrow s.$$

By conditions (B.1)–(B.3),

$$\frac{1}{\sqrt{n}} \int_{B_{m(n)}(s)} \frac{D_n^2(t)}{V_n^2(t)} dt \xrightarrow[n \rightarrow \infty]{P} y(s)$$

and

$$\frac{1}{\sqrt{n}} \int_{B_{m(n)}(s)} \frac{D_n^2(t)}{V_n^2(t)} (\alpha(t) - \alpha(s)) dt \xrightarrow[n \rightarrow \infty]{P} 0,$$

which implies that (3.6) converges to zero in probability.

Proof of Theorem 3.2. We will first show the convergence of one-dimensional distributions. One can write

$$(3.7) \quad n^{1/4}(\widehat{\alpha}_n(s) - \alpha(s)) \\ = \frac{n^{1/4} \left(\frac{1}{\sqrt{n}} \int_{B_{m(n)}(s)} \frac{D_n^2(t)}{V_n^2(t)} (\alpha(t) - \alpha(s)) dt \right) + n^{1/4} \left(\frac{1}{\sqrt{n}} \int_{B_{m(n)}(s)} \frac{D_n(t)}{V_n(t)} dW_n(t) \right)}{\frac{1}{\sqrt{n}} \int_{B_{m(n)}(s)} \frac{D_n^2(t)}{V_n^2(t)} dt}.$$

We have shown that the denominator of (3.7) converges in probability to $y(s)$. Note that by (B.4) the first term in the numerator of (3.7) converges to zero in probability.

By Rebolledo's theorem the sequence $M_n(\cdot)$ is convergent in distribution in $D([0, 1])$ to an element $\widetilde{M}(\cdot)$ which is a continuous Gaussian martingale with independent increments. Denote by Q_n, Q the measures generated by $M_n(\cdot)$ and $\widetilde{M}(\cdot)$ in $D([0, 1])$ respectively.

Let \mathcal{A} denote the following class of subsets of $D([0, 1])$:

$$\mathcal{A} = \left\{ A \subset D([0, 1]) : A = \bigcap_k \{x : x(t_k(s)) \leq x_k\} \cap \bigcap_l \{x : x(t_l(s)) \leq y_l\} \right\},$$

where $\{t_k(s)\}$ and $\{t_l(s)\}$ are finite or infinite subsets of the sets $\{t_{m(n)}(s)\}$ and $\{t_{m(n)+1}(s)\}$ respectively, with arbitrary real numbers $\{x_k\}$, $\{y_l\}$. From Theorem 3 in Topsøe (1967) it follows that \mathcal{A} is a Q -uniformity class. Thus

$$\sup_{A \in \mathcal{A}} |Q_n(A) - Q(A)| \xrightarrow{n \rightarrow \infty} 0.$$

It is easy to see that the σ -algebra $\sigma(\mathcal{A})$ generated by the class \mathcal{A} is also a Q -uniformity class.

Let F_n and G_n denote the distribution functions of $M_n(B_{m(n)}(s))$ and $\widetilde{M}(B_{m(n)}(s))$ respectively. Since $\sigma(M_n(t_{m(n)}(s)); M_n(t_{m(n)+1}(s)), n \geq 1)$ is contained in $\sigma(\mathcal{A})$, which is a Q -uniformity class, we obtain

$$(3.8) \quad \sup_x |F_n(x) - G_n(x)| \xrightarrow{n \rightarrow \infty} 0.$$

Obviously, $n^{1/4} \widetilde{M}(B_{m(n)}(s)) \xrightarrow{D} U$ as $n \rightarrow \infty$, where U is a random variable normally distributed with zero expectation and variance $y(s)$. Thus, from (3.8) it follows that also $n^{1/4} M_n(B_{m(n)}(s)) \xrightarrow{D} U$ as $n \rightarrow \infty$.

Now, we have proved that the numerator of (3.7) converges in distribution to the random variable U and the denominator of (3.7) converges in probability to $y(s)$.

This shows the asymptotical normality of (3.7) with zero expectation and variance $1/y(s)$.

Now we will show the convergence of the two-dimensional distributions. We can write

