



HAL
open science

On the Goodness-of-fit Testing of Composite Hypothesis for Dynamical Systems with Small Noise

A Rabhi

► **To cite this version:**

A Rabhi. On the Goodness-of-fit Testing of Composite Hypothesis for Dynamical Systems with Small Noise. *Annales de l'ISUP*, 2009, LIII (2-3), pp.31-48. hal-03634256

HAL Id: hal-03634256

<https://hal.science/hal-03634256>

Submitted on 7 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Pub. Inst. Stat. Univ. Paris
LIII, fasc. 2-3, 2009, 31 à 48

On the Goodness-of-fit Testing of Composite Hypothesis for Dynamical Systems with Small Noise

A. RABHI

L.S.T.A., Université Pierre et Marie Curie ,
 175, rue du Chevaleret, bât. A, 75013 Paris FRANCE

and

Laboratoire de Statistique et Processus, Université du Maine
 72085 Le Mans, Cédex 9, France

Abstract

The problem of the goodness of-fit testing for stochastic differential equation is considered. A test statistic of the Cramér-von Mises type is proposed and its asymptotic behavior is studied. The asymptotic corresponds to the “small noise” approach, i.e., the diffusion coefficient tends to zero. The basic hypothesis is supposed to be composite parametric and the test statistics depends on the maximum likelihood estimator of the parameter. We consider two types of problems (under hypothesis): smooth (regular) and change point (singular) types. The behaviour of the power function of the corresponding tests under nonparametric alternatives is discussed too.

MSC 2000 Classification : 62M02, 62G10, 62G20.

Keywords: Hypotheses testing, Cramér-von Mises test, diffusion process, goodness-of-fit tests.

1 Introduction

The goodness-of-fit testing is one of the main subjects of the classical mathematical statistics. These tests allow to verify if the mathematical model fits well to the given data. Our goal is to present some goodness-of-fit tests for the dynamical system with small noise. The approach is similar to that used

for the tests of Kolmogorov-Smirnov, Cramér-von Mises and Chi-square (see, e.g., Durbin [6], Lehmann and Romano [14] and Greenwood and Nikulin [9]). In fact the mentioned classical tests are *distribution free*, i.e., their limit distributions do not depend on the basic hypothesis and therefore for different statistical models the choice of the threshold can be the same.

Let us remind the properties of the classical tests. Suppose that we observe n independent identically distributed random variables $(X_1, \dots, X_n) = X^n$ with continuous distribution function $F(x)$ and the basic hypothesis is simple :

$$\mathcal{H}_0 : F(x) = F_*(x), \quad x \in \mathfrak{R}.$$

Then the Cramér-von Mises W_n^2 and Kolmogorov-Smirnov D_n statistics are

$$W_n^2 = n \int_{-\infty}^{\infty} [\hat{F}_n(x) - F_*(x)]^2 dF_*(x), \quad D_n = \sup_x \sqrt{n} |\hat{F}_n(x) - F_*(x)|,$$

where $\hat{F}_n(x)$ is the empirical distribution function. The corresponding Cramér-von Mises and Kolmogorov-Smirnov tests are

$$\psi_n(X^n) = \mathbb{1}_{\{W_n^2 > c_\alpha\}}, \quad \phi_n(X^n) = \mathbb{1}_{\{D_n > d_\alpha\}}$$

and the thresholds c_α, d_α are defined as follows. Let us introduce the Brownian bridge $\{W_0(s), 0 \leq s \leq 1\}$, i.e., continuous Gaussian process with:

$$\mathbf{E}W_0(s) = 0, \quad \mathbf{E}W_0(s)W_0(t) = t \wedge s - st.$$

Then the limit behaviour of these statistics can be described with the help of this process as follows

$$W_n^2 \implies \int_0^1 W_0(s)^2 ds, \quad D_n \implies \sup_{0 \leq s \leq 1} |W_0(s)|.$$

Hence to fix the asymptotic probability $\alpha \in (0, 1)$ of the first type error we take the constants c_α, d_α as solutions of the equations:

$$\mathbf{P}\left\{\int_0^1 W_0(s)^2 ds > c_\alpha\right\} = \alpha, \quad \mathbf{P}\left\{\sup_{0 \leq s \leq 1} |W_0(s)| > d_\alpha\right\} = \alpha.$$

These tests are distribution-free. It can be shown that they are consistent against any fixed alternative (see Durbin [6]).

Remind that in i.i.d. case with the basic parametric hypothesis

$$\mathcal{H}_0 : F(x) = F_*(\vartheta, x), \quad \vartheta \in \Theta, \quad x \in \mathfrak{R}$$

the Cramér-von Mises and Kolmogorov-Smirnov tests are no more distribution free see [6].

For a somewhat different hypothesis \mathcal{H}_0 ,

$$\mathcal{H}_0 : F(x) = F_*(\vartheta_0, x), \quad \text{for a specified } \vartheta_0 \in \Theta, \quad x \in \mathbb{R}$$

the following test of Cramér [3], Smirnov [16], is available, and the test function is then

$$W_n^2 = n \int_{-\infty}^{+\infty} [F_n(x) - F(x, \vartheta_0)]^2 dF(x, \vartheta_0),$$

The limiting distribution of W_n^2 , if \mathcal{H}_0 is true, was given by Smirnov; it has been tabulate [1]. This test is distribution free for all n , and it is consistent.

In an effort to modify the W_n^2 test to treat the hypothesis \mathcal{H}_0 of (1), Darling [5] propose the test function

$$C_n^2 = n \int_{-\infty}^{+\infty} [F_n(x) - F(x, \hat{\vartheta}_n)]^2 dF(x, \hat{\vartheta}_n)$$

where $\hat{\vartheta}_n$ is a suitable estimator for the unknown parameter ϑ in $F(x, \vartheta)$, and by some condition we have

$$\lim_{n \rightarrow \infty} \mathbf{P}\{C_n^2 < x\} = \mathbf{P}\left\{\int_0^1 Y^2(u) du < x\right\},$$

where $Y(u)$ is a Gaussian process with mean 0 and covariance $\rho(s, t) = k(u, v) - \varphi(u)\varphi(v)$, see section 4 in Darling [5].

