UNIVERSITÀ DEGLI STUDI DI BERGAMO
DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE
E METODI MATEMATICI°

QUADERNI DEL DIPARTIMENTO

Department of Information Technology and Mathematical Methods

Working Paper

Series "*Mathematics and Statistics*"

n. 1/MS – 2011

Asymptotically distribution free test for parameter change in a diffusion process model

by

Ilia Negri, Yoichi Nishiyama

° Viale Marconi. 5, I – 24044 Dalmine (BG), ITALY, Tel. +39-035-2052339; Fax. +39-035-562779

Asymptotically distribution free test for parameter change in a diffusion process model *

Ilia Negri[†] Yoichi Nishiyama[‡]

January 7, 2011

Abstract

A test procedure to detect a change in the value of the parameter in the drift of a diffusion process is proposed. The test statistic is asymptotically distribution free under the null hypothesis that the true parameter does not change. Also, the test is shown to be consistent under the alternative that there exists a change point.

Key words: consistent test, empirical process, asymptotically distribution free tests.
2010 MSC: 62M02; 60G10; 60G15.

*This work was supported by Grant-in-Aid for Scientific Research (C), 21540157, from Japan Society for the Promotion of Science and by the Department of Information Technology and Mathematical Methods, Grant 2010.

[†]Corresponding author: Department of Information Technology and Mathematical Methods, University of Bergamo, Viale Marconi, 5, 24044, Dalmine (BG), Italy. ilia.negri@unibg.it.

[‡]The Institute of Statistical Mathematics, 10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan. nisiyama@ism.ac.jp

1 Introduction

Testing on structural change problems has been an important issue in statistics. It originally starts in quality control context, where one is concerned about the output of a production line and wants to find any departure from an acceptable standard of the products. Rapidly the problem of abrupt changes moved to various fields such as economics, finance, biology and environmental sciences. From the statistical point of view, the problem consists in testing whether there is a statistically significant change point in a sequence of chronologically ordered data. The problem for an i.i.d sample was first considered in the paper of Page [18], see also Hinkley [10], and, for a general survey of the change point detection and estimation, see Chen and Gupta [4]. The parameter chance point problem became very popular in regression and time series models. This is because these models can be used to describe structural changes that often occur in financial and economic phenomena (due for example to a change of political situation or to a change of economic policy) or in environmental phenomena (due to sudden changes in weather situation or the happen of catastrophic natural events). In such kind of phenomena the first problem one has deal with is to test if a change of parameter is occurred in the factor of interest. For regression models see for example Hinkley [9], Quandt [20], Brown, Durbin and Evans [1], Chen [5]. For time series models, Picard [19] considers the problem of detecting a change-point occurring in the mean or in the covariance of an autoregressive process. Ling [14] deals with detecting structural changes in a general time series framework that includes ARMA and GARCH models between others. For a general review we refer also to Csörgő and Horváth [6] and to Chen and Gupta [3] for parametric methods and analysis. Diffusion process can be considered as the most popular continuos time stochastic process and it has been playing a central role in modeling phenomena in many fields, not only in finance and more generally in economics science, but also in other fields such as biology, medicine, physics and engineering. Despite the fact of their importance in applications, few works are devoted to testing change point in parameter for diffusions. For a complete reference on statistical problems for ergodic diffusions based on continuous time observations see Kutoyants [12]. In Lee, Nishiyama and Yoshida [13], the cusum test based on one-step estimator is considered and up to our knowledge there are no existing literature on this subject and on this framework.

In this work a test for detecting if a change of the parameter in the drift of a diffusion process takes place is proposed. The test is based on the continuous observation of the process up to time T. The interest for this test is that it may be used for the most common family of diffusion process because the conditions on the coefficients of the diffusion process are very general. Moreover the asymptotic distribution of the test statistics does not depend on the unknown parameter, so the test is asymptotically distribution free. It is also proved that the test is consistent against any alternative where the alternative means that at a certain instant the parameter specifying the drift coefficient change.

The rest of the paper is organized as follow. In Section 2 the model, the conditions and some preliminary result needed later are presented. The main result, consisting in the asymptotic distribution of the test statistic and the construction of an asymptotically distribution free and consistent test is given in Section 3. Finally the necessary lemmas needed to prove the main theorems are proved in the Section 4.