$$\begin{aligned} & [n^{1/4}(\widehat{\alpha}_n(s_1) - \alpha(s_1)), n^{1/4}(\widehat{\alpha}_n(s_2) - \alpha(s_2))] \\ &= \left[\frac{n^{1/4} \left(\frac{1}{\sqrt{n}} \int_{B_{m(n)}(s_1)} \frac{D_n^2(t)}{V_n^2(t)} (\alpha(t) - \alpha(s_1)) dt \right)}{\frac{1}{\sqrt{n}} \int_{B_{m(n)}(s_1)} \frac{D_n^2(t)}{V_n^2(t)} dt}, \right. \\ & \quad \left. \frac{n^{1/4} \left(\frac{1}{\sqrt{n}} \int_{B_{m(n)}(s_2)} \frac{D_n^2(t)}{V_n^2(t)} (\alpha(t) - \alpha(s_2)) dt \right)}{\frac{1}{\sqrt{n}} \int_{B_{m(n)}(s_2)} \frac{D_n^2(t)}{V_n^2(t)} dt} \right] \\ &+ \left[\frac{n^{1/4} \left(\frac{1}{\sqrt{n}} \int_{B_{m(n)}(s_1)} \frac{D_n(t)}{V_n(t)} dW_n(t) \right)}{\frac{1}{\sqrt{n}} \int_{B_{m(n)}(s_1)} \frac{D_n^2(t)}{V_n^2(t)} dt}, \frac{n^{1/4} \left(\frac{1}{\sqrt{n}} \int_{B_{m(n)}(s_2)} \frac{D_n(t)}{V_n(t)} dW_n(t) \right)}{\frac{1}{\sqrt{n}} \int_{B_{m(n)}(s_2)} \frac{D_n^2(t)}{V_n^2(t)} dt} \right]. \end{aligned}$$

By earlier considerations the first term of the above sum is convergent to 0 in probability.

In order to show the asymptotical normality of the second term we introduce the following notations:

$$Y_{n,i} = n^{1/4} \left(\frac{1}{\sqrt{n}} \int_{B_{m(n)}(s_i)} \frac{D_n(t)}{V_n(t)} dW_n(t) \right),$$

$$Z_{n,i} = \frac{1}{\sqrt{n}} \int_{B_{m(n)}(s_i)} \frac{D_n^2(t)}{V_n^2(t)} dt, \quad i = 1, 2.$$

Using the same arguments as for the one-dimensional case, including Rebolledo's theorem and P -uniformity one can show that for all $\underline{t} = (t_1, t_2)$,

$$\langle \underline{t}, \underline{Y}_n \rangle = t_1 Y_{n,1} + t_2 Y_{n,2} \xrightarrow[n \rightarrow \infty]{D} t_1 Y_{0,1} + t_2 Y_{0,2},$$

where $Y_{0,1}$ and $Y_{0,2}$ are independent random variables, normally distributed with zero expectation and variance $y(s_1)$ and $y(s_2)$ respectively. By the Cramer–Wold device this shows the asymptotical normality of the vector $[Y_{n,1}, Y_{n,2}]$.

Define $R_{n,i} = t_i/Z_{n,i}$, $i = 1, 2$. We may write

$$\left\langle \underline{t}, \left[\frac{Y_{n,1}}{Z_{n,1}}, \frac{Y_{n,2}}{Z_{n,2}} \right] \right\rangle = R_{n,1} Y_{n,1} + R_{n,2} Y_{n,2} = \langle \underline{R}_n, \underline{Y}_n \rangle.$$

Applying Slutsky's lemma we observe that

$$\langle \underline{R}_n, \underline{Y}_n \rangle \xrightarrow[n \rightarrow \infty]{D} \frac{t_1}{y(s_1)} Y_{0,1} + \frac{t_2}{y(s_2)} Y_{0,2}$$

and using again the Cramer–Wold device we get the convergence for the two-dimensional case.

The proof of the convergence for the p -dimensional case goes along the same lines.

Proof of Theorem 3.3. Assumption (B.5) allows us to write

$$E|\widehat{\alpha}_n(s) - \alpha(s)| \leq \sigma^{-1} n^{-1/2} E \left| \int_{B_{m(n)}(s)} \frac{D_n^2(t)}{V_n^2(t)} (\alpha(t) - \alpha(s)) dt + \int_{B_{m(n)}(s)} \frac{D_n(t)}{V_n(t)} dW_n(t) \right|.$$

From condition (B.6) we see that there exists a positive constant C_1 such that for any $t \in [0, 1]$ and $n \in \mathbb{N}$,

$$E \frac{D_n^2(t)}{V_n^2(t)} \leq C_1 n.$$

Applying now the Hölder condition (B.7) and Cauchy inequality we observe that $E|\widehat{\alpha}_n(t) - \alpha(t)| \leq \sigma^{-1} (C(\alpha) C_1 n^{-\beta/2} + \sqrt{C_1} n^{-1/4})$. Therefore, we get

$$\lim_{n \rightarrow \infty} E|\widehat{\alpha}_n(t) - \alpha(t)| = 0.$$

Using the dominated convergence theorem and Fubini's theorem we obtain

$$\lim_{n \rightarrow \infty} E \int_0^1 |\hat{\alpha}_n(s) - \alpha(s)| ds = 0,$$

which by the Chebyshev inequality concludes the proof of the theorem.

3.3. Histogram estimator for Ito processes. In this section we extend the results obtained in Section 3.2 to a more general class of Ito processes.

