On Cramér-von Mises test based on local time of switching diffusion process

Anis Gassem
Laboratoire de Statistique et Processus,
Université du Maine, 72085 Le Mans Cedex 9, France
e-mail: Anis.Gassem@univ-lemans.fr

Abstract
We consider a Cramér-von Mises test for hypothesis that the observed diffusion process has sign-type trend coefficient. It is shown that the limit distribution of the proposed test statistic is defined by the integral type functional of continuous Gaussian process. We provide the Karhunen-Loève expansion on [0,1] of the corresponding limiting process. The representation for the limit statistic allow us to find the threshold.

AMS Classification subjects: 62G10, 62G20, 62M02.
Key Words: Cramér-von Mises test, Gaussian process, asymptotic properties, Karhunen-Loève expansion.

1 Introduction
We consider the Cramér-von Mises (C-vM) test, when the basic model is an ergodic diffusion process, i.e., the observations $X^T = \{X_t, 0 \leq t \leq T\}$ are form the stochastic differential equation

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

with initial value $X_0$, Wiener process $\{W_t, t \geq 0\}$ and unknown to the observer trend coefficient $S(\cdot)$. Diffusion process of this type is widely used as model in many different fields such as biology, physics, economics and finance. Let us introduce the stochastic differential equation

$$dX_t = S_0(X_t) \, dt + dW_t, \quad X_0, \quad 0 \leq t \leq T,$$
where $S_0(x)$ is some known function. We have two hypotheses:

\[ H_0 : S(\cdot) = S_0(\cdot), \quad \text{against alternative} \quad H_1 : S(\cdot) \neq S_0(\cdot). \]

We observe the trajectory $X^T$ of (1) and we test the hypothesis $H_0$. Our goal is to study the C-vM test in this problem. The problem considered for this stochastic model is similar to the well-known in classical statistic C-vM test (see, e.g., Anderson and Darling [1, 2], Durbin and Knott [9], Durbin [10] and Stephens [24]). Indeed, in situation of i.i.d. observations $X^n = (X_1, \ldots, X_n)$ with c.d.f. $F(x)$ and the basic hypothesis $H_0$: $F(x) = F_0(x)$. The C-vM statistic is,

\[
W_n^2 = n \int_{-\infty}^{\infty} \left( \hat{F}_n(x) - F_0(x) \right)^2 dF_0(x), \quad \hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} 1\{X_i < x\},
\]

where $\hat{F}_n(x)$ is the empirical distribution function. Anderson-Darling generalize the previous test by adding a (nonnegative) weight function. Let us denote by $\beta(t)$ a Brownian bridge, i.e., a continuous Gaussian process with

\[
E\beta(t) = 0, \quad K(t, s) = E\beta(t)\beta(s) = t \wedge s - ts.
\]

Then the limit behavior of this statistic can be described with the help of this process as follows

\[
W_n^2 = n \int_{0}^{1} (\tilde{F}_n(t) - t)^2 dt = \int_{0}^{1} \beta^2_n(t) dt \implies \int_{0}^{1} \beta^2(t) dt.
\]

Hence the C-vM test $\psi_n(X^n) = 1_{\{W_n^2 > c_\alpha\}}$, with constant $c_\alpha$ defined by the equation

\[
P\left\{ \int_{0}^{1} \beta^2(t) dt > c_\alpha \right\} = \alpha.
\]

is of asymptotic size $\alpha$. The classical solution of this problem requires a preliminary computation of the sequence of eigenvalues $\lambda_1 > \lambda_2 > \ldots$ and eigenfunctions $f_1(t), f_2(t), \ldots$ of the Fredholm operator

\[
f(t) \rightarrow \tilde{f}(t) = \int_{0}^{1} K(t, s) f(s) ds.
\]

By Mercer’s theorem we have $K(t, s) = \sum_{k=1}^{\infty} \lambda_k f_k(s)f_k(t)$. Moreover, it is possible to expand $\beta(\cdot)$ into the Karhunen-Loève (KL) series (see, e.g., Kac and Siegert [15], Kac [14], Ash and Gardner [3])

\[
\beta(t) = \sum_{k=1}^{\infty} \sqrt{\lambda_k} \xi_k f_k(t), \quad W^2 = \sum_{k=1}^{\infty} \lambda_k \xi_k^2,
\]

where $S_0(x)$ is some known function. We have two hypotheses:

\[ H_0 : S(\cdot) = S_0(\cdot), \quad \text{against alternative} \quad H_1 : S(\cdot) \neq S_0(\cdot). \]