Note that the limit distribution of C_n not depend of the parameter ϑ .

This work is a continuation of the study started in [4], where the following goodness of fit test is proposed. Let us consider the problem of hypotheses testing for the stochastic differential equation

$$dX_t = S(X_t) dt + \varepsilon dW_t, \quad X_0 = x_0, \quad 0 \leq t \leq T,$$

where W_t , $0 \leq t \leq T$ is a Wiener process. The basic hypothesis is simple

$$\mathcal{H}_0 : S(x) = S_*(x), \quad x_0 \leq x \leq x_T^*,$$

where $S_*(\cdot)$ is known function and the alternative, as usual in goodness of fit testing, is nonparametric:

$$\mathcal{H}_1 : S(x) \neq S_*(x), \quad x_0 \leq x \leq x_T^*.$$

Here $S_*(x) > 0$ for $x > x_0$ and x_t^* is the solution of the ordinary differential equation under the hypothesis \mathcal{H}_0 :

$$\frac{dx_t^*}{dt} = S_*(x_t^*), \quad x_0, \quad 0 \leq t \leq T.$$

Then the following Cramér-von Mises (W_ε^2) and Kolmogorov-Smirnov (D_ε) type statistics are proposed :

$$W_\varepsilon^2 = \left[\int_0^T \frac{dt}{S_*(x_t^*)^2} \right]^{-2} \int_0^T \left(\frac{X_t - x_t^*}{\varepsilon S_*(x_t^*)^2} \right)^2 dt,$$

$$D_\varepsilon = \left[\int_0^T \frac{dt}{S_*(x_t^*)^2} \right]^{-1/2} \sup_{0 \leq t \leq T} \left| \frac{X_t - x_t^*}{\varepsilon S_*(x_t^*)^2} \right|.$$

It is shown that these statistics converge (as $\varepsilon \rightarrow 0$) to the following functionals

$$W_\varepsilon^2 \Rightarrow \int_0^1 W(s)^2 ds, \quad D_\varepsilon \Rightarrow \sup_{0 \leq t \leq T} |W(s)|,$$

where $\{W(s), 0 \leq s \leq 1\}$ is the Wiener process. Hence the corresponding tests

$$\psi_\varepsilon(X^\varepsilon) = \mathbb{I}_{\{W_\varepsilon^2 > c_\alpha\}}, \quad \phi_\varepsilon(X^\varepsilon) = \mathbb{I}_{\{\varepsilon^{-1} D_\varepsilon > d_\alpha\}}$$

with the constants c_α, d_α defined by the equations:

$$\mathbf{P}\left\{ \int_0^1 W(s)^2 ds > c_\alpha \right\} = \alpha, \quad \mathbf{P}\left\{ \sup_{0 \leq s \leq t} |W(s)| > d_\alpha \right\} = \alpha$$

are of asymptotic size α . Note that the choice of the thresholds does not depend on the hypothesis (distribution free) and these tests are consistent against any fixed alternative.

At the present work, we consider the similar problem but with the composite basic hypothesis. We suppose that under hypothesis the trend coefficient depends on the unknown one-dimensional parameter, i.e.,

$$dX_t = S(\vartheta, X_t) dt + \varepsilon dW_t, \quad X_0 = x_0, \quad 0 \leq t \leq T,$$

and we describe the limit ($\varepsilon \rightarrow 0$) distributions of the Cramér-von Mises type statistic in this parametric case in two situations: smooth and change-point type dependence of the trend coefficient on the unknown parameter.

2 Model of observation

Suppose that we have an observation $X^\varepsilon = \{X_t, 0 \leq t \leq T\}$ in continuous time of the solution of the stochastic differential equation:

$$dX_t = S(X_t) dt + \varepsilon dW_t, \quad X_0 = x_0, \quad 0 \leq t \leq T, \quad (1)$$

where $\{W_t, 0 \leq t \leq T\}$ is the Wiener process and x_0 is deterministic initial value. The trend coefficient is in general unknown and usually the statistical problems concern this coefficient. We suppose that the trend coefficient satisfies the global Lipschitz condition, i.e., *there exist a constant L such that*

$$|S(x) - S(y)| \leq L|x - y|. \quad (2)$$

for all $x, y \in \mathbb{R}$. This condition provides the existence and uniqueness of the strong solution of the equation (1) [15]. We consider the statistical problem of hypotheses testing in the asymptotics of *small noise*, i.e., $\varepsilon \rightarrow 0$. It is well known that the process $X = \{X_t, 0 \leq t \leq T\}$ converges uniformly in $t \in [0, T]$ to the deterministic function $\{x_t, 0 \leq t \leq T\}$, which is solution of the ordinary differential equation (see [8], [10]):

$$\frac{dx_t}{dt} = S(x_t), \quad x_0, \quad 0 \leq t \leq T. \quad (3)$$

We have two hypotheses: the basic hypothesis is

$$\mathcal{H}_0 : S(x) \in \mathcal{F}_\Theta,$$

where $\mathcal{F}_\Theta = \{S(\vartheta, x), x \in [x_0, x_T(\vartheta)], \vartheta \in \Theta = [a, b]\}$, i.e., we suppose that under this hypothesis the observations come from the stochastic differential equation

$$dX_t = S(\vartheta, X_t) dt + \varepsilon dW_t, \quad X_0 = x_0, \quad 0 \leq t \leq T,$$

where $\vartheta \in \Theta$ and $S(\cdot, \cdot)$ is some known positive function. The limit function $x_t, 0 \leq t \leq T$ in (3) now depends on ϑ , i.e., $x_t = x_t(\vartheta)$ because it satisfies the equation

$$\frac{\partial x_t}{\partial t} = S(\vartheta, x_t), \quad x_0, \quad 0 \leq t \leq T. \quad (4)$$

The test statistic below is constructed with the help of the maximum likelihood estimator, MLE $\hat{\vartheta}_\varepsilon$, which is introduced as usual by the equation

$$L(\hat{\vartheta}_\varepsilon, X^\varepsilon) = \sup_{\vartheta \in \Theta} L(\vartheta, X^\varepsilon).$$

Here $L(\vartheta, X^\varepsilon)$ is the likelihood ratio function defined by

$$L(\vartheta, X^\varepsilon) = \exp \left\{ \int_0^T \frac{S(\vartheta, X_t)}{\varepsilon^2} dX_t - \int_0^T \frac{S(\vartheta, X_t)^2}{2\varepsilon^2} dt \right\}, \quad \vartheta \in \Theta.$$

The properties of the MLE for this model of observations are studied in [10].
As alternative we take

$$\mathcal{H}_1 : S(x) \notin \mathcal{F}_\Theta.$$

We have a parametric hypothesis against the nonparametric alternative.