2 Preliminaries

Let $(\Omega, \mathcal{A}, \mathbf{P})$ be a probability space and $\{\mathcal{A}_t\}_{t\geq 0}$ a filtration of $\mathcal{A}$. Let $\{S(\cdot,\theta) : \theta \in \Theta\}$ be a family of $\mathbb{R}$-valued measurable functions on $\mathbb{R}$ indexed by a subset Θ of $\mathbb{R}^k$. With $||\cdot||$ we denote the usual norm in $\mathbb{R}^k$. Let $\sigma : \mathbb{R} \to (0,\infty)$ be a measurable function which is known to statisticians. Suppose that the functions $S(\cdot,\theta)$ and $\sigma(\cdot)$ are such that there exists a solution X^θ to the stochastic differential equation (SDE)

$$X_t = X_0 + \int_0^t S(X_s,\theta)ds + \int_0^t \sigma(X_s)dW_s, \quad t \geq 0, \tag{1}$$

where $W = \{W_s : \ s \geq 0\}$ is a standard Wiener process and the initial value X_0 is independent of W_t, $t \geq 0$. Denote by $\mathbf{E}$ the expectation with respect to $\mathbf{P}$. The *scale function* of the diffusion (1) is defined by $p_\theta(x) = \int_0^x \exp\left\{-2\int_0^y \frac{S(v,\theta)}{\sigma^2(v)}\,\mathrm{d}v\right\}\mathrm{d}y$. The *speed measure* of the diffusion (1) is defined by

$$m_\theta(A) = \int_A \frac{1}{\sigma(x)^2}\exp\left(2\int_0^x \frac{S(y,\theta)}{\sigma(y)^2}dy\right)dx, \quad A \in \mathcal{B}(\mathbb{R}),$$

where $\mathcal{B}(\mathbb{R})$ is the Borel sigma algebra on $\mathbb{R}$. Furthermore let us suppose that X^θ is regular, $\lim_{x\to\pm\infty} p_\theta(x) = \pm\infty$ and the speed measure is finite, that is, $m_\theta(\mathbb{R}) < \infty$. Then, the solution process X^θ is ergodic with the invariant distribution function $F_\theta(\cdot)$ given by $F_\theta(x) = m_\theta((-\infty,x])/m_\theta(\mathbb{R})$, that is, it holds for any $dF_\theta(x)$-integrable function g that with probability one,

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T g(X_t)dt = \int_{\mathbb{R}} g(z)dF_\theta(z).$$

See for example Gikhman and Skorohod [8]. For the definition of regular process see for example Rogers and Williams [21] or Durrett [7].

Now for a given function $\theta_T : [0, T] \to \Theta$, let us consider the SDE

$$X_t = X_0 + \int_0^t S(X_s, \theta_T(s))ds + \int_0^t \sigma(X_s)dW_s, \quad t \in [0, T]. \tag{2}$$

Although the solution X to this SDE may not exist in general, at least it does exist in the following hypotheses which we wish to test:

H_0: there exists a certain $\theta_0 \in \Theta$ such that $\theta_T(s) = \theta_0$ for all $s \in [0, T]$;

H_1: there exist two different values $\theta_0 \neq \theta_1$ both belonging to Θ, and a certain $u_* \in (0, 1)$, such that $\theta_T(s) = \theta_0$ for $s \in [0, Tu_*]$ and $\theta_T(s) = \theta_1$ for $s \in (Tu_*, T]$.

Let us introduce the following regularity conditions on the functions S and σ. We denote with a dot the derivative with respect to θ and with a double dot the second derivative with respect to θ. We suppose that S is two times differentiable with respect to θ and the derivatives satisfy the following conditions:

$$\int_{\mathbb{R}} \frac{||\dot{S}(z, \theta_0)||}{\sigma(z)} dF_{\theta_0}(z) < \infty, \quad \forall \theta_0 \in \Theta, \tag{3}$$

and

$$\int_{\mathbb{R}} \frac{\sup_\theta ||\ddot{S}(z, \theta)||}{\sigma(z)} dF_{\theta_0}(z) < \infty, \quad \forall \theta_0 \in \Theta. \tag{4}$$

In this paper, we present a very simple method to construct a test statistic which is asymptotically distribution free under H_0, and consistent under H_1.