We observe the trajectory $X^T$ of (1) and we test the hypothesis $H_0$. Our goal is to study the C-vM test in this problem. The problem considered for this stochastic model is similar to the well-known in classical statistic C-vM test (see, e.g., Anderson and Darling [1, 2], Durbin and Knott [9], Durbin [10] and Stephens [24]). Indeed, in situation of i.i.d. observations $X^n = (X_1, \ldots, X_n)$ with c.d.f. $F(x)$ and the basic hypothesis $H_0$: $F(x) = F_0(x)$. The C-vM statistic is,

\[
W_n^2 = n \int_{-\infty}^{\infty} \left[ \hat{F}_n(x) - F_0(x) \right]^2 dF_0(x), \quad \hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} 1\{X_i < x\},
\]

where $\hat{F}_n(x)$ is the empirical distribution function. Anderson-Darling generalize the previous test by adding a (nonnegative) weight function. Let us denote by $\beta(t)$ a Brownian bridge, i.e., a continuous Gaussian process with

\[
E\beta(t) = 0, \quad K(t, s) = E\beta(t)\beta(s) = t \wedge s - ts.
\]

Then the limit behavior of this statistic can be described with the help of this process as follows

\[
W_n^2 = n \int_{0}^{1} (\tilde{F}_n(t) - t)^2 dt = \int_{0}^{1} \beta^2_n(t) dt \implies \int_{0}^{1} \beta^2(t) dt.
\]

Hence the C-vM test $\psi_n(X^n) = 1_{\{W_n^2 > c_\alpha\}}$, with constant $c_\alpha$ defined by the equation

\[
P\left\{ \int_{0}^{1} \beta^2(t) dt > c_\alpha \right\} = \alpha.
\]

is of asymptotic size $\alpha$. The classical solution of this problem requires a preliminary computation of the sequence of eigenvalues $\lambda_1 > \lambda_2 > \ldots$ and eigenfunctions $f_1(t), f_2(t), \ldots$ of the Fredholm operator

\[
f(t) \rightarrow \tilde{f}(t) = \int_{0}^{1} K(t, s) f(s) ds.
\]

By Mercer’s theorem we have $K(t, s) = \sum_{k=1}^{\infty} \lambda_k f_k(s)f_k(t)$. Moreover, it is possible to expand $\beta(\cdot)$ into the Karhunen-Loève (KL) series (see, e.g., Kac and Siegert [15], Kac [14], Ash and Gardner [3])

\[
\beta(t) = \sum_{k=1}^{\infty} \sqrt{\lambda_k} \xi_k f_k(t), \quad W^2 = \sum_{k=1}^{\infty} \lambda_k \xi_k^2,
\]

where $S_0(x)$ is some known function. We have two hypotheses:

\[ H_0 : S(\cdot) = S_0(\cdot), \quad \text{against alternative} \quad H_1 : S(\cdot) \neq S_0(\cdot). \]

We observe the trajectory $X^T$ of (1) and we test the hypothesis $H_0$. Our goal is to study the C-vM test in this problem. The problem considered for this stochastic model is similar to the well-known in classical statistic C-vM test (see, e.g., Anderson and Darling [1, 2], Durbin and Knott [9], Durbin [10] and Stephens [24]). Indeed, in situation of i.i.d. observations $X^n = (X_1, \ldots, X_n)$ with c.d.f. $F(x)$ and the basic hypothesis $H_0$: $F(x) = F_0(x)$. The C-vM statistic is,

\[
W_n^2 = n \int_{-\infty}^{\infty} \left[ \hat{F}_n(x) - F_0(x) \right]^2 dF_0(x), \quad \hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} 1\{X_i < x\},
\]

where $\hat{F}_n(x)$ is the empirical distribution function. Anderson-Darling generalize the previous test by adding a (nonnegative) weight function. Let us denote by $\beta(t)$ a Brownian bridge, i.e., a continuous Gaussian process with

\[
E\beta(t) = 0, \quad K(t, s) = E\beta(t)\beta(s) = t \wedge s - ts.
\]

Then the limit behavior of this statistic can be described with the help of this process as follows

\[
W_n^2 = n \int_{0}^{1} (\tilde{F}_n(t) - t)^2 dt = \int_{0}^{1} \beta^2_n(t) dt \implies \int_{0}^{1} \beta^2(t) dt.
\]

Hence the C-vM test $\psi_n(X^n) = 1_{\{W_n^2 > c_\alpha\}}$, with constant $c_\alpha$ defined by the equation

\[
P\left\{ \int_{0}^{1} \beta^2(t) dt > c_\alpha \right\} = \alpha.
\]

is of asymptotic size $\alpha$. The classical solution of this problem requires a preliminary computation of the sequence of eigenvalues $\lambda_1 > \lambda_2 > \ldots$ and eigenfunctions $f_1(t), f_2(t), \ldots$ of the Fredholm operator

\[
f(t) \rightarrow \tilde{f}(t) = \int_{0}^{1} K(t, s) f(s) ds.
\]