In order to take advantage of the results obtained for diffusion processes it is necessary to put some restriction on $D_n(t)$ and $\alpha(t)$ in addition to assumptions (A.1)–(A.5) given in Section 3.1.

Namely, we will assume that

$$(A.10) \quad \int_0^1 E|\alpha(t)D_n(t)| dt < \infty.$$

By assumption (A.10) we may represent the Ito process given by (3.1) in the following form (see Liptser and Shiriyayev (1981), Theorem 7.17):

$$(3.9) \quad dX_n(t) = \alpha(t)\tilde{D}_n(X_n, t)dt + V_n(X_n, t)d\tilde{W}_n(t), \quad X_n(0) = 0,$$

where

$$\tilde{D}_n(X_n, t) = E(D_n(t) | \mathcal{F}_t^{X_n})$$

and $\mathcal{F}_t^{X_n} = \sigma\{X_n(s) : s \leq t\}$. Moreover, by (A.4) the Wiener process $\tilde{W}_n(t)$ is adapted to the filtration $(\mathcal{F}_t^{X_n})$, $0 \leq t \leq 1$.

Replacing $D_n(t, X_n(t))$ by $\tilde{D}_n(t, X_n(t))$ and formulating the assumptions analogous to (A.6)–(A.9) one may obtain the following

LEMMA 3.2. *The maximum likelihood estimator $\hat{\alpha}_n$ based on the histogram sieve $S(n)$ is*

$$(3.10) \quad \hat{\alpha}_n(s) = \sum_{l=1}^{m(n)} \frac{\int_{B_{l,m(n)}} \frac{\tilde{D}_n(t)}{V_n^2(t)} dX_n(t)}{\int_{B_{l,m(n)}} \frac{\tilde{D}_n^2(t)}{V_n^2(t)} dt} 1_{B_{l,m(n)}}(s) 1_{C_{l,m(n)}}(s).$$

Along the same lines as in Section 3.2.1 it is possible to obtain results on consistency and asymptotic normality of the estimator (3.10).

4. General semimartingale model. The results of Section 3 for diffusion and Ito processes can be generalized to semimartingale models. In this case we assume that the process $X_n(t)$ admits the following representation:

$$(4.1) \quad X_n(t) = \int_0^t \alpha(s)D_n(s) ds + \int_0^t V_n(s) dM_n(s), \quad X_n(0) = 0.$$

Instead of (A.1) and (A.2) assume that

(A.1') M_n is a cadlag martingale for which the predictable variation process satisfies $d\langle M_n \rangle(t) = q(t)dt$, where q is a continuous function.

(A.2') $D_n(t)$ and $V_n(t)$ are adapted to the filtration $\{\mathcal{F}_{n,t}\}$ and $\frac{D_n^2(t)}{q(t)V_n^2(t)}$ is locally integrable with respect to t .

Writing $M'_n(t) = \frac{1}{\sqrt{n}} \int_0^t \frac{D_n(v)}{V_n(v)} dM_n(v)$ we assume that for the martingale $M'_n(t)$ the following condition holds:

(A.3') $\forall_{t \in [0,1], \varepsilon > 0} \langle M'_{\varepsilon,n} \rangle(t) \xrightarrow{P} 0$ as $n \rightarrow \infty$, where $M'_{\varepsilon,n}$ is a locally square integrable martingale containing all the jumps of the martingale M'_n larger than ε in absolute value.

Unfortunately, in such a general semimartingale model with discontinuous realizations it is difficult to derive the form of the likelihood function analogous to (3.3). Instead, we will follow the ideas of Hutton and Nelson (1986) and construct a histogram sieve estimator which is a solution of an estimating equation in the set of the histogram sieve for an appropriately chosen estimating function. Let $S(n)$ be a histogram sieve as in (2.1). For $\alpha(s) = \sum_{l=1}^{m(n)} x_l 1_{B_{l,m(n)}}(s)$, define the *estimating function* $\dot{Q}_n(\theta_{m(n)}) = (\dot{Q}_{n,1}(\theta_{m(n)}), \dots, \dot{Q}_{n,l}(\theta_{m(n)}), \dots, \dot{Q}_{n,m(n)}(\theta_{m(n)}))$ as

$$\begin{aligned}
 (4.2) \quad \dot{Q}_{n,l}(\theta_{m(n)}) &= \int_0^1 1_{B_{l,m(n)}}(t) \frac{D_n(t)}{q(t)V_n^2(t)} dX_n(t) \\
 &\quad - \int_0^1 1_{B_{l,m(n)}}(t) \left(\sum_{k=1}^{m(n)} x_k 1_{B_{k,m(n)}}(t) \right) \frac{D_n^2(t)}{q(t)V_n^2(t)} dt \\
 &= \int_{B_{l,m(n)}} \frac{D_n(t)}{q(t)V_n^2(t)} dX_n(t) - x_l \int_{B_{l,m(n)}} \frac{D_n^2(t)}{q(t)V_n^2(t)} dt,
 \end{aligned}$$