The log-likelihood function of the process (1) observed up to time T, is given by

$$L_T(\theta) = \int_0^T \frac{S(X_t, \theta)}{\sigma^2(X_t)} dX_t - \frac{1}{2} \int_0^T \frac{S^2(X_t, \theta)}{\sigma^2(X_t)} dt. \tag{5}$$

We define $\hat{\theta}_T$ the maximizer of (5) over Θ. We suppose that for every

$\theta_i \in \Theta$, θ_i is the unique local and global minimizer of

$$\theta \mapsto g(\theta, \theta_i) = \int_{\mathbb{R}} \frac{(S(x,\theta) - S(x,\theta_i))^2}{\sigma^2(x)} dF_{\theta_i}(x), \tag{6}$$

over Θ. Actually, we suppose that the function

$$\theta \mapsto \frac{\partial}{\partial \theta} g(\theta, \theta_i) = \int_{\mathbb{R}} \frac{(S(x,\theta) - S(x,\theta_i))\dot{S}(x,\theta)}{\sigma^2(x)} dF_{\theta_i}(x), \tag{7}$$

is zero if and only if $\theta = \theta_i$. Hereafter, we suppose that the order of integration and differentiation is exchangeable. Let θ^* be the minimizer of

$$\theta \mapsto G(\theta, \theta_0, \theta_1) = u_* g(\theta, \theta_0) + (1 - u_*) g(\theta, \theta_1) \tag{8}$$

over Θ. Here θ_0, θ_1 and u_* are the same as specified under H_1.

Later on we will suppose that $\sqrt{T}(\hat{\theta}_T - \theta_0) = O_{\mathbf{P}}(1)$ under H_0 and that $\hat{\theta}_T \to^p \theta^*$ under H_1. Let us explain how natural and mild these assumptions are. The former is really standard. The latter follows from Corollary 3.2.3 of van der Vaart and Wellner [22] because, under H_1, the following almost sure convergence holds

$$\frac{1}{T} L_T(\theta) \to \int_{\mathbb{R}} \frac{S(z,\theta_0)^2}{\sigma^2(z)} dF_{\theta_0}(z) + \int_{\mathbb{R}} \frac{S(z,\theta_1)^2}{\sigma^2(z)} dF_{\theta_1}(z) - G(\theta, \theta_0, \theta_1),$$

and under some mild conditions this convergence is uniform in $\theta \in \Theta$.

3 Main result

In this section, we construct a test statistic which is asymptotically distribution free under H_0, and consistent under H_1. Here on we suppose that all the conditions stated in previous Section 2 hold. In order to construct a statistic for this testing problem, we introduce the random field $\{\hat{V}_T(u,x) : (u,x) \in [0,1] \times \mathbb{R}\}$ given by

$$\widehat{V}_T(u,x) = \frac{1}{\sqrt{T}} \int_0^T (\mathbf{1}_{\{s \le Tu\}} - u)\mathbf{1}_{\{X_s \le x\}} \frac{1}{\sigma(X_s)}(dX_s - S(X_s, \hat{\theta}_T)ds),$$

where $\hat{\theta}_T$ satisfies $\sqrt{T}(\hat{\theta}_T - \theta_0) = O_{\mathrm{P}}(1)$ under H_0.

Before stating our results, we illustrate our approach. First we will prove that $\sup_{u,x} |V_T(u,x)|$, where

$$V_T(u,x) = \frac{1}{\sqrt{T}} \int_0^T (\mathbf{1}_{\{s \le Tu\}} - u)\mathbf{1}_{\{X_s \le x\}} \frac{1}{\sigma(X_s)}(dX_s - S(X_s, \theta_0)ds),$$

is asymptotically distribution free under H_0. In order to prove this result in Lemma 1 we obtain the weak convergence of V_T to the random field $(u,x) \rightsquigarrow B^\circ(u, F_{\theta_0}(x))$ where $B^\circ = \{B^\circ(s,t) : (s,t) \in [0,1] \times [0,1]\}$ is a centered Gaussian random field with the covariance

$$\mathbf{E}[B^\circ(s_1,t_1)B^\circ(s_2,t_2)] = (s_1 \wedge s_2 - s_1 s_2)(t_1 \wedge t_2).$$

The crucial point of our approach then will be to prove that $\sup_{u,x} |\hat{V}_T(u,x) - V_T(u,x)| \to^p 0$.

The main result of the paper is the following.