By Mercer’s theorem we have $K(t, s) = \sum_{k=1}^{\infty} \lambda_k f_k(s)f_k(t)$. Moreover, it is possible to expand $\beta(\cdot)$ into the Karhunen-Loève (KL) series (see, e.g., Kac and Siegert [15], Kac [14], Ash and Gardner [3])

\[
\beta(t) = \sum_{k=1}^{\infty} \sqrt{\lambda_k} \xi_k f_k(t), \quad W^2 = \sum_{k=1}^{\infty} \lambda_k \xi_k^2,
\]
where \( \{\xi_k : k \geq 1\} \) is a sequence of independent normal \( \mathcal{N}(0,1) \) random variables. The relations (2) are used to approximate the value \( c_\alpha \).

For certain Gaussian processes, such expansions were obtained with the help of special functions. In particular, the trigonometric functions (case of Brownian bridge), Anderson-Darling process expansion has expression in Legendre polynomial functions. Recently, Bessel functions are used by Deheuvels and Martynov [7], Deheuvels [6] and Gassem [12]. More about Cramér-von Mises theory and the use of KL expansion can be found in [10], [21], [8].

In this present work we are interested by the C-vM test based on the observation \( X^T=\{X_t, 0 \leq t \leq T\} \) solution of (1)

\[
W^2_T = T \int_{-\infty}^{\infty} \left[ f^*_T(x) - f_{S_0}(x) \right]^2 dF_{S_0}(x), \quad f^*_T(x) = \frac{\Lambda_T(x)}{T},
\]

where \( f^*_T(x) \) is the local time type estimator of the density (LTE) (see [18] Section 1.1.3). This test was proposed in [4], but it is not distribution free.

Despite the fact that this statistic converges in distribution (under hypothesis) to a functional of a Gaussian process [18], the choice of the threshold \( c_\alpha \) for the test \( \Phi_T(X^T) = 1_{\{W^2_T > c_\alpha\}} \) is not easy due to the structure of the covariance. To avoid such difficulty, a weighting of this statistic was introduced to make this test asymptotically distribution free (see kutoyants [19]). The purpose of the paper is to study the C-vM test for hypothesis that the observed process (1) is switching diffusion, i.e., the trend coefficient \( S_0(x) = -\text{sgn}(x) \) is discontinuous function and taking just two values \(+1\) and \( -1 \), the observed process has the form

\[
dx_t = -\text{sgn}(x) \, dt + dW_t, \quad X_0, \quad 0 \leq t \leq T, \tag{3}
\]

This work is a continuation of a study started by Gassem in [12], where the limit distribution of the test for this model was studied. Here we provide the KL expansion of the corresponding limiting process and we present the threshold \( c_\alpha \) of the test.

## 2 Model of observation

Suppose that the observed process is (1), where the trend coefficient \( S(x) \) satisfies the conditions of the existence and uniqueness of the solution and this solution has ergodic properties, i.e., there exists an invariant probability distribution \( F_S(x) \), and for any integrable w.r.t. this distribution function
The law of large numbers holds

\[ \frac{1}{T} \int_0^T h(X_t) \, dt \longrightarrow \int_{-\infty}^{\infty} h(x) \, dF_S(x). \]

These conditions can be found, in Durrett [11] and Kutoyants [18].

The invariant density function \( f_S(x) \) is defined by the equality

\[ f_S(x) = \frac{1}{G(S)} \exp \left\{ 2 \int_0^x S(v) \, dv \right\}, \quad x \in \mathbb{R}, \]

where \( G(S) \) is the normalizing constant. This density function can be estimated by the LTE \( f_T^T(x) = \Lambda_T(x)/T \). The local time \( \Lambda_T(x) \) admits the Tanaka-Meyer representation (see [22])

\[ \Lambda_T(x) = |X_T - x| - |X_0 - x| - \int_0^T \text{sgn}(X_t - x) \, dX_t, \]

and recently started to play an important role in statistical inference [18], [5], [16]. Note that this estimator is \( \sqrt{T} \) asymptotically normal as \( T \to \infty \) (see [18], Proposition 1.51)

\[ \sqrt{T} \left( f_T^T(x) - f_S(x) \right) \Rightarrow \mathcal{N}(0, R_{fS}(x, x)). \]

Moreover, this normed difference converge weakly to the limit Gaussian process (see [18], Theorem 4.13) with zero mean and covariance function

\[ R_{fS}(x, y) = 4f_S(x)f_S(y) \mathbb{E} \left( \frac{[1_{\{\xi>x\}} - F_S(\xi)][1_{\{\xi>y\}} - F_S(\xi)]}{f_S(\xi)^2} \right). \]

### 3 Choice of the threshold

#### 3.1 C-vM test

It is easy to see that (3) is an ergodic diffusion process with the stationary density \( f_{S_0}(x) = e^{-2|x|} \), for \( x \in \mathbb{R} \).