3 Smooth case

3.1 Hypothesis

First we consider the hypotheses testing problem in the regular case, i.e., we suppose that the trend coefficient of the observed stochastic process is a smooth function of the unknown parameter.

Let us denote by \mathcal{A} the following conditions (under \mathcal{H}_0):

1. The function $S(\vartheta, x)$ has two continuous bounded derivatives w.r.t. ϑ and x .
2. The trend coefficient $S(\cdot) > 0$ and the Fisher information

$$I(\vartheta) = \int_0^T \dot{S}(\vartheta, x_t)^2 dt > 0.$$

3. The function

$$G(y, \vartheta) = \int_0^T [S(y, x_t(\vartheta)) - S(\vartheta, x_t(\vartheta))]^2 dt, \quad y \in \Theta,$$

has a unique minimum on $y \in \Theta$ at point $y = \vartheta$.

We denote by $\dot{S}(\cdot, \cdot)$ the derivative of $S(\vartheta, x)$ with respect to ϑ , and by $S'(\cdot, \cdot)$ the derivative of $S(\vartheta, x)$ with respect to x .

Note that these conditions of regularity allow us to show the consistency

$$\mathbf{P}_\vartheta - \lim_{\varepsilon \rightarrow 0} \hat{\vartheta}_\varepsilon = \vartheta,$$

and the asymptotic normality (for $\vartheta \in (a, b)$) of the maximum likelihood estimator:

$$\mathcal{L}_\vartheta\{\varepsilon^{-1}I(\vartheta)^{1/2}(\hat{\vartheta}_\varepsilon - \vartheta)\} \implies \mathcal{N}(0, 1), \quad (5)$$

(see [10], Theorem 2.2).

To construct the test we use the following statistic:

$$\delta_\varepsilon = \int_0^T \left(\frac{X_t - x_t(\hat{\vartheta}_\varepsilon)}{\varepsilon} \right)^2 dt.$$

Below $c_\alpha(\vartheta)$ is solution of the equation

$$\mathbf{P} \left\{ \int_0^T \xi_t^2(\vartheta) dt > c_\alpha(\vartheta) \right\} = \alpha,$$

ξ_t is a stochastic integral defined by

$$\xi_t(\vartheta) = \int_0^t \Phi_\vartheta(s, t) dW_s, \quad (6)$$

where

$$\Phi_\vartheta(s, t) = \frac{S(\vartheta, x_t)}{S(\vartheta, x_s)} \mathbb{I}_{\{0 < s < t\}} - \dot{x}_t(\vartheta) \frac{\dot{S}(\vartheta, x_s)}{I(\vartheta)}. \quad (7)$$

Let us fix a number $\alpha \in (0, 1)$ and define the class \mathcal{K}_α of tests of asymptotic level $1 - \alpha$:

$$\mathcal{K}_\alpha = \{ \Psi(\cdot), \lim_{\varepsilon \rightarrow 0} \mathbf{E}_\vartheta \psi(X^\varepsilon) = \alpha, \vartheta \in \Theta \}.$$

Remind that $\mathbf{E}_\vartheta \Psi(X^\varepsilon)$ is the probability to reject the hypothesis \mathcal{H}_0 .

Theorem 3.1 *Let the condition \mathcal{A} be fulfilled, then the test*

$$\Psi = \mathbb{I}_{\{\delta_\varepsilon > c_\alpha(\hat{\vartheta}_\varepsilon)\}} \in \mathcal{K}_\alpha.$$

Proof. Let us write

$$\frac{X_t - x_t(\hat{\vartheta}_\varepsilon)}{\varepsilon} = \frac{X_t - x_t(\vartheta)}{\varepsilon} + \frac{x_t(\vartheta) - x_t(\hat{\vartheta}_\varepsilon)}{\varepsilon}$$

and study these two terms separately. Here ϑ is the true value.

Recall that the stochastic process $\varepsilon^{-1}(X_t - x_t(\vartheta))$ converges uniformly on $t \in [0, T]$ to the limit Gaussian process $x_t^{(1)} = x_t^{(1)}(\vartheta)$, solution of the linear equation:

$$dx_t^{(1)} = S'(\vartheta, x_t)x_t^{(1)} dt + dW_t, \quad x_0^{(1)} = 0. \quad (8)$$

This solution can be written explicitly

$$x_t^{(1)} = \int_0^t e^{\int_s^t S'(\vartheta, x_u) du} dW_s.$$

Using the equation (4), we can write

$$\begin{aligned} \int_s^t S'(\vartheta, x_u) du &= \int_s^t \frac{S'(\vartheta, x_u)}{S(\vartheta, x_u)} S(\vartheta, x_u) du = \int_s^t \frac{S'(\vartheta, x_u)}{S(\vartheta, x_u)} dx_u = \\ &= \int_{x_s}^{x_t} \frac{S'(\vartheta, v)}{S(\vartheta, v)} dv = \int_{x_s}^{x_t} (\ln S(\vartheta, v))' dv = \ln \frac{S(\vartheta, x_t)}{S(\vartheta, x_s)}. \end{aligned}$$

For the details see [10], p. 136.