where $\theta_{m(n)} = (x_1, \dots, x_{m(n)})$ and $l = 1, \dots, m(n)$. It is easily seen that the solution $\hat{\theta}_{m(n)} = (\hat{x}_1, \dots, \hat{x}_{m(n)})$ of the *estimating equation*

$$\dot{Q}_n(\theta_{m(n)}) = 0$$

has the form

$$\hat{x}_{l,n} = \frac{\int_{B_{l,m(n)}} \frac{D_n(t)}{q(t)V_n^2(t)} dX_n(t)}{\int_{B_{l,m(n)}} \frac{D_n^2(t)}{q(t)V_n^2(t)} dt}.$$

Thus, the histogram sieve estimator derived through the estimating function (4.2) may be written in the form

$$\hat{\alpha}_n(t) = \sum_{l=1}^{m(n)} \hat{x}_{l,n} 1_{B_{l,m(n)}}(t) 1_{C_{l,m(n)}}(t),$$

where

$$C_{l,m(n)} = \left\{ \int_{B_{l,m(n)}} \frac{D_n^2(t)}{q(t)V_n^2(t)} dt > 0 \right\}, \quad l = 1, \dots, m(n).$$

For fixed $s \in [0, 1]$ we obtain

$$(4.3) \quad \widehat{\alpha}_n(s) = \frac{\int_{B_{m(n)}(s)} \frac{D_n(t)}{q(t)V_n^2(t)} dX_n(t)}{\int_{B_{m(n)}(s)} \frac{D_n^2(t)}{q(t)V_n^2(t)} dt} 1_{C_{m(n)}(s)},$$

in the same way as in (3.5).

In what follows, we assume that all conditions (A.1'), (A.2'), (A.3') and (A.4) hold and that for $D_n(t)$ and $V_n(t)$ satisfying these conditions some of assumptions (B.1)–(B.6) are fulfilled.

Along the same lines as in the previous section we obtain the following theorems.

THEOREM 4.1. *If conditions (B.1)–(B.3) hold then the histogram sieve estimator $\widehat{\alpha}_n(s)$ defined in (4.3) is pointwise consistent for each $s \in [0, 1]$.*

THEOREM 4.2. *Let $\{s_1, \dots, s_p\}$ be an arbitrary finite collection of points from the interval $[0, 1]$. If conditions (B.1)–(B.4) are fulfilled then the sequence of random vectors*

$$n^{1/4}(\widehat{\alpha}_n(s_1) - \alpha(s_1), \dots, \widehat{\alpha}_n(s_p) - \alpha(s_p))$$

converges in distribution, as $n \rightarrow \infty$, to the p -dimensional normal distribution with zero expectation and diagonal covariance matrix with diagonal entries $\sigma_i = q(s_i)/y(s_i)$, $i = 1, \dots, p$.

THEOREM 4.3. *Assume that conditions (B.1)–(B.3), (B.5) and (B.6) are satisfied. Then the histogram sieve estimator $\widehat{\alpha}_n(s)$ is L^1 consistent, i.e.*

$$\int_0^1 |\widehat{\alpha}_n(s) - \alpha(s)| ds \xrightarrow[n \rightarrow \infty]{P} 0.$$

5. Examples and simulation results. (i) Let $Z_n(t)$ denote a sequence of diffusion processes satisfying the following stochastic differential equation:

$$dZ_n(t) = \alpha(t) dt + \frac{\sigma}{\sqrt{n}} dW_n(t), \quad Z_n(0) = 0, \quad t \in [0, 1].$$

In this case, estimation of the unknown drift function $\alpha(t)$ is asymptotically equivalent to estimation of the regression function in the model

$$Y_i = \alpha(t_i) + \sigma \varepsilon_i, \quad i = 1, \dots, n,$$

where $t_i = i/n$ and ε_i is an i.i.d. sequence with $E\varepsilon_i = 0$ and $\text{Var } \varepsilon_i = 1$.

(ii) Let $Y_k(t)$, $k = 1, \dots, n$, be a sequence of independent copies of a diffusion process $Y(t)$ for which

$$dY(t) = \alpha(t)a(t, Y(t))dt + dW(t),$$

where $Y(0) = \tilde{Y}$ is a random variable and $W(t)$ is a Wiener process. Assume that the process $Y(t)$ satisfies Conditions I–III from Leśkow and Róžański (1989a). Define a sequence $X_n(t)$ of semimartingales by

$$\begin{aligned} dX_n(t) &= \sum_{k=1}^n a(t, Y_k(t))dY_k(t) \\ &= \alpha(t) \left(\sum_{k=1}^n a^2(t, Y_k(t)) \right) dt + \sum_{k=1}^n a(t, Y_k(t))dW_k(t). \end{aligned}$$

Then the process $X_n(t)$ can be written in the form of a semimartingale model considered in Section 4. Moreover, if Conditions I–III from Leśkow and Róžański (1989a) hold then Conditions (A1'), (A2'), (A3'), (A4), (B1) are satisfied.