Suppose that we observed the trajectory \( X^T = \{X_t, 0 \leq t \leq T\} \) of (1) then to test the hypothesis \( \mathcal{H}_0 \) we can use the estimator LTE for construction of goodness-of-fit test based on the C-vM statistic

\[ W_T^2 = T \int_{-\infty}^{\infty} [f_T^T(x) - f_{S_0}(x)]^2 \, dF_{S_0}(x) = \int_0^1 \eta_T^2(s) \, ds, \]
The transformation $Y_t = F_{S_0}(X_t)$ simplifies the writing, because by the Itô formula the diffusion process $Y_t$ satisfies the differential equation

$$Y_t = f_{S_0}(X_t) [2S_0(X_t)dt + dW_t], \quad Y_0 = F_{S_0}(X_0)$$

with reflecting bounds in 0 and 1 and under hypothesis $H_0$ has uniform on $[0, 1]$ invariant distribution.

Fix a number $\alpha \in (0, 1)$ and define the class $K_\alpha$ of tests of asymptotic level $1 - \alpha$ as follows:

$$K_\alpha = \{ \phi_T : \lim_{T \to \infty} \mathbb{E}_{S_0} \phi_T(X^T) \leq \alpha \}.$$

Let us denote by $c_\alpha$ the value defined by the equation

$$\mathbb{P} \left\{ \int_0^1 \eta^2_f(s) ds > c_\alpha \right\} = \mathbb{P} (W^2 > c_\alpha) = \alpha,$$  \hspace{1cm} (4)

We have the following

**Proposition 3.1.** The C-vM test $\Phi_T(X^T) = 1_{\{W^2_T > c_\alpha\}}$, belongs to $K_\alpha$.

**Proof.** The main idea of the proof is to show (under $H_0$) by using the method of Ibragimov-Khasminskii (Theorem A22, in [13]) that the distribution of $W^2_T$ converge to the distribution of $W^2$. The detailed proof is given in [12].

### 3.2 KL expansion of Gaussian process $\eta_f(t)$

In this section, we establish the KL expansion of the Gaussian process $\eta_f(t)$, for $0 \leq t \leq 1$, the representation for the limit statistic defined by the equation (4) allow to find the threshold $c_\alpha$.

Under hypothesis $H_0$ the process $\eta_f(t)$ admits the following representation

$$\eta_f(t) = 2 (1 - |2t - 1|) \int_0^1 \frac{1_{\{s > t\}} - s}{1 - |2s - 1|} dW_s,$$

The last integral is with respect to double-sided Wiener process, i.e., $W_s = W^+(s)$, $s \geq 1/2$ and $W_s = W^-(s)$, $s \leq 1/2$, where $W^+(\cdot)$ and $W^-(\cdot)$ are two independent Wiener processes.

By the law of large numbers we have

$$\sqrt{T} \left( f^2_T(x) - f_{S_0}(x) \right) = 2f_{S_0}(x) W \left( \int_0^T \left( \frac{1_{\{X_t > x\}} - F_{S_0}(X_t)}{f_{S_0}(X_t)} \right)^2 dt \right)$$

$$\rightarrow 2f_{S_0}(x) W \left( \int_{-\infty}^{\infty} \left( \frac{1_{\{y > x\}} - F_{S_0}(y)}{f_{S_0}(y)} \right)^2 dF_{S_0}(y) \right).$$

5
Proposition 3.2. The Gaussian process \( \eta_f(t) \) has a KL expansion given by

\[
\eta_f(t) = \sum_{n=1}^{\infty} \frac{\xi_n}{\nu_n} \sqrt{2} \text{sgn}(1/2 - t) \sin n\pi(1 - |2t - 1|)
\]

\[
+ \sum_{n=1}^{\infty} \frac{\xi_n}{\nu_n} \left\{ \frac{\sqrt{2}}{\text{Si}(\nu_n)} \left[ \left( \frac{\alpha(\nu_n)}{\nu_n} - \bar{\alpha}(\nu_n) \right) \sin(\nu_n(1 - |2t - 1|)) 
\right.
\]