Theorem 1. *(i) Under H_0, if $\sqrt{T}(\hat{\theta}_T - \theta_0) = O_{\mathbf{P}}(1)$, it holds that*

$$\sup_{u,x} |\hat{V}_T(u,x)| \to^d \sup_{(s,t)\in[0,1]^2} |B^\circ(s,t)|$$

where B° is a centered Gaussian random field with the covariance

$$\mathbf{E}[B^\circ(s_1,t_1)B^\circ(s_2,t_2)] = (s_1 \wedge s_2 - s_1 s_2)(t_1 \wedge t_2).$$

(ii) Under H_1, if $\hat{\theta}_T \to^p \theta^*$, it holds that

$$\mathbf{P}\left(\sup_{u,x} |\hat{V}_T(u,x)| > K\right) \to 1, \quad \forall K > 0.$$

Proof. The weak convergence of the random field $\widehat{V}_T$ follows from the weak convercence of the random field V_T to the random field $(u,x) \rightsquigarrow B^\circ(u, F_{\theta_0}(x))$, which is proved in Lemma 1 of Section 4 and from the uniform convergence $\sup_{u,x} |\hat{V}_T(u,x) - V_T(u,x)| \to^p 0$ which is proved in Lemma 2 of Section 4. Now by the continuous mapping theorem $\sup_{u,x} |\hat{V}_T(u,x)| \to^d \sup_{u,x} B^\circ(u, F_{\theta_0}(x))$ and by the following equality in distribution

$$\sup_{(u,x)\in[0,1]\times\mathbb{R}} B^\circ(u, F_{\theta_0}(x)) = \sup_{(u,t)\in[0,1]\times[0,1]} B^\circ(u,t)$$

part (i) is proved. Part (ii) follows from Lemma 3 of Section 4. □

The distribution of the limit process is well known (see for example Brownrigg [2]), so we can reject H_0 at a fixed level $0 < \alpha < 1$ if the test statistics is grater of the critical value c_α given by $\mathbf{P}(\sup_{(s,t)} |B^\circ(s,t)| > c_\alpha) = \alpha$.

4 Auxiliary results

In this section we prove the lemmas needed in the proof of the main result. Let start with the behavior of the random field V_T.

Lemma 1. *The random field V_T converges weakly, as T goes to infinity, in $\ell^\infty([0,1] \times \mathbb{R})$ to the centered Gaussian random field $\tilde{B}^\circ = \{B^\circ(u, F_{\theta_0}(x)) : (u,x) \in [0,1] \times \mathbb{R}\}$.*

Proof. The finite dimensional convergence is immediate from the martingale central limit theorem. Indeed the random field V_T under H_0 can be written as $V_T(u,x) = M_T^{T,(u,x)}$ where

$$M_t^{T,(u,x)} = \frac{1}{\sqrt{T}} \int_0^t (\mathbf{1}_{\{s \le Tu\}} - u) \frac{1}{\sigma(X_s)} (dX_s - S(X_s, \theta_0) ds), \quad t \in [0,T].$$

Let us calculate the quadratic variation:

$$\begin{aligned} \langle M^{T,(u,x)}, M^{T,(v,y)} \rangle_T &= \frac{1}{T} \int_0^T (\mathbf{1}_{\{s \le uT\}} - u)(\mathbf{1}_{\{s \le vT\}} - v) \mathbf{1}_{\{X_s \le x\}} \mathbf{1}_{\{X_s \le y\}} ds \\ &= \frac{1}{T} \int_0^T (\mathbf{1}_{\{s \le (u \wedge v)T\}} - v\mathbf{1}_{\{s \le uT\}} - u\mathbf{1}_{\{s \le vT\}} + uv) \mathbf{1}_{\{X_s \le x \wedge y\}} ds \end{aligned}$$

For the strong law of large number, with probability one it holds

$$\begin{aligned} &\frac{1}{T} \int_0^T (\mathbf{1}_{\{s \le (u \wedge v)T\}} - v\mathbf{1}_{\{s \le uT\}} - u\mathbf{1}_{\{s \le vT\}} + uv) \mathbf{1}_{\{X_s \le x \wedge y\}} ds \\ &\qquad \to (u \wedge v - uv) F_{\theta_0}(x \wedge y). \end{aligned}$$

Hence we have the finite-dimensional convergence of the random field $V_T(u,x)$ to the random field $B^\circ(u, F_{\theta_0}(x))$. For the asymptotic tightness, it is sufficient to show that there exists a semimetric ρ on $[0,1] \times \mathbb{R}$ such that

$$\sup_{\rho((u,x),(v,y))>0} \frac{\sqrt{\langle M^{T,(u,x)} - M^{T,(v,y)} \rangle_T}}{\rho((u,x),(v,y))} = O_{\mathrm{P}}(1) \tag{9}$$

and that

$$\int_0^1 \sqrt{\log N(\varepsilon, [0,1] \times \mathbb{R}, \rho)} d\varepsilon < \infty. \tag{10}$$