Simulations. In order to gain insight into the behaviour of the histogram sieve estimator we carried out some computer experiments. In our simulation study we considered Examples (i) and (ii) described above. The data were generated according to the following models:

$$(M1) \quad dZ_n(t) = \alpha(t)dt + \frac{\sigma}{\sqrt{n}} dW_n(t), \quad Z_n(0) = 0, \quad \sigma = 1,$$

$$(M2) \quad dX_n(t) = \sum_{k=1}^n Y_k(t)dY_k(t),$$

where Y_k are independent copies of a diffusion process $Y(t)$ for which

$$dY_k(t) = \alpha(t)Y_k(t)dt + dW_k(t), \quad Y_k(0) \sim N(0, 1).$$

For both models two different drift functions were used: $\alpha(t) = \sin(4\pi t)$ and $\alpha(t) = t^2$.

Figures 1–4 show the histogram sieve estimator of the drift function (dotted line) for models M1 and M2, obtained for $n = 500$ and $m(n) = \sqrt{n}$.

We have also drawn 95% pointwise confidence intervals (dashed lines) based on asymptotic normality and variance estimator. Performance of the constructed intervals was investigated in terms of coverage probabilities based on 1000 Monte Carlo trials. Simulations were carried out for three different choices of the sequence $m(n)$, namely $m(n) = \sqrt{n}$, $m(n) = n^{4/9}$ and $m(n) = n^{2/5}$. Note that the latter two satisfy the conditions given in Remark in Section 3.

Tables 1–6 contain coverage percentages with estimated standard errors in parentheses (in percentages). These results were obtained for selected central points of the subintervals defining the partition of the interval $[0,1]$.