\[
- \left( \frac{\sin(\nu_n)}{\nu_n} - \cos(\nu_n) \right) \alpha(\nu_n(1 - |2t - 1|)) \right\}, \quad 0 \leq t \leq 1,
\]

where \( W(u), u \geq 0 \) is some Wiener process. Making the transformation \( t = F_{S_0}(x) \) we obtain

\[
\eta_f(t) = 2(1 - |2t - 1|) W \left( \int_0^1 \frac{1_{(s \geq t)} - s}{1 - |2s - 1|} \, ds \right)
\]

\[
= 2 (1 - |2t - 1|) \int_0^1 \frac{1_{(s \geq t)} - s}{1 - |2s - 1|} \, dW_s = \eta_f(t)
\]

The covariance function of \( \eta_f(\cdot) \), for \( 0 \leq s, t \leq 1 \), can be written as follows:

\[
R_f(s, t) = 4(1 - |2t - 1|)(1 - |2s - 1|) \int_0^1 \frac{(1_{(s \geq t)} - u)(1_{(u \geq s)} - u)}{(1 - |2u - 1|)^2} \, du
\]

\[
= 4(1 - |2t - 1|)(1 - |2s - 1|) \left\{ \int_0^{s \wedge t} \frac{u(1-u)}{(1 - |2u - 1|)^2} \, du + \int_{s \vee t}^1 \frac{(1-u)^2}{(1 - |2u - 1|)^2} \, du \right\}.
\]

Let \( s \leq t \), then by a direct calculation we have

\[
R_f(s, t) = \left\{ \begin{array}{cl}
4st \ln(4st) + 4s(1-t), & s, t \leq 1/2, \\
4s(1-t) \ln(4s(1-t)) + 4s(1-t), & s \leq 1/2, t > 1/2, \\
4(1-s)(1-t) \ln((4s)(1-t)) + 4s(1-t), & s, t > 1/2.
\end{array} \right.
\]

Therefore, we can write

\[
R_f(s, t) = (1 - |2s - 1|)(1 - |2t - 1|) \ln[(1 - |2s - 1|)(1 - |2t - 1|)]
\]

\[
- 4(1_{(s \geq t)} - s)(1_{(u > s)} - t), \quad 0 \leq s, t \leq 1.
\]  

(5)

Of course,

\[
R_f(t, t) = 2(1 - |2t - 1|)^2 \ln(1 - |2t - 1|) + 4t(1-t), \quad 0 \leq t \leq 1.
\]  

(6)

By the Karhunen-Loève Theorem, we have the following

Proposition 3.2. The Gaussian process \( \eta_f(t) \) has a KL expansion given by
where \( \{\xi_n, n \geq 1\}, \{\xi_n^*, n \geq 1\} \) denote two independent sequences of independent and identically distributed \( \mathcal{N}(0, 1) \) random variables and \( \nu_n, n = 1, 2, \ldots \), are the positive zeros of \( f(\cdot) \), defined by the equation

\[
f(t) = Si(t) [\alpha(t) - t \dot{\alpha}(t)] - G(t) [\sin(t) - t \cos(t)],
\]

where

\[
\alpha(t) = Ci(t) \sin(t) - Si(t) \cos(t), \quad \dot{\alpha}(t) = \frac{d}{dt} \alpha(t), \quad G(t) = \int_0^t \frac{\alpha(s)}{s} ds,
\]

with

\[
Ci(t) = \gamma + \ln(t) + \int_0^t \frac{\cos(s)}{s} ds, \quad Si(t) = \int_0^t \frac{\sin(s)}{s} ds,
\]

the cosine and sine integral respectively, \( \gamma = 0.577215... \) the Euler’s constant.

**Proof.** To determine the eigenfunctions, one must solve the integral equation:

\[
\lambda_n \psi_n(t) = \int_0^1 R_f(t, s) \psi_n(s) ds, \quad t \in [0, 1]. \tag{8}
\]

The fact that we have an absolute value in the covariance function \( R_f(s, t) \), we will take \( \psi_n(t) = \psi_{1,n}(t)\mathbb{1}_{\{t<1/2\}} + \psi_{2,n}(t)\mathbb{1}_{\{t\geq1/2\}} \).

