

Cramér-von Mises Gaussianity test in Hilbert space

Gennady MARTYNOV

Institute for Information Transmission Problems
of the Russian Academy of Sciences
Higher School of Economics, Russia, Moscow

Statistique Asymptotique des Processus Stochastiques-X

Le Mans, Mars 2015

CONTENTS

One-dimensional CvM test:

$$\omega_n^2 = n \int_0^1 \psi^2(t) (F_n(t) - t)^2 dt$$



Multidimensional CvM test:

$$\omega_n^2 = n \int_{[0, 1]^d} \psi^2(t) (F_n(t) - t)^2 dt$$



CvM test that the process is Gaussian:

$$\omega_n^{*2} = n \int_{[0, 1]^\infty} \psi^2(t) (F_n^*(t) - F^*(t))^2 dt$$

WEIGHTED CRAMÉR-VON MISES STATISTICS

One-dimensional weighted Cramér-von Mises statistic is

$$\omega_n^2 = n \int_0^1 \psi^2(t)(F_n(t) - t)^2 dt, \quad (1)$$

where $F_n(t)$ is the empirical distribution function based on the sample X_1, X_2, \dots, X_n from the uniform distribution on $[0, 1]$, and $\psi(t)$ is a weight function. The statistic (1) designed to test the hypothesis

$$H_0 : F(x) = t$$

against the alternative

$$H_1 : F(x) \neq t,$$

where $F(x)$ is continuous distribution function.

.

If the condition

$$\int_0^1 \psi^2(t) t(1-t) dt < \infty$$

is fulfilled then the statistic ω_n^2 converges in probability to

$$\omega^2 = \int_0^1 \xi^2(t) dt, \quad (2)$$

where $\xi(t)$, $t \in [0, 1]$, is the Gaussian process with zero mean and the covariance function

$$K_\psi(t, \tau) = \psi(t)\psi(\tau)(\min(t, \tau) - t\tau)$$

(see for example Van der Vaart and Wellner [1996, p. 50]).

The Gauss process $\xi(t)$ can be developed in the Karhunen-Loève series

$$\xi(t) = \sum_{i=1}^{\infty} \frac{x_k \varphi_k(t)}{\sqrt{\lambda_k}},$$

where $x_k \sim N(0, 1)$, $k = 1, 2, \dots$, are independent random variables, and λ_k and $\varphi_k(t)$, $i = 1, 2, \dots$, are the eigenvalues and eigenfunctions of the Fredholm integral equation

$$\varphi(t) = \lambda \int_0^1 \psi(t)\psi(\tau)(\min(t, \tau) - t\tau)\varphi(\tau)d\tau. \quad (3)$$

By twice differentiation (3) with respect to t , we obtain the differential equation

$$h''(t) + \lambda\psi^2(t)h(t) = 0$$

with the conditions $h(0) = h(1) = 0$. Here $h(t) = \varphi(t)/\psi(t)$.

.

CRAMÉR-VON MISES STATISTIC WITH THE POWER WEIGHT FUNCTION

$$\psi(t) = t^\beta$$

Deheuvels and Martynov (2003) described the follows result. Let $\{B(t) : 0 \leq t \leq 1\}$ be the **Brownian bridge**. Then, for each $\beta = \frac{1}{2\nu} - 1 > -1$, the **Karhunen-Loeve expansions** of $\{t^\beta B(t) : 0 < t \leq 1\}$ is given by

$$t^\beta B(t) = \sum_{k=1}^{\infty} \sqrt{\lambda_{kB}} \omega_k e_{kB}(t).$$

Here, $\{\omega_k : k \geq 1\}$ are **i.i.d. $N(0, 1)$ random variables**, and, for $k = 1, 2, \dots$, the **eigenvalues** are $\lambda_k = (2\nu/z_{\nu,k})^2$, corresponding **eigenfunctions** are

$$e_{B,k}(t) = \frac{t^{\frac{1}{2\nu}-\frac{1}{2}} J_\nu(z_{\nu,k} t^{\frac{1}{2\nu}})}{\sqrt{\nu} J_{\nu-1}(z_{\nu,k})}, \quad 0 < t \leq 1,$$

and $z_{\nu,k}$, $k = 1, 2, \dots$, are **zeros of the Bessel functions $J_\nu(z)$** .

.

Let $\{W(t) : 0 \leq t \leq 1\}$ be the Wiener process. Then, for each $\beta = \frac{1}{2\nu} - 1 > -1$, the Karhunen-Loeve expansions of $\{t^\beta W(t) : 0 < t \leq 1\}$ is

$$t^\beta W(t) = \sum_{k=1}^{\infty} \sqrt{\lambda_{kW}} \omega_k e_{kW}(t).$$

Here, the eigenvalues are $\left(\frac{2\nu}{z_{\nu-1,i}}\right)^2$, $k = 1, 2, \dots$, and corresponding eigenfunctions are

$$e_{W,i}^*(t) = \frac{\frac{1}{t^{2\nu}} - \frac{1}{2}}{\nu J_\nu^2(z_{\nu-1,i})} J_\nu\left(z_{\nu-1,i} t^{1/2\nu}\right),$$

Deheuvels, P., Martynov, G. (2003) Karhunen-Loeve expansions for weighted Wiener processes and Brownian bridges via Bessel functions. *Progress in Probability*. 55, 57-93. Birkhäuser