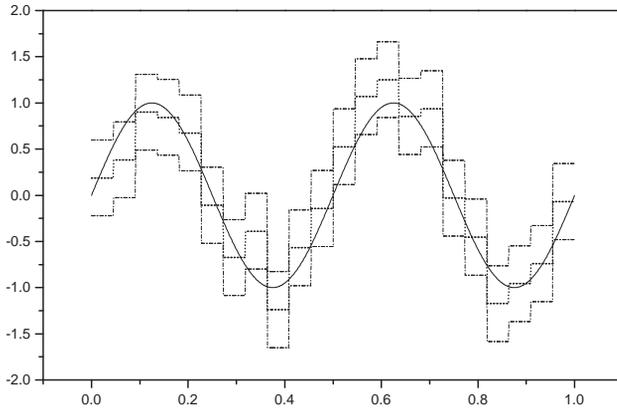


Fig. 1. Histogram sieve estimator for M1 and $\alpha(t) = \sin(4\pi t)$

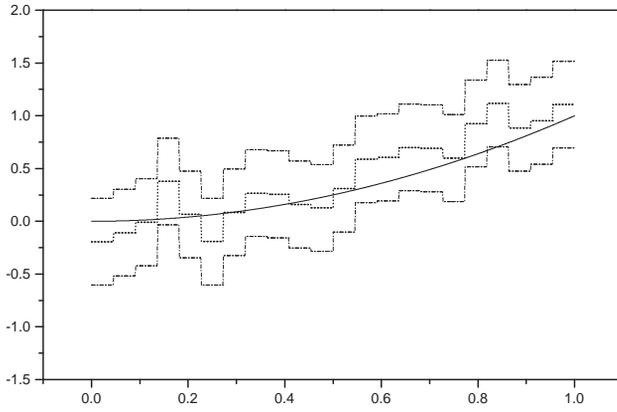


Fig. 2. Histogram sieve estimator for M1 and $\alpha(t) = t^2$

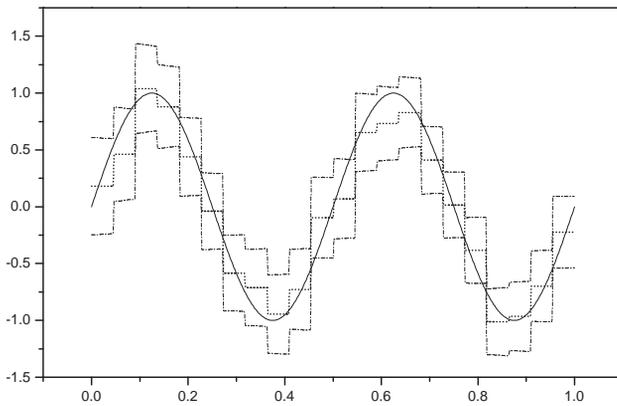


Fig. 3. Histogram sieve estimator for M2 and $\alpha(t) = \sin(4\pi t)$

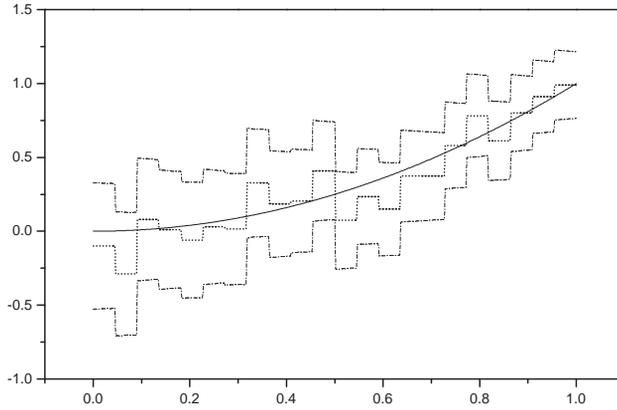


Fig. 4. Histogram sieve estimator for M2 and $\alpha(t) = t^2$

Table 1. Empirical coverage for model M1 and $m(n) = \sqrt{n}$

t	$\alpha(t) = \sin(4\pi t)$	$\alpha(t) = t^2$
0.0227	93.9%(0.757)	95.8%(0.634)
0.0680	95.7%(0.641)	93.3%(0.791)
0.1133	94.8%(0.702)	95.9%(0.627)
0.1593	95.3%(0.669)	94.8%(0.702)
0.2047	94.1%(0.745)	94.1%(0.745)
0.2500	95.8%(0.634)	96.1%(0.612)
0.2953	94.6%(0.715)	95.3%(0.669)
0.3407	94.9%(0.696)	94.3%(0.733)
0.3867	94.4%(0.727)	95.7%(0.641)
0.4320	94.4%(0.727)	95.3%(0.669)
0.4773	94.8%(0.702)	94.8%(0.702)
0.5227	95.2%(0.676)	94.4%(0.727)
0.5680	94.0%(0.751)	94.7%(0.708)
0.6133	94.9%(0.696)	94.7%(0.708)
0.6593	95.4%(0.662)	96.0%(0.620)
0.7047	95.1%(0.683)	95.1%(0.683)
0.7500	94.2%(0.739)	94.4%(0.727)
0.7953	95.9%(0.627)	94.1%(0.745)
0.8407	94.4%(0.727)	95.4%(0.662)
0.8867	95.5%(0.656)	95.2%(0.676)
0.9320	95.0%(0.689)	95.9%(0.627)
0.9773	95.1%(0.683)	94.7%(0.708)

Table 2. Empirical coverage for model M1 and $m(n) = n^{4/9}$

t	$\alpha(t) = \sin(4\pi t)$	$\alpha(t) = t^2$
0.0313	94.5%(0.721)	95.6%(0.649)
0.0940	94.1%(0.745)	95.7%(0.641)
0.1560	96.2%(0.605)	95.3%(0.669)
0.2187	95.0%(0.689)	94.8%(0.702)
0.2813	95.2%(0.676)	94.9%(0.696)
0.3440	95.1%(0.683)	95.0%(0.689)
0.4060	94.3%(0.733)	94.6%(0.715)
0.4687	95.1%(0.683)	94.4%(0.727)
0.5313	93.6%(0.774)	94.5%(0.721)
0.5940	95.7%(0.641)	94.3%(0.733)
0.6560	95.6%(0.649)	95.8%(0.634)
0.7187	95.9%(0.627)	95.2%(0.676)
0.7813	95.7%(0.641)	93.3%(0.791)
0.8440	95.0%(0.689)	95.7%(0.641)
0.9060	93.8%(0.763)	93.8%(0.763)
0.9687	95.2%(0.676)	94.2%(0.739)

Table 3. Empirical coverage for model M1 and $m(n) = n^{2/5}$

t	$\alpha(t) = \sin(4\pi t)$	$\alpha(t) = t^2$
0.0413	94.8%(0.702)	94.6%(0.715)
0.1253	94.1%(0.745)	95.5%(0.656)
0.2080	94.9%(0.696)	95.4%(0.662)
0.2920	94.8%(0.702)	94.6%(0.715)
0.3753	93.4%(0.785)	95.0%(0.689)
0.4580	95.1%(0.683)	94.7%(0.708)
0.5413	93.8%(0.763)	94.7%(0.708)
0.6247	93.7%(0.768)	95.4%(0.662)
0.7080	95.3%(0.669)	94.2%(0.739)
0.7913	95.4%(0.662)	93.9%(0.757)
0.8747	92.6%(0.828)	94.4%(0.727)
0.9580	93.7%(0.768)	95.4%(0.662)

We observe that the performance of the confidence intervals for models M1 and M2 with sinusoidal drift function $\alpha(t) = \sin(4\pi t)$ is not satisfactory when we choose $m(n) = n^{2/5}$ tending to infinity too slowly. Other choices of the sequence $m(n)$ yield rather similar coverages close to the nominal.