Hence

$$x_t^{(1)} = \int_0^t \frac{S(\vartheta, x_t)}{S(\vartheta, x_s)} dW_s.$$

We can write

$$\frac{X_t - x_t(\vartheta)}{\varepsilon} \longrightarrow \int_0^t \frac{S(\vartheta, x_t)}{S(\vartheta, x_s)} dW_s. \quad (9)$$

To study the convergence of $\frac{x_t(\hat{\vartheta}_\varepsilon) - x_t(\vartheta)}{\varepsilon}$, we need the following lemma:

Lemma 3.1 *The function $x_t(\vartheta)$ is differentiable with respect to ϑ and the derivative is*

$$\dot{x}_t(\vartheta) = S(\vartheta, x_t) \int_0^t \frac{\dot{S}(\vartheta, x_s)}{S(\vartheta, x_s)} ds. \quad (10)$$

The proof is given in the appendix below.

According to this lemma we can write

$$\frac{x_t(\hat{\vartheta}_\varepsilon) - x_t(\vartheta)}{\varepsilon} = \left(\frac{\hat{\vartheta}_\varepsilon - \vartheta}{\varepsilon} \right) \dot{x}_t(\tilde{\vartheta}),$$

where $\tilde{\vartheta}$ satisfies $|\tilde{\vartheta} - \vartheta| \leq |\hat{\vartheta}_\varepsilon - \vartheta|$, and from the consistency of the MLE it follows that

$$\sup_{t \in [0, T]} |\dot{x}_t(\tilde{\vartheta}) - \dot{x}_t(\vartheta)| \longrightarrow 0.$$

The asymptotic normality of the MLE (5) yields the weak convergence

$$\frac{x_t(\vartheta) - x_t(\hat{\vartheta}_\varepsilon)}{\varepsilon} \longrightarrow \frac{\dot{x}_t(\vartheta)}{I(\vartheta)} \int_0^T \dot{S}(\vartheta, x_t) dW_t. \quad (11)$$

It can be shown (by a strangling of the proof of the Theorem 2.2 in [10]) that we have the joint asymptotic normality of the MLE and the derivative $x_t^{(1)}$. Hence, using (11) and (9), we obtain

$$\begin{aligned}
 \delta_\varepsilon &\longrightarrow \int_0^T \left[\int_0^t e^{\int_s^t S'(\vartheta, x_u) du} dW_s - \frac{\dot{x}_t(\vartheta)}{I(\vartheta)} \int_0^T \dot{S}(\vartheta, x_t) dW_t \right]^2 dt \\
 &\longrightarrow \int_0^T \left[\int_0^t [e^{\int_s^t S'(\vartheta, x_u) du} \mathbb{I}_{\{0 < s < t\}} - \dot{x}_t(\vartheta) \frac{\dot{S}(\vartheta, x_s)}{I(\vartheta)}] dW_s \right]^2 dt \\
 &\longrightarrow \int_0^T \left[\int_0^t \left[\frac{S(\vartheta, x_t)}{S(\vartheta, x_s)} \mathbb{I}_{\{0 < s < t\}} - \dot{x}_t(\vartheta) \frac{\dot{S}(\vartheta, x_s)}{I(\vartheta)} \right] dW_s \right]^2 dt \\
 &\longrightarrow \int_0^T \left[\int_0^t \Phi_\vartheta(s, t) dW_s \right]^2 dt = \int_0^T \xi_t(\vartheta)^2 dt \equiv \xi(\vartheta)^2.
 \end{aligned} \tag{12}$$

The function $\Phi_\vartheta(s, t)$ was defined above (see (7)) and $\xi_t(\vartheta) = \int_0^T \Phi_\vartheta(s, t) dW_s$ is a stochastic integral with a null mean and the covariance function which depends to ϑ . Hence this test is not distribution free even asymptotically. The continuity of the function $c(\vartheta)$ can be verified as follows. Let us denote by $F(\vartheta, x)$ the distribution function of $\xi(\vartheta)^2$. Then $c_\alpha(\vartheta)$ is an implicate function defined by the equation

$$1 - F(\vartheta, c) = \alpha.$$

Hence its derivative

$$\dot{c}(\vartheta) = - \frac{\dot{F}(\vartheta, c)}{f(\vartheta, c)},$$

where $f(\vartheta, x)$ is the density function of the r.v. ξ . Now the continuity of $c(\vartheta)$ follows from the existence of these two functions and $f(\vartheta, c_\alpha(\vartheta)) > 0$.

The limit expression can be slightly simplified if we consider another test function. Let us introduce a family of stochastic processes $\{\hat{X}_t(\vartheta), \vartheta \in \Theta\}$ by the relation

$$\hat{X}_t(\vartheta) = x_0 + \int_0^t S(\vartheta, X_s) ds, \quad 0 \leq t \leq T,$$

i.e., $\hat{X}_t(\vartheta)$ is a solution of the differential equation

$$\frac{d\hat{X}_t(\vartheta)}{dt} = S(\vartheta, X_t), \quad \hat{X}_0(\vartheta) = x_0, \quad 0 \leq t \leq T.$$

Note that such family was used in [12], for construction of a class of trajectory fitting estimators. Introduce the statistic

$$\tilde{\delta}_\varepsilon = \int_0^T \left(\frac{X_t - \hat{X}_t(\hat{\vartheta}_\varepsilon)}{\varepsilon} \right)^2 dt$$

and the test $\tilde{\psi}(X^\varepsilon) = \mathbb{I}_{\{\tilde{\delta}_\varepsilon > \tilde{c}(\hat{\vartheta}_\varepsilon)\}}$. Then the first term in the representation

$$\frac{X_t - \hat{X}_t(\hat{\vartheta}_\varepsilon)}{\varepsilon} = \frac{X_t - \hat{X}_t(\vartheta)}{\varepsilon} + \frac{X_t(\vartheta) - \hat{X}_t(\hat{\vartheta}_\varepsilon)}{\varepsilon}$$

is $(X_t - \hat{X}_t(\vartheta))/\varepsilon = W_t$. For the second term we have

$$\frac{X_t(\vartheta) - \hat{X}_t(\hat{\vartheta}_\varepsilon)}{\varepsilon} = \frac{\vartheta - \hat{\vartheta}_\varepsilon}{\varepsilon} \tilde{x}_t(\vartheta) (1 + o(1)), \quad \tilde{x}_t(\vartheta) = \int_0^t \dot{S}(\vartheta, x_s) ds.$$

Hence

$$\tilde{\delta}_\varepsilon \longrightarrow \int_0^T \left[W_t - \frac{\tilde{x}_t(\vartheta)}{I(\vartheta)} \int_0^T \dot{S}(\vartheta, x_s) dW_s \right]^2 dt.$$

This last expression have to be used in the definition of the threshold $\tilde{c}_\alpha(\vartheta)$.