Here, the \( \lambda_n, n = 1 \ldots \infty \) are eigenvalues yet to be determined. We have the boundary conditions

\[
\psi_{1,n}(0) = \psi_{2,n}(1) = 0, \quad \psi_{1,n}(1/2) = \psi_{2,n}(1/2), \tag{9}
\]

\[
\lambda_n \psi_{n}(1/2) = \int_0^1 (1 - |2s - 1|) [1 + \ln(1 - |2s - 1|)] \psi_n(s) ds. \tag{10}
\]

Taking derivative of equation (8) with respect to \( t \), we obtain

\[
\lambda_n \psi_n'(t) = \int_0^1 \frac{\partial R_f(t, s)}{\partial t} \psi_n(s) ds, \quad t \in [0, 1].
\]

Evaluating the derivative of the last equation we obtain

\[
\lambda_n \psi_n''(t) + 4 \psi_n(t) = \frac{4C_n}{1 - |2t - 1|}, \quad t \in [0, 1]. \tag{11}
\]
where
\[ C_n = \int_0^1 (1 - |2s - 1|) \psi_n(s) \, ds. \] (12)

Let \( \nu_n = 1/\sqrt{\lambda_n} \), the last equation can be written as follow
\[
\begin{align*}
\psi''_{1,n}(t) + 4 \nu_n^2 \psi_{1,n}(t) &= \frac{2 \nu_n^2 C_n}{t}, & t < 1/2, \\
\psi''_{2,n}(t) + 4 \nu_n^2 \psi_{2,n}(t) &= \frac{2 \nu_n^2 C_n}{1 - t}, & t \geq 1/2.
\end{align*}
\]

When \( C_n = 0 \) the equation (11) has solution
\[
\psi_n(t) = \begin{cases} 
\psi_{1,n}(t) = A_{1,n} \sin(2\nu_n t) + B_{1,n} \cos(2\nu_n t), & t < 1/2, \\
\psi_{2,n}(t) = A_{2,n} \sin(2\nu_n t) + B_{2,n} \cos(2\nu_n t), & t \geq 1/2.
\end{cases}
\]

Now, using the Lagrange method we have
\[
\begin{align*}
\dot{A}_{1,n} \sin(2\nu_n t) + \dot{B}_{1,n} \cos(2\nu_n t) &= 0, \\
\dot{A}_{1,n} \cos(2\nu_n t) - \dot{B}_{1,n} \sin(2\nu_n t) &= \nu_n C_n \frac{\cos(2\nu_n t)}{t}, \\
\dot{A}_{2,n} \sin(2\nu_n t) + \dot{B}_{2,n} \cos(2\nu_n t) &= 0, \\
\dot{A}_{2,n} \cos(2\nu_n t) - \dot{B}_{2,n} \sin(2\nu_n t) &= \nu_n C_n \frac{\sin(2\nu_n t)}{1 - t}.
\end{align*}
\]

Then, the first equation has a solution of the following form
\[
\psi_{1,n}(t) = A_{1,n} \sin(2\nu_n t) + B_{1,n} \cos(2\nu_n t) + \nu_n C_n \alpha(2\nu_n t), & t < 1/2,
\]
where
\[
\alpha(t) = \text{Ci}(t) \sin(t) - \text{Si}(t) \cos(t)
\]
with
\[
\text{Ci}(t) = \gamma + \ln(t) + \int_0^t \frac{\cos(s)}{s} \, ds, \quad \text{Si}(t) = \int_0^t \frac{\sin(s)}{s} \, ds.
\]

In the same manner for the second equation, we have
\[
\begin{align*}
\dot{A}_{2,n} \sin(2\nu_n t) + \dot{B}_{2,n} \cos(2\nu_n t) &= 0, \\
\dot{A}_{2,n} \cos(2\nu_n t) - \dot{B}_{2,n} \sin(2\nu_n t) &= \nu_n C_n \frac{\cos(2\nu_n t)}{1 - t}, \\
\dot{A}_{2,n} \sin(2\nu_n t) + \dot{B}_{2,n} \cos(2\nu_n t) &= 0, \\
\dot{A}_{2,n} \cos(2\nu_n t) - \dot{B}_{2,n} \sin(2\nu_n t) &= \nu_n C_n \frac{\sin(2\nu_n t)}{1 - t}.
\end{align*}
\]

Using the following equalitie
\[
\begin{align*}
\cos(2\nu_n t) &= \cos(2\nu_n (1 - t)) \cos(2\nu_n) + \sin(2\nu_n (1 - t)) \sin(2\nu_n), \\
\sin(2\nu_n t) &= \cos(2\nu_n (1 - t)) \sin(2\nu_n) - \sin(2\nu_n (1 - t)) \cos(2\nu_n).
\end{align*}
\]
Then, the second equation has a solution of the following form

\[ \psi_{1,n}(t) = A_{2,n} \sin(2\nu_n t) + B_{2,n} \cos(2\nu_n t) + \nu_n C_n \alpha(2\nu_n (1 - t)), \quad t \geq 1/2. \]