Here, we denote by $N(\varepsilon, \mathcal{X}, \rho)$ the smallest number of open balls with ρ-radius

ε which cover the space $\mathcal{X}$. See Nishiyama [16], [17]. Notice that

$$\begin{aligned}
&\langle M^{T,(u,x)} - M^{T,(v,y)} \rangle_T = \\
&= \frac{1}{T}\int_0^T ((\mathbf{1}_{\{s\le uT\}} - u)\mathbf{1}_{\{X_s\le x\}} - (\mathbf{1}_{\{s\le vT\}} - v)\mathbf{1}_{\{X_s\le y\}})^2 ds \\
&\le \frac{1}{T}\int_0^T 2(\mathbf{1}_{\{s\le uT\}} - u)^2(\mathbf{1}_{\{X_s\le x\}} - \mathbf{1}_{\{X_s\le y\}})^2 ds + \\
&\quad + \frac{1}{T}\int_0^T 2\left((\mathbf{1}_{\{s\le uT\}} - u - \mathbf{1}_{\{s\le vT\}} + v)\mathbf{1}_{\{X_s\le y\}}\right)^2 ds \\
&\le \frac{1}{T}\int_0^T 2(\mathbf{1}_{\{X_s\le x\}} - \mathbf{1}_{\{X_s\le y\}})^2 ds + \\
&\quad + \frac{1}{T}\int_0^T 2(\mathbf{1}_{\{s\le uT\}} - \mathbf{1}_{\{s\le vT\}} - (u-v))^2 ds
\end{aligned}$$

The first term is bounded by

$$\begin{aligned}
&\frac{2}{T}\int_0^T \mathbf{1}_{\{x\wedge y\le X_s\le x\vee y\}} ds = \\
&\qquad = \frac{2}{T}\int_{x\wedge y}^{x\vee y} l_T(z) m_{\theta_0}(dz) \le \frac{2}{T}\sup_z l_T(z) m_{\theta_0}([x\wedge y, x\vee y]),
\end{aligned}$$

where $l_T(z)$ is the local time of X with respect to the speed measure m_{θ_0}. For the second term

$$\begin{aligned}
&\frac{1}{T}\int_0^T 2(\mathbf{1}_{\{s\le uT\}} - \mathbf{1}_{\{s\le vT\}} - (u-v))^2 ds \le \\
&\qquad \frac{1}{T}\int_0^T \{4|\mathbf{1}_{\{s\in[T(u\wedge v), T(u\vee v)]\}}|^2 + 4|u-v|^2\} ds \le 8|u-v|.
\end{aligned}$$

Since $\sup_z l_T(z) = O_{\mathrm{P}}(T)$ (see Theorem 4.2 of Van der Vaart and Van Zanten [23] and also Van Zanten [24]), the above conditions (9) and (10) are satisfied if we define $\rho((u,x),(v,y)) = \sqrt{|u-v| \vee m_{\theta_0}([x\wedge y, x\vee y])}$.

□

Lemma 2. *Under H_0, if $\sqrt{T}(\hat{\theta}_T - \theta_0) = O_{\mathbf{P}}(1)$, then $\sup_{u,x} |\hat{V}_T(u,x) - V_T(u,x)| \to^p 0$, as $T \to \infty$.*

Proof. Developing $S(x,\theta)$ around θ_0, and denoting the transpose with a prime, we can write, for a good choice of $\tilde{\theta}$,

$$S(x,\theta) = S(x,\theta_0) + \dot{S}(x,\theta_0)'(\theta - \theta_0) + (\theta - \theta_0)'\ddot{S}(x,\tilde{\theta})(\theta - \theta_0).$$

So we have

$$\begin{aligned}
\hat{V}_T(u,x) - V_T(u,x) = \\
&= \frac{1}{\sqrt{T}} \int_0^T (\mathbf{1}_{\{s \le uT\}} - u)\mathbf{1}_{\{X_s \le x\}} \frac{S(X_s,\theta_0) - S(X_s,\hat{\theta}_T)}{\sigma(X_s)} ds \\
&= \frac{1}{T} \int_0^T (\mathbf{1}_{\{s \le uT\}} - u)\mathbf{1}_{\{X_s \le x\}} \frac{\dot{S}(X_s,\theta_0)'}{\sigma(X_s)} \sqrt{T}(\theta_0 - \hat{\theta}_T) ds + \\
&\quad + \frac{1}{\sqrt{T}} \int_0^T (\mathbf{1}_{\{s \le uT\}} - u)\mathbf{1}_{\{X_s \le x\}} (\theta_0 - \hat{\theta}_T)' \frac{\ddot{S}(X_s,\tilde{\theta})}{\sigma(X_s)} (\theta_0 - \hat{\theta}_T) ds
\end{aligned}$$