3.2 Alternative

Suppose now that $S \notin \mathcal{F}_\Theta$ with $\Theta = [a, b]$ but the function $S(x)$, as before, has two continuous bounded derivative. Then the solution of the corresponding equation :

$$dX_t = S(X_t)dt + \varepsilon dW_t, \quad X_0 = x_0, \quad 0 \leq t \leq T$$

converges to another determinist function y_t , which is the solution of the following equation:

$$\frac{dy_t}{dt} = S(y_t), \quad y_0 = x_0.$$

We now verify that the test $\psi(X^\varepsilon) = \mathbb{I}_{\{\delta_\varepsilon > c_\alpha(\hat{\vartheta}_\varepsilon)\}}$ is consistent.

Let us introduce the quantity

$$g_S^2 = \inf_{\vartheta \in \Theta} \int_0^T (y_t - x_t(\vartheta))^2 dt = \inf_{\vartheta \in \Theta} \|y_t - x_t(\vartheta)\|^2 \quad (13)$$

and note that $g_S^2 > 0$. Indeed, if $g_S^2 = 0$, then there exists a value $\vartheta \in \Theta$ such that $y_t = x_t(\vartheta)$ for all $t \in [0, T]$, which contradicts the condition $S \notin \mathcal{F}_\Theta$.

We have

$$\sqrt{\delta_\varepsilon} = \left\| \frac{X_t - y_t}{\varepsilon} + \frac{y_t - x_t(\hat{\vartheta}_\varepsilon)}{\varepsilon} \right\| \geq \left\| \frac{X_t - y_t}{\varepsilon} \right\| - \left\| \frac{y_t - x_t(\hat{\vartheta}_\varepsilon)}{\varepsilon} \right\|. \quad (14)$$

Remind that

$$\sup_{t \in [0, T]} \left| \frac{X_t - y_t}{\varepsilon} - y_t^{(1)} \right| \rightarrow 0, \quad (15)$$

where $y_t^{(1)}$ is a Gaussian linear process:

$$dy_t^{(1)} = S'(y_t)y_t^{(1)} dt + dW_t, \quad y_0^{(1)} = 0, \quad 0 \leq t \leq T.$$

Hence the first term in (14) is bounded in probability. For the second term we have the estimate

$$\left\| \frac{y_t - x_t(\hat{\vartheta}_\varepsilon)}{\varepsilon} \right\| \geq \frac{g_S}{\varepsilon}. \quad (16)$$

Hence by (15) and (16) we obtain,

$$\sqrt{\delta_\varepsilon} \rightarrow \infty.$$

Therefore

$$\mathbf{P}_S \left\{ \delta_\varepsilon > c_\alpha \left(\hat{\vartheta}_\varepsilon \right) \right\} \rightarrow 1 \quad (17)$$

as $\varepsilon \rightarrow 0$ for all $S(\cdot) \notin \mathcal{F}_\Theta$. Hence, this test is consistent against any fixed alternative.

Remark. It can be shown, that like the case of simple basic hypothesis studied in [4], this test is not uniformly consistent and

$$\lim_{\varepsilon \rightarrow 0} \inf_{S \notin \mathcal{F}_\Theta} \mathbf{P}_S \left\{ \delta_\varepsilon > c_\alpha \left(\hat{\vartheta}_\varepsilon \right) \right\} = \alpha.$$

3.3 Example of Ornstein-Uhlenbeck Process

Suppose now that under the basic hypothesis \mathcal{H}_0 the observed process is Ornstein-Uhlenbeck :

$$dX_t = \vartheta X_t dt + \varepsilon dW_t, \quad X_0 = x_0, \quad 0 \leq t \leq T \quad (18)$$

where $\vartheta \in (a, b)$, $\vartheta \neq 0$, $|a| + |b| < \infty$ and $x_0 \neq 0$. The set $\mathcal{F}_\Theta = \{S(x) = \vartheta x, x \in [x_0, x_T(\vartheta)], \vartheta \in \Theta\}$. Here $x_T(\vartheta) = x_0 e^{\vartheta T}$.

Hence we have the hypotheses testing problem

$$\mathcal{H}_0: S(x) \in \mathcal{F}_\vartheta \quad \text{against} \quad \mathcal{H}_1: S(x) \notin \mathcal{F}_\vartheta.$$

The solution of limit equation (under \mathcal{H}_0)

$$\frac{\partial x_t}{\partial t} = \vartheta x_t(\vartheta), \quad x_0, \quad 0 \leq t \leq T.$$

is, of course, $x_t(\vartheta) = x_0 e^{\vartheta t}$.

The goodness of fit test studied in the Section 3.1 is not distribution free and here we consider a different statement of the problem.

Let us suppose that T is large and we try to find the limit distribution (as $T \rightarrow \infty$) of the limit statistics obtained for $\varepsilon \rightarrow 0$ in first step and $T \rightarrow \infty$ in second step (see [11]).

Let us introduce the following statistics:

$$\gamma_{\varepsilon, T} = \left[\int_0^T e^{-2\hat{\vartheta}_\varepsilon t} dt \right]^{-2} \int_0^T \left(\frac{X_t - x_t(\hat{\vartheta}_\varepsilon)}{\varepsilon e^{2\hat{\vartheta}_\varepsilon t}} \right)^2 dt, \quad (19)$$

where the MLE $\hat{\vartheta}_\varepsilon = \left(\int_0^T X_t^2 dt \right)^{-1} \int_0^T X_t dX_t$.

Below c_α is solution of the following equation

$$\mathbf{P} \left\{ \int_0^1 w_z^2 dz > c_\alpha \right\} = \alpha,$$

where $w_z, 0 < z < 1$ is a Wiener process.