Therefore, using the boundary conditions (9), the general solution is of the following form

\[ \psi_n(t) = (A_{1,n} 1_{t \leq 1/2} - A_{2,n} 1_{t \geq 1/2}) \sin(\nu_n (1 - |2t - 1|)) + \nu_n C_n \alpha(\nu_n (1 - |2t - 1|)), \quad t \in [0, 1]. \]

Now, returning to the equation (10), replacing \( \psi_n(s) \) by this general form and making the change of variable \( u = \nu_n s \), we have

\[ \psi_n(1/2) = \frac{\nu_n^2}{2} \int_0^1 s (1 + \ln(s)) [(A_{1,n} - A_{2,n}) \sin(\nu_n s) + 2 C_n \nu_n \alpha(\nu_n s)] ds \]

\[ = \frac{1}{2} (A_{1,n} - A_{2,n}) \int_0^{\nu_n} u [1 - \ln(\nu_n) + \ln(u)] \sin(u) du \]

\[ + C_n \nu_n \int_0^{\nu_n} u [1 - \ln(\nu_n) + \ln(u)] \alpha(u) du. \]

By a simple integration by part we have

\[ \int_0^{\nu_n} u [1 - \ln(\nu_n) + \ln(u)] \sin(u) du = 2 \sin(\nu_n) - \nu_n \cos(\nu_n) - \text{Si}(\nu_n). \]

\[ \int_0^{\nu_n} u [1 - \ln(\nu_n) + \ln(u)] \alpha(u) du = 2 \alpha(\nu_n) - \nu_n \dot{\alpha}(\nu_n) - G(\nu_n). \]

where

\[ \dot{\alpha}(t) = \frac{d}{dt} \alpha(t) = \text{Si}(t) \sin(t) + \text{Ci}(t) \cos(t), \quad G(t) = \int_0^t \frac{\alpha(s)}{s} ds. \]

Finally we have

\[ (A_{1,n} - A_{2,n}) [\sin(\nu_n) - \nu_n \cos(\nu_n) - \text{Si}(\nu_n)] \]

\[ + 2 C_n \nu_n [\alpha(\nu_n) - \nu_n \dot{\alpha}(\nu_n) - G(\nu_n)] = 0. \]

(14)

In the same manner for equality (12), we obtain

\[ C_n = \frac{1}{2 \nu_n^2} \int_0^{\nu_n} u [(A_{1,n} - A_{2,n}) \sin(u) + 2 C_n \nu_n \alpha(u)] du \]

\[ = \frac{1}{2 \nu_n^2} (A_{1,n} - A_{2,n}) [\sin(\nu_n) - \nu_n \cos(\nu_n)] \]

\[ + \frac{1}{\nu_n} C_n \nu_n [\alpha(\nu_n) - \nu_n \dot{\alpha}(\nu_n) + \nu_n], \]
and we have
\[
(A_{1,n} - A_{2,n}) (\sin(\nu_n) - \nu_n \cos(\nu_n)) + 2 C_n \nu_n (\alpha(\nu_n) - \nu_n \dot{\alpha}(\nu_n)) = 0. \tag{15}
\]
Using the boundary conditions (9) and equalities (14), (15), we get the linear system
\[
\begin{bmatrix}
\sin(\nu_n) & \sin(\nu_n) & 0 \\
\text{Si}(\nu_n) & -\text{Si}(\nu_n) & 2 \nu_n G(\nu_n) \\
\sin(\nu_n) - \nu_n \cos(\nu_n) & - (\sin(\nu_n) - \nu_n \cos(\nu_n)) & 2 \nu_n (\alpha(\nu_n) - \nu_n \dot{\alpha}(\nu_n))
\end{bmatrix}
\begin{bmatrix}
A_{1,n} \\
A_{2,n} \\
C_n
\end{bmatrix}
= \begin{bmatrix}
0 \\
0 \\
0
\end{bmatrix},
\tag{16}
\]
which has non-zero solutions in \((A_{1,n}, A_{2,n}, C_n)\), if and only if
\[
\det
\begin{bmatrix}
\sin(\nu_n) & \sin(\nu_n) & 0 \\
\text{Si}(\nu_n) & -\text{Si}(\nu_n) & 2 \nu_n G(\nu_n) \\
\sin(\nu_n) - \nu_n \cos(\nu_n) & - (\sin(\nu_n) - \nu_n \cos(\nu_n)) & 2 \nu_n (\alpha(\nu_n) - \nu_n \dot{\alpha}(\nu_n))
\end{bmatrix}
= 4 \nu_n \sin(\nu_n) f(\nu_n) = 0,
\tag{17}
\]
where
\[
f(\nu_n) = \text{Si}(\nu_n) (\alpha(\nu_n) - \nu_n \dot{\alpha}(\nu_n)) - G(\nu_n) (\sin(\nu_n) - \nu_n \cos(\nu_n)). \tag{18}
\]
Next we consider the function \(\psi_n(t)\) of the form (13) and \(A_{1,n}, A_{2,n}, C_n\) fulfilling (16), and by \(\nu_{1,n}, \nu_{2,n}, n = 1, 2, \ldots\), respectively solutions of \(\sin(\nu_n) = 0\) and \(f(\nu_n) = 0\). We have the following two cases:

**Case 1:** When \(\nu_n = \nu_{1,n} = n \pi\), for some \(n = 1, 2, \ldots\), \(\sin(\nu_n) = 0\), so that (16) implies that \(A_{1,n} = A_{2,n}\) and \(C_n = 0\). The requirement that \(\int_0^1 \psi_n^2(t) dt = 1\) yields, therefore
\[
\psi_n(t) = \sqrt{2} \text{sgn}(1/2 - t) \sin(n \pi(1 - |2t - 1|)), \quad t \in [0, 1].
\]

**Case 2:** When \(\nu_n = \nu_{2,n}\), for some \(n = 1, 2, \ldots, f(\nu_n) = 0\), so that (16) implies that \(A_{1,n} = -C_n (\alpha(\nu_{2,n}) - \nu_{2,n} \dot{\alpha}(\nu_{2,n})) / (\sin(\nu_{2,n}) - \nu_{2,n} \cos(\nu_{2,n})) = -A_{2,n}\). The requirement that \(\int_0^1 \psi_n^2(t) dt = 1\) yields, therefore
\[
\psi_n(t) = \frac{\sqrt{2}}{\text{Si}(\nu_{2,n})} \left[ \left( \frac{\alpha(\nu_{2,n})}{\nu_{2,n}} - \dot{\alpha}(\nu_{2,n}) \right) \sin(\nu_{2,n}(1 - |2t - 1|)) 
- \left( \frac{\sin(\nu_{2,n})}{\nu_{2,n}} - \cos(\nu_{2,n}) \right) \alpha(\nu_{2,n}(1 - |2t - 1|)) \right], \quad t \in [0, 1].
\]
and this ends the proof.
3.3 Simulation

As a direct consequence of (7), the distribution of \( \int_0^1 \eta^2_f(t) \, dt \), coincide with the distribution of the quadratic form (under hypothesis \( \mathcal{H}_0 \)), we obtain the identity

\[
\int_0^1 \eta^2_f(t) \, dt = \sum_{n=1}^{\infty} \frac{\xi^2_n}{n^2 \pi^2} + \sum_{n=1}^{\infty} \frac{\xi^2_n}{\nu^2_n},
\]

(19)

allowing to evaluate the distribution of \( \int_0^1 \eta^2_f(t) \, dt \) by a number of methods (see, e.g., Smirnov [23], Martynov [20], and the references therein).

We may check from the formulas (7) and (19) that

\[
\int_0^1 R_f(t, t) \, dt = E \int_0^1 \eta^2_f(t) \, dt = \sum_{n=1}^{\infty} \frac{1}{n^2 \pi^2} + \sum_{n=1}^{\infty} \frac{1}{\nu^2_n} = \frac{1}{6} + \frac{5}{18} = \frac{4}{9},
\]

(20)

where for the first sum we have used the Euler’s formula \( \sum_{n=1}^{\infty} 1/n^2 = \pi^2/6 \) and numerically (Secant Method) for the second sum.

To check (20), we infer from (6) that

\[
\int_0^1 R_f(t, t) \, dt = 2 \int_0^1 [2t(1-t) + t^2 \ln(t)] \, dt = 2 \left( \frac{1}{3} - \frac{1}{9} \right) = \frac{4}{9},
\]

(21)

We see therefore that (20),(21) are in agreement.

The random variable \( \int_0^1 \eta^2_f(t) \, dt \) is a weighted sum of independent \( \chi_1^2 \) components, Thus, calculating the \( 1 - \alpha \) quantiles of this random variable we obtaining the threshold \( c_\alpha \) of C-vM test defined in equation (4).

![Threshold choice of \( \int_0^1 \eta^2_f(t) \, dt \)](image)

Particularly, if \( \alpha = 0.05 \) we obtain \( c_{0.05} = 2.4202 \).
Acknowledgments

The author is very grateful to Professor Y. Kutoyants for suggesting this problem, and fruitful discussion and Professor M. Kleptsyna for important help.

References