The second term, remembering the condition (4), is $o_{\mathbf{P}}(1)$ as T goes to infinity. Let us consider the first term. Let us write $\dot{S}^{(l)}(x,\theta_0) = \dot{S}^{(l)}(x,\theta_0)^+ - \dot{S}^{(l)}(x,\theta_0)^-$, where $\dot{S}^{(l)}(x,\theta_0)^+$ and $\dot{S}^{(l)}(x,\theta_0)^-$ are the positive and the negative part respectively of the l component of $\dot{S}(x,\theta_0)$, $l = 1,\ldots,k$. Let us consider the l component.

For every $\varepsilon > 0$ there exists $N = N(\varepsilon)$, such that one can choose $0 = u_0 < u_1 < \ldots < u_h < \ldots < u_N = 1$, such that for $K = \int \frac{\dot{S}^{(l)}(z,\theta_0)^+}{\sigma(z)} dF_{\theta_0}(z)$, which is finite by (3), it holds $\varepsilon_h = |u_h - u_{h-1}|K < \varepsilon$ for every $h = 1,2,\ldots N$.

For the same ε there exists $M = M(\varepsilon)$, such that one can choose $-\infty = x_0 < x_1 < \ldots < x_j < \ldots < x_M = +\infty$, such that $\varepsilon_j = |\psi(x_j) - \psi(x_{j-1})| < \varepsilon$ for every $j = 1,2,\ldots M$, where $\psi(x) = \int_{-\infty}^x \frac{\dot{S}^{(l)}(z,\theta_0)^+}{\sigma(z)} dF_{\theta_0}(z)$.

Let us consider

$$\frac{1}{T}\int_0^T (\mathbf{1}_{\{s\le uT\}} - u)\mathbf{1}_{\{X_s\le x\}}\frac{\dot S^{(l)}(X_s,\theta_0)^+}{\sigma(X_s)}\sqrt{T}(\theta_0-\hat\theta_T)ds.$$

As $\sqrt{T}(\hat\theta_T-\theta_0)$ is bounded in probability, we consider

$$\begin{aligned}
&\sup_{(x,u)\in\mathbb{R}\times[0,1]}\frac{1}{T}\int_0^T (\mathbf{1}_{\{s\le uT\}} - u)\mathbf{1}_{\{X_s\le x\}}\frac{\dot S^{(l)}(X_s,\theta_0)^+}{\sigma(X_s)}ds\\
&\quad\le \max_{h,j}\left(\frac{1}{T}\int_0^{u_hT}\mathbf{1}_{\{X_s\le x_j\}}\frac{\dot S^{(l)}(X_s,\theta_0)^+}{\sigma(X_s)}ds+\right.\\
&\quad\quad\left.-\frac{u_{h-1}}{T}\int_0^T\mathbf{1}_{\{X_s\le x_{j-1}\}}\frac{\dot S^{(l)}(X_s,\theta_0)^+}{\sigma(X_s)}ds\right)\\
&\quad\to^p \max_{h,j}\left((u_h-u_{h-1})\psi(x_j)+u_{h-1}(\psi(x_j)-\psi(x_{j-1}))\right)\\
&\quad\le \max_{h,j}\left((u_h-u_{h-1})\int_{-\infty}^{+\infty}\frac{\dot S^{(l)}(z,\theta_0)^+}{\sigma(X_s)}dF_{\theta_0}(z)+(\psi(x_j)-\psi(x_{j-1})\right)\\
&\quad\le \max_h|u_h-u_{h-1}|K+\varepsilon<2\varepsilon.
\end{aligned}$$

Analogously we obtain a similar bound from below. The same arguments hold for each component and for the negative part of $\dot S$. So, as ε is arbitrary we get the uniform convergence $\sup_{(x,u)\in\mathbb{R}\times[0,1]}|\hat V_T(u,x)-V_T(u,x)|\to^p 0$. □

The next Lemma gives the behavior of $\hat V_T$ under the alternative.