Proposition 3.2 *The test $\phi_\varepsilon = \mathbb{I}_{\{\gamma_{\varepsilon, T} > c_\alpha\}}$ belongs to the class \mathcal{K}_α in the following sense*

$$\lim_{T \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} \mathbf{P}_\vartheta \{ \gamma_{\varepsilon, T} > c_\alpha \} = \alpha, \quad \forall \vartheta \neq 0.$$

Proof. According to the equation (12), we have (under \mathcal{H}_0) the following limit

$$\begin{aligned} \int_0^T \left(\frac{X_t - x_t(\hat{\vartheta}_\varepsilon)}{\varepsilon} \right)^2 dt &\longrightarrow \int_0^T e^{2\vartheta t} \left[\int_0^T e^{-\vartheta s} (\mathbb{I}_{\{s < t\}} - \frac{t x_0^2}{I_T(\vartheta)} e^{2\vartheta s}) dW_s \right]^2 dt \\ &\equiv J_T(\vartheta), \end{aligned}$$

where the Fisher information $I_T(\vartheta) = \frac{x_0^2}{2\vartheta} (e^{2\vartheta T} - 1)$. It is easy to see that as $T \rightarrow \infty$ we obtain the relations

$$I_T(\vartheta) \simeq \begin{cases} \frac{x_0^2}{2\vartheta} e^{2\vartheta T}, & \vartheta > 0 \\ -\frac{x_0^2}{2\vartheta}, & \vartheta < 0. \end{cases}$$

and for any $\vartheta \in (a, b)$, $\vartheta \neq 0$, we have

$$e^{-\vartheta s} (\mathbb{1}_{\{s < t\}} - \frac{t x_0^2}{I_T(\vartheta)} e^{2\vartheta s}) \simeq e^{-\vartheta s} \mathbb{1}_{\{s < t\}}.$$

Hence for large T

$$J_T(\vartheta) \simeq \int_0^T e^{2\vartheta t} \left(\int_0^t e^{-\vartheta s} dW_s \right)^2 dt,$$

and we can write (as $\varepsilon \rightarrow 0$ and large T), see [11])

$$\begin{aligned} \int_0^T \left(\frac{X_t - x_t(\hat{\vartheta}_\varepsilon)}{\varepsilon e^{2\hat{\vartheta}_\varepsilon t}} \right)^2 dt &\implies \int_0^T e^{-2\vartheta t} W^2 \left(\int_0^t e^{-2\vartheta s} ds \right) dt (1 + o(1)) \\ &= \int_0^T W^2 \left(\int_0^t e^{-2\vartheta s} ds \right) d \left(\int_0^t e^{-2\vartheta s} ds \right) (1 + o(1)) \\ &= \int_0^{Q_T} W(u)^2 du (1 + o(1)) = Q_T^2 \int_0^1 w_v^2 dv (1 + o(1)), \end{aligned}$$

where

$$u = \int_0^t e^{-2\vartheta s} ds = v \int_0^T e^{-2\vartheta s} ds = v Q_T.$$

Here $w_v = Q_T^{-1/2} W(Q_T v)$, $0 \leq v \leq 1$ is a Wiener process.

Hence (as $\varepsilon \rightarrow 0$ and then $T \rightarrow \infty$)

$$\gamma_{\varepsilon, T} \longrightarrow \int_0^1 w_v^2 dv.$$

Now the choice of thresholds c_α does not depend on the hypothesis and the test is distribution free.

Remark. Note that for $\vartheta = 0$, and under the null hypothesis, the solution of the limit equation is

$$x_t = x_0,$$

hence for large T

$$J_T \simeq \int_0^T \left(\frac{X_t - x_0}{\varepsilon} \right)^2 dt \longrightarrow \int_0^T W_t^2 dt.$$

and

$$\frac{1}{T^2} \int_0^T \left(\frac{X_t - x_0}{\varepsilon} \right)^2 dt \longrightarrow \int_0^1 w_u^2 du.$$

where w_u , $0 \leq u \leq 1$ is the Wiener process.

4 Change point testing (Non smooth case)

4.1 Hypothesis

Suppose that under the null hypothesis \mathcal{H}_0 we observe a trajectory $X^\varepsilon = \{X_t, 0 \leq t \leq T\}$ of diffusion process

$$dX_t = [g(X_t)1_{\{t < \vartheta\}} + h(X_t)1_{\{t \geq \vartheta\}}]dt + \varepsilon dW_t, \quad X_0 = x_0, \quad 0 \leq t \leq T,$$

where $\vartheta \in \Theta = [a, b]$, $0 < a < b < T$. The parameter ϑ is an unknown instant of switching of the trend coefficient (change point). The set of possible trend coefficients (under this hypothesis) is

$$\mathcal{F}_\Theta = \{S(\vartheta, x_t) = g(x_t)1_{\{t < \vartheta\}} + h(x_t)1_{\{t \geq \vartheta\}}, x_t \in [x_0, x_T(\vartheta)], \vartheta \in \Theta\},$$

where $x_t = x_t(\vartheta)$ is the solution of the deterministic switching system:

$$\frac{dx_t}{dt} = \begin{cases} g(x_t), & t < \vartheta \\ h(x_t), & t \geq \vartheta. \end{cases}$$

Let us introduce the **condition** \mathcal{B} :

The functions $h(x)$ and $g(x)$ have continuous bounded non negatives derivatives and $h(x) > g(x) > 0$ for $x \in [x_0, x_T(a)]$.

Note that $\sup_{\vartheta \in \Theta} x_T(\vartheta) = x_T(a)$.