Table 4. Empirical coverage for model M2 and $m(n) = \sqrt{n}$

t	$\alpha(t) = \sin(4\pi t)$	$\alpha(t) = t^2$
0.0233	94.0%(0.751)	94.4%(0.727)
0.0683	95.4%(0.662)	95.5%(0.656)
0.1133	93.4%(0.785)	94.0%(0.751)
0.1583	95.4%(0.662)	95.6%(0.649)
0.2050	94.9%(0.696)	95.1%(0.683)
0.2500	94.7%(0.708)	94.6%(0.715)
0.2950	93.5%(0.780)	93.2%(0.796)
0.3417	94.8%(0.702)	94.9%(0.696)
0.3867	95.8%(0.634)	95.5%(0.656)
0.4317	95.2%(0.676)	95.2%(0.676)
0.4767	95.5%(0.656)	95.4%(0.662)
0.5233	95.8%(0.634)	96.0%(0.620)
0.5683	93.9%(0.757)	94.4%(0.727)
0.6133	95.2%(0.676)	95.6%(0.649)
0.6583	95.0%(0.689)	95.2%(0.676)
0.7050	95.8%(0.634)	95.8%(0.634)
0.7500	96.9%(0.548)	96.5%(0.581)
0.7950	94.6%(0.715)	94.7%(0.708)
0.8417	94.8%(0.702)	94.4%(0.727)
0.8867	94.7%(0.708)	95.0%(0.689)
0.9317	94.3%(0.733)	94.3%(0.733)
0.9767	94.6%(0.715)	95.0%(0.689)

For the remaining models (with parabolic drift function) the results of empirical coverage obtained for all sequences $m(n)$ considered are quite similar and choosing $m(n)$ different from \sqrt{n} generally does not improve coverage significantly. Nevertheless, the coverage probabilities obtained for $m(n) = \sqrt{n}$ seem to exhibit more variability across different points of the interval $[0, 1]$ than for other choices of $m(n)$.

Additionally, it is worth noting that for all cases considered, smaller $m(n)$ yields narrower confidence intervals.

One can also analyse the accuracy of the histogram sieve estimator with respect to some measure of goodness of fit such as mean integrated absolute error $MIAE = E \int_0^1 |\hat{\alpha}(u) - \alpha(u)| du$, or mean integrated squared error $MISE = E \int_0^1 (\hat{\alpha}(u) - \alpha(u))^2 du$. We recall that L^1 consistency was proved in Section 3 (Theorem 3.3).

Table 5. Empirical coverage for model M2 and $m(n) = n^{4/9}$

t	$\alpha(t) = \sin(4\pi t)$	$\alpha(t) = t^2$
0.0317	95.0%(0.689)	95.1%(0.683)
0.0933	95.7%(0.641)	96.0%(0.620)
0.1567	95.0%(0.689)	94.5%(0.721)
0.2183	94.0%(0.751)	94.2%(0.739)
0.2817	94.5%(0.721)	94.4%(0.727)
0.3433	95.0%(0.689)	95.6%(0.649)
0.4067	95.0%(0.689)	95.0%(0.689)
0.4683	95.8%(0.634)	96.1%(0.612)
0.5317	95.4%(0.662)	95.5%(0.656)
0.5933	94.1%(0.745)	95.3%(0.669)
0.6567	95.4%(0.662)	95.2%(0.676)
0.7183	95.5%(0.656)	95.6%(0.649)
0.7817	94.9%(0.696)	95.3%(0.669)
0.8433	94.9%(0.696)	95.3%(0.669)
0.9067	94.8%(0.702)	95.3%(0.669)
0.9683	94.3%(0.733)	94.6%(0.715)

Table 6. Empirical coverage for model M2 and $m(n) = n^{2/5}$

t	$\alpha(t) = \sin(4\pi t)$	$\alpha(t) = t^2$
0.0417	95.3%(0.669)	95.8%(0.634)
0.1250	92.8%(0.817)	94.0%(0.751)
0.2083	94.9%(0.696)	95.0%(0.689)
0.2917	93.8%(0.763)	94.6%(0.715)
0.3750	93.1%(0.801)	94.2%(0.739)
0.4583	94.6%(0.715)	95.3%(0.669)
0.5417	94.2%(0.739)	94.8%(0.702)
0.6250	93.2%(0.796)	94.4%(0.727)
0.7083	95.1%(0.683)	95.0%(0.689)
0.7917	94.8%(0.702)	94.5%(0.721)
0.8750	93.0%(0.807)	95.5%(0.656)
0.9583	94.3%(0.733)	95.1%(0.683)

A closely related issue is choosing an optimal sequence $m(n)$ (determining the speed of growth of the sieve) which minimizes a given criterion. A detailed discussion of these problems is beyond the scope of this paper and will be given elsewhere. We restrict ourselves to presenting some numerical results.