Lemma 3. *Under H_1, if $\hat\theta_T\to^p\theta^*$, it holds that*

$$\mathbf{P}\left(\sup_{u,x}|\hat V_T(u,x)|>K\right)\to 1,\quad \forall K>0.$$

Proof. Let θ^* be the minimizer of (8), then θ^* has to satisfy the equation

$$u_*\int\frac{(S(\theta_0)-S(\theta))\dot S(\theta)}{\sigma^2}dF_{\theta_0}+(1-u_*)\int\frac{(S(\theta_1)-S(\theta))\dot S(\theta)}{\sigma^2}dF_{\theta_1}=0,\quad (11)$$

with obvious notations. Let us introduce the following function

$$C(x,\theta) = \int_{-\infty}^{x} \frac{S(z,\theta_0) - S(z,\theta)}{\sigma(z)} dF_{\theta_0}(z) - \int_{-\infty}^{x} \frac{S(z,\theta_1) - S(z,\theta)}{\sigma(z)} dF_{\theta_1}(z)$$

and let us consider $C(x,\theta^*)$. We state that it exists at least a x_* such that $C(x_*,\theta^*) \neq 0$. Indeed if $C(x,\theta^*) = 0$ for every x, then

$$(S(x,\theta_0) - S(x,\theta^*)) f_{\theta_0}(x) = (S(x,\theta_1) - S(x,\theta^*)) f_{\theta_1}(x), \tag{12}$$

for every $x \in \mathbb{R}$. Now, taking in account (12) and considering (11), θ^* has to satisfy these two equations

$$\int_{\mathbb{R}} \frac{(S(z,\theta_0) - S(z,\theta^*))\dot{S}(z,\theta^*)}{\sigma^2(z)} dF_{\theta_0}(z) = 0$$

and

$$\int_{\mathbb{R}} \frac{(S(z,\theta_1) - S(z,\theta^*))\dot{S}(z,\theta^*)}{\sigma^2(z)} dF_{\theta_1}(z) = 0.$$

This means that θ^* is the solution of the two following minimizing problems

$$\min_{\theta\in\Theta} \int_{\mathbb{R}} \frac{(S(z,\theta) - S(z,\theta_0))^2}{\sigma^2(z)} dF_{\theta_0}(z)$$

and

$$\min_{\theta\in\Theta} \int_{\mathbb{R}} \frac{(S(z,\theta) - S(z,\theta_1))^2}{\sigma^2(z)} dF_{\theta_1}(z).$$

Now (6) has a unique minimizer, so the solution of each of the last two minimizing problems is unique. This imply that $\theta^* = \theta_0$ and $\theta^* = \theta_1$ but this is impossible as $\theta_0 \neq \theta_1$.

Let us read $\theta_{\text{true}} = \theta_0$ when $t \leq u_* T$ and and $\theta_{\text{true}} = \theta_1$ when $t > u_* T$

Observe that

$$
\begin{aligned}
\sup_{x,u} |\hat{V}_T(u,x)| \geq & |\hat{V}_T(u_*, x_*)| \\
& = \left| \frac{1}{\sqrt{T}} \int_0^T (\mathbf{1}_{\{s \leq T u_*\}} - u_*) \mathbf{1}_{\{X_s \leq x_*\}} \frac{1}{\sigma(X_s)} (dX_s - S(X_s, \hat{\theta}_T) ds) \right| \\
& \geq \left| \frac{1}{\sqrt{T}} \int_0^T (\mathbf{1}_{\{s \leq u_* T\}} - u_*) \mathbf{1}_{\{X_s \leq x_*\}} \frac{S(X_s, \theta_{\text{true}}) - S(X_s, \hat{\theta}_T)}{\sigma(X_s)} ds \right| + \\
& \quad - \left| \frac{1}{\sqrt{T}} \int_0^T (\mathbf{1}_{\{s \leq T u_*\}} - u_*) \frac{1}{\sigma(X_s)} (dX_s - S(X_s, \theta_{\text{true}}) ds) \right|
\end{aligned}
$$