We test the following two hypotheses:

$$\mathcal{H}_0: \quad S(x) \in \mathcal{F}_\Theta \quad \text{against} \quad \mathcal{H}_1: \quad S(x) \notin \mathcal{F}_\Theta.$$

To construct the test, we use the statistic :

$$\Gamma_\varepsilon = [V_T(\hat{\vartheta}_\varepsilon)]^{-2} \int_0^T D(x_t, \hat{\vartheta}_\varepsilon) \left[\frac{X_t - x_t(\hat{\vartheta}_\varepsilon)}{\varepsilon} \right]^2 dt, \quad (20)$$

where

$$V_T(\vartheta) = g(x_T)^2 \int_0^\vartheta \frac{ds}{g(x_s)^2} + h(x_T)^2 \int_\vartheta^T \frac{ds}{h(x_s)^2},$$

$$D(x_t, \vartheta) = 2g(x_t)^2 g'(x_t) \int_0^\vartheta \frac{ds}{g(x_s)^2} + 2h(x_t)^2 h'(x_t) \int_\vartheta^t \frac{ds}{h(x_s)^2} + 1,$$

and $\hat{\vartheta}_\varepsilon$ is the MLE.

The constant c_α is solution of the equation:

$$\mathbf{P} \left\{ \int_0^1 w_z^2 dz > c_\alpha \right\} = \alpha, \quad (21)$$

where w_z , $0 \leq z \leq 1$ is some Wiener process. Note that the choice of the threshold c_α does not depend on the hypothesis (distribution free).

Theorem 4.1 *Let the condition \mathcal{B} be fulfilled, then the test $\psi_\varepsilon = \mathbf{1}_{\{\Gamma_\varepsilon > c_\alpha\}}$ belongs to the class \mathcal{K}_α*

Proof. Let us write

$$\frac{X_t - x_t(\hat{\vartheta}_\varepsilon)}{\varepsilon} = \frac{X_t - x_t(\vartheta)}{\varepsilon} - \frac{x_t(\hat{\vartheta}_\varepsilon) - x_t(\vartheta)}{\varepsilon},$$

and remind that the first term converges uniformly on $t \in [0, T]$ to the Gaussian process (see details in [10], p. 160)

$$\frac{X_t - x_t(\vartheta)}{\varepsilon} \rightarrow x_t^{(1)}(\vartheta), \quad (22)$$

where $x_t^{(1)}$, is a solution of the following equation for $0 \leq t < \vartheta$,

$$dx_t^{(1)} = g'(x_t) x_t^{(1)} dt + dW_t, \quad x_0^{(1)} = 0,$$

Hence

$$x_t^{(1)} = \int_0^t \frac{g(x_s)}{g(x_s)} dW_s, \quad 0 \leq t < \vartheta. \quad (23)$$

and for $t \geq \vartheta$, $x_t^{(1)}$ is a solution of another equation

$$dx_t^{(1)} = h'(x_t) x_t^{(1)} dt + dW_t,$$

and

$$x_t^{(1)} = x_\vartheta^{(1)} + \int_\vartheta^t \frac{h(x_s)}{h(x_s)} dW_s, \quad t \geq \vartheta. \quad (24)$$

To study the term $x_t(\hat{\vartheta}_\varepsilon) - x_t(\vartheta)$ we need the derivative $\partial x_t(\vartheta) / \partial \vartheta$. Its existence can be shown by the following way. We have

$$x_t(\vartheta) = \begin{cases} x_0 + \int_0^t g(x_s) ds, & t < \vartheta, \\ x_0 + \int_0^\vartheta g(x_s) ds + \int_\vartheta^t h(x_s) ds, & t \geq \vartheta \end{cases}$$

Hence

$$\frac{\partial x_t(\vartheta)}{\partial \vartheta} = \begin{cases} 0, & t < \vartheta \\ g(x_\vartheta) - h(x_\vartheta), & t \geq \vartheta \end{cases}$$

then if we suppose that $\vartheta < \hat{\vartheta}_\varepsilon$, we obtain

$$\begin{aligned} \int_0^T \left(\frac{X_t - x_t(\hat{\vartheta}_\varepsilon)}{\varepsilon} \right)^2 dt &= \int_0^\vartheta \left(\frac{X_t - x_t(\vartheta)}{\varepsilon} \right)^2 dt \\ &\quad + \int_\vartheta^{\hat{\vartheta}_\varepsilon} \left(\frac{X_t - x_t(\vartheta)}{\varepsilon} - \frac{\hat{\vartheta}_\varepsilon - \vartheta}{\varepsilon} \dot{x}_t(\vartheta) \right)^2 dt \\ &\quad + \int_{\hat{\vartheta}_\varepsilon}^T \left(\frac{X_t - x_t(\vartheta)}{\varepsilon} - \frac{\hat{\vartheta}_\varepsilon - \vartheta}{\varepsilon} \dot{x}_t(\vartheta) \right)^2 dt. \end{aligned} \quad (25)$$

and if $\vartheta > \hat{\vartheta}_\varepsilon$

$$\int_0^T \left(\frac{X_t - x_t(\hat{\vartheta}_\varepsilon)}{\varepsilon} \right)^2 dt = \int_0^{\hat{\vartheta}_\varepsilon} \left(\frac{X_t - x_t(\vartheta)}{\varepsilon} \right)^2 dt + \int_{\hat{\vartheta}_\varepsilon}^{\vartheta} \left(\frac{X_t - x_t(\vartheta)}{\varepsilon} \right)^2 dt \quad (26)$$

$$+ \int_{\vartheta}^T \left(\frac{X_t - x_t(\vartheta)}{\varepsilon} - \frac{\hat{\vartheta}_\varepsilon - \vartheta}{\varepsilon} \dot{x}_t(\vartheta) \right)^2 dt.$$