Table 7. Empirical MIAE for model M1

$\alpha(t)$	n	$m(n)$		
		\sqrt{n}	$n^{4/9}$	$n^{2/5}$
$\sin(4\pi t)$	500	0.192	0.191	0.208
	1000	0.156	0.151	0.163
t^2	500	0.167	0.144	0.125
	1000	0.142	0.118	0.102

Table 8. Empirical MIAE for model M2

$\alpha(t)$	n	$m(n)$		
		\sqrt{n}	$n^{4/9}$	$n^{2/5}$
$\sin(4\pi t)$	500	0.161	0.169	0.193
	1000	0.129	0.131	0.150
t^2	500	0.131	0.112	0.099
	1000	0.111	0.093	0.081

Table 9. Empirical MISE for model M1

$\alpha(t)$	n	$m(n)$		
		\sqrt{n}	$n^{4/9}$	$n^{2/5}$
$\sin(4\pi t)$	500	0.058	0.057	0.068
	1000	0.038	0.036	0.041
t^2	500	0.044	0.033	0.025
	1000	0.032	0.022	0.016

Table 10. Empirical MISE for model M2

$\alpha(t)$	n	$m(n)$		
		\sqrt{n}	$n^{4/9}$	$n^{2/5}$
$\sin(4\pi t)$	500	0.041	0.045	0.059
	1000	0.026	0.027	0.035
t^2	500	0.028	0.020	0.016
	1000	0.020	0.014	0.010

Tables 7–10 contain results of empirical MIAE and MISE based on 1000 Monte Carlo realizations obtained for models M1 and M2 and for three different choices of the sequence $m(n)$. Roughly speaking, both criteria behave

in a similar manner with respect to the choice of $m(n)$. Namely, for both models and drift function $\alpha(t) = t^2$ we observe that MIAE as well as MISE is decreasing when $m(n)$ decreases. On the other hand, for the sinusoidal drift function the behaviour is not so regular. Namely, for M1 and $\alpha(t) = \sin(4\pi t)$ the best result is attained for the sequence $m(n) = n^{4/9}$ whereas for M2 the values of both criteria are minimal when $m(n) = \sqrt{n}$.

References

- P. Billingsley (1968), *Convergence of Probability Measures*, Wiley.
- S. Geman and C. R. Hwang (1982), *Nonparametric maximum likelihood estimation by the method of sieves*, Ann. Statist. 10, 401–414.
- U. Grenander (1981), *Abstract Inference*, Wiley, New York.
- J. E. Hutton and P. L. Nelson (1986), *Quasi-likelihood estimation for semimartingales*, Stochastic Process. Appl., 22, 245–257.
- A. F. Karr (1987), *Maximum likelihood estimation in the multiplicative intensity model via sieves*, Ann. Statist. 15, 473–490.
- J. Leśkow and R. Róžański (1989a), *Maximum likelihood estimator of the drift function for a diffusion process*, Statist. Decisions 7, 243–262.
- J. Leśkow and R. Róžański (1989b), *Histogram maximum likelihood estimator in the multiplicative intensity model*, Stochastic Process. Appl. 31, 151–159.
- R. Liptser and A. Shiriyayev (1981), *Statistics of Random Processes*, PWN, Warszawa (Polish translation).
- I. McKeague (1986), *Estimation for a semimartingale regression model using the method of sieves*, Ann. Statist. 14, 579–589.
- H. T. Nguyen and T. D. Pham (1982), *Identification of nonstationary diffusion model by the method of sieves*, SIAM J. Control Optim. 20, 603–611.
- B. L. S. Prakasa Rao (1999), *Semimartingales and Their Statistical Inference*, Chapman.
- R. Rebolledo (1980), *Central limit theorems for local martingales*, Z. Wahrsch. Verw. Gebiete 51, 269–286.
- R. Róžański and A. Zagdański (2001), *On estimation in the multiplicative intensity model via histogram sieve*, Probab. Math. Statist. 21, 89–99.
- H. J. Stone and J. Z. Huang (2003), *Statistical modeling of diffusion processes with free knot splines*, J. Statist. Plann. Inference 116, 451–474.
- F. Topsøe (1967), *On the connection between P -continuity and P -uniformity in weak convergence*, Teor. Veroyatnost. i Primenen. 12, 279–288 (in Russian).

Institute of Mathematics and Computer Science
 Wrocław University of Technology
 Wybrzeże Wyspiańskiego 27
 50-370 Wrocław, Poland
 E-mail: roman.rozanski@pwr.wroc.pl

Received on 8.10.2002;
 revised version on 7.12.2005

(1657)