The second term weakly converge to a random variable, so is tight. By (3) and (4) we can easily prove that the almost sure convergence

$$
\begin{aligned}
& \frac{1}{T} \int_0^{u_* T} (1 - u_*) \mathbf{1}_{\{X_s \leq x\}} \frac{S(X_s, \theta_0) - S(X_s, \theta)}{\sigma(X_s)} ds + \\
& \qquad + \frac{1}{T} \int_{u_* T}^T u_* \mathbf{1}_{\{X_s \leq x\}} \frac{S(X_s, \theta_0) - S(X_s, \theta)}{\sigma(X_s)} ds \\
& \to u_*(1 - u_*) C(x, \theta)
\end{aligned}
$$

is uniformly on θ. So, for the first term, it follows that

$$
\frac{1}{T} \int_0^T (\mathbf{1}_{\{s \leq u_* T\}} - u_*) \mathbf{1}_{\{X_s \leq x_*\}} \frac{S(X_s, \theta_{\text{true}}) - S(X_s, \hat{\theta}_T)}{\sigma(X_s)} ds \to^p u_*(1 - u_*) C(x_*, \theta^*).
$$

As $u_*(1 - u_*) C(x_*, \theta^*) \neq 0$, this complete the proof. □

5 Acknowledgements

The first author thanks the Institute of Statistical Mathematics for the financial support during her stay as visiting professor.

References

[1] Brown, R. L., Durbin, J. and Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. With discussion. *J. Roy. Statist. Soc. Ser. B.* **37**, 149-192.

[2] Brownrigg, R.D. (2005). Brownian Motion in Space - Tables of Suprema. At `http://homepages.ecs.vuw.ac.nz/~ray/Brownian/`

[3] Chen, J. and Gupta, A. K. (2000). *Parametric Statistical Change Point Analysis*. Borkäuser, Boston.

[4] Chen, J. and Gupta, A. K. (2001). On change point detection and estimation, *Communications in Statistics - Simulation and Computation*, **30**, 665-697.

[5] Chen, J. (1998). Testing for a change point in linear regression models. *Comm. Statist. Theory Methods*, **27**, 2481-2493.

[6] Csörgő, M. and Horváth L. (1997). *Limit Theorems in Change-Points analysis*. Wiley, New York.

[7] Durrett, R. (1996). *Stochastic Calculus: A Practical Introduction.* CRC Press, Boca Raton.

[8] Gikhman, I.I. and Skorohod, A.V. (1972). *Stochastic Differential Equations*. Springer-Verlag, New York.

[9] Hinkley, D. V. (1969). Inference about the intersection in two-phase regression. *Biometrika.* **56**, 495-504.

[10] Hinkley, D. V. (1970). Inference about the change-point in a sequence of random variables. *Biometrika.* **57**, 1-17.

[11] Itô, K. and McKean, H.P. (1965). *Diffusion Processes and Their Sample Paths*, Springer-Verlag, Berlin.

[12] Kutoyants, Y.A. (2004). *Statistical Inference for Ergodic Diffusion Processes*. Springer-Verlag, New York.

[13] Lee, S., Nishiyama, Y. and Yoshida, N. (2006). Test for parameter change in diffusion processes by cusum statistics based on one-step estimators. *Ann. Inst. Statist. Math.* **58**, 211-222.

[14] Ling, S. (2007). Testing for change points in time series models and limiting theorems for NED sequences. *Ann. Statist.* **35**, 1213-1237.

[15] Negri, I. and Nishiyama, Y. (2009). Goodness of fit test for ergodic diffusion processes. *Ann. Inst. Statist. Math.* **61**, 919-928.

[16] Nishiyama, Y. (1999). A maximal inequality for continuous martingales and M-estimation in a Gaussian white noise model. *Ann. Statist.* **27**, 675-696.

[17] Nishiyama, Y. (2000). *Entropy Methods for Martingales*. CWI Tract **128**, Centrum voor Wiskunde en Informatica, Amsterdam.

[18] Page, E. S. (1955). A test for a change in a parameter occurring at an unknown point. *Biometrika*. **42**, 523-527.

[19] Picard, D. (1985). Testing and estimating change-points in time series. *Adv. in Appl. Probab.* **17**, 841-867.

[20] Quandt, R. E. (1960). Tests of the hypothesis that a linear regression system obeys two separate regimes. *J. Amer. Statist. Assoc.* **55**, 324-330.

[21] Rogers, L.C.G. and Williams, D. (2000). *Diffusion, Markov Processes and Martingales*, Vol. II, 2nd ed. Cambridge University Press, Cambridge.

[22] van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes: With Applications to Statistics*. Springer-Verlag, New York.

[23] van der Vaart, A.W. and van Zanten, H. (2005). Donsker theorems for diffusions: necessary and sufficient conditions. *Ann. Probab.* **33**, 1422-1451.

[24] van Zanten H. (2003). On empirical processes for ergodic diffusions and rates of convergence of M-estimators. *Scandinavian Journal of Statistics.* **30**, 443-458.