Using theorem 5.7 of Kutoyants[10] and for all ϑ we obtain

$$\int_0^T \left(\frac{X_t - x_t(\hat{\vartheta}_\varepsilon)}{\varepsilon} \right)^2 dt \longrightarrow \int_0^T \left(\frac{X_t - x_t(\vartheta)}{\varepsilon} \right)^2 dt \quad (27)$$

combined with (4.1) one finds

$$\int_0^T \left(\frac{X_t - x_t(\hat{\vartheta}_\varepsilon)}{\varepsilon} \right)^2 dt$$

$$\longrightarrow \int_0^{\vartheta} \left(\int_0^t \frac{g(x_t)}{g(x_s)} dW_s \right)^2 dt + \int_{\vartheta}^T \left(x_{\vartheta}^{(1)} + \int_{\vartheta}^t \frac{h(x_t)}{h(x_s)} dW_s \right)^2 dt$$

$$= \int_0^{\vartheta} \left(\int_0^t \frac{g(x_t)}{g(x_s)} dW_s \right)^2 dt + \int_{\vartheta}^T \left(\int_0^{\vartheta} \frac{g(x_{\vartheta})}{g(x_s)} dW_s + \int_{\vartheta}^t \frac{h(x_t)}{h(x_s)} dW_s \right)^2 dt$$

$$= \int_0^T \left[\int_0^t \left(\frac{g(x_t)}{g(x_s)} \mathbb{I}_{\{s < \vartheta\}} + \frac{h(x_t)}{h(x_s)} \mathbb{I}_{\{s \geq \vartheta\}} \right) dW_s \right]^2 dt$$

$$= \int_0^T W^2 \left[\int_0^t \left(\frac{g(x_t)}{g(x_s)} \mathbb{I}_{\{s < \vartheta\}} + \frac{h(x_t)}{h(x_s)} \mathbb{I}_{\{s \geq \vartheta\}} \right) ds \right]^2 dt = \int_0^T W^2[v(t)] dt.$$

where $v(t) = \int_0^t \left(\frac{g(x_t)}{g(x_s)} \mathbb{I}_{\{s < \vartheta\}} + \frac{h(x_t)}{h(x_s)} \mathbb{I}_{\{s \geq \vartheta\}} \right)^2 ds$, in connection with the **condition** \mathcal{B} , it is easy to see that $v(t)$ is an increasing function (see appendix), which guaranteed the existence of the Wiener process,

Therefore we can write

$$\int_0^T D(x_t, \hat{\vartheta}_\varepsilon) \left(\frac{X_t - x_t(\hat{\vartheta}_\varepsilon)}{\varepsilon} \right)^2 dt \longrightarrow \int_0^T D(x_t, \vartheta) W^2(v(t)) dt$$

$$= \int_0^{V_T(\vartheta)} W^2(v) dv = V_T(\vartheta)^2 \int_0^1 w_z^2 dz,$$

where $w_z = V_T(\vartheta)^{-1/2} W(V_T z)$, $0 \leq z \leq 1$ is a Wiener process.

Hence (as $\varepsilon \rightarrow 0$)

$$\Gamma_\varepsilon \implies \int_0^1 w_z^2 dz,$$

and the test is distribution free. Therefore for the first type errors we have

$$\alpha_\varepsilon(\Gamma_\varepsilon) = \mathbf{P}_{\mathcal{H}_0} \{ \Gamma_\varepsilon > c_\alpha \} \longrightarrow \mathbf{P} \left\{ \int_0^1 w_z^2 dz > c_\alpha \right\} = \alpha. \quad (28)$$

The theorem (4.1) is proved.

4.2 Alternative

Suppose now that $S \notin \mathcal{F}_\Theta$ with $\Theta = [a, b]$ but the function $S(\cdot)$, as before, has continuous bounded derivative. Then the solution of the corresponding equation :

$$dX_t = S(X_t)dt + \varepsilon dW_t, \quad X_0 = x_0, \quad 0 \leq t \leq T, \quad (29)$$

converges to another determinist function y_t , which is solution of the following equation:

$$\frac{dy_t}{dt} = S(y_t), \quad y_0 = x_0,$$

with the same consideration that in the alternative for the smooth case, the test $\Psi(X^\varepsilon) = \mathbb{I}_{\{\Gamma_\varepsilon > c_\alpha\}}$ is consistent and we have (17).

References

- [1] Anderson, T.W and Darling, D.A. (1952). *Asymptotic theory of certain goodness of fit criteria based on stochastic processes*, Ann. Math. Stat., Vol. 23, pp. 193-212.
- [2] Cramér, H. (1928). *On the composition of elementary errors*, II skand. Aktuarietds., vol.11 , pp. 141-180.
- [3] Cramér, H. (1946). *Mathematical Methods of statistics*, Princeton University Press .
- [4] Dachian, S. and Kutoyants, Yu.A. (2008). *On the goodness-of-fit tests for some continuous time processes*, in *Statistical Models and Methods for Biomedical and Technical Systems*, F.Vonta et al. (Eds), Birkhäuser, Boston, 385-423.
- [5] Darling, D.A. (1955). *The Cramér-Smirnov test in the parametric test*. Ann. Math. Statist. 42, 147-156.

- [6] Durbin, J. (1973). *Distribution Theory for Tests Based on the Sample Distribution Function*. Conference Board of the Mathematical sciences Regional conference series in Applied mathematics, No.9. Society for Industrial and Applied Mathematics, Philadelphia, Pa., .
- [7] Durbin, J. (1973). *Weak convergence of the sample distribution function when parameters are estimated*. Ann. Statist. 1, 279-290.
- [8] Freidlin, M.I., Wentzell, A.D. (1984). *Random Perturbations of Dynamical Systems*, Springer, New York.
- [9] Greenwood, P.E., Nikulin, M. (1996). *A Guide to Chi-Squared Testing*, John Wiley and Sons, New York.
- [10] Kutoyants, Yu. (1994). *Identification of Dynamical Systems with Small Noise*, Kluwer, Dordrecht.
- [11] Kutoyants, Yu.A. and Pilibossian, P. (1994). *On minimum L_1 -norm estimates of the parameter of Ornstein-Uhlenbeck process*, Statist. Probab. Lett., 20, 117-123.
- [12] Kutoyants, Y. A. (2004). *Statistical inference for ergodic diffusion processes*. Springer Series in Statistics. Springer-Verlag London Ltd., London.
- [13] Kutoyants, Y. (2008). *Goodness of fit tests for perturbed dynamical systems*. Universit du maine, Le Mans.
- [14] Lehmann, E.L. and Romano, J.P. (2005). *Testing Statistical Hypotheses*. (3rd ed.) Springer, New York.
- [15] Lipster, R.S. and Shiryaev, A.N. (2001). *Statistics of Random Process*. Applications, (2nd ed.) Springer, New York.
- [16] Smirnov, N. (1936). *Sur la distribution de w^2* , C. R. Acad. Sci. Paris, Vol. 202 , pp. 449-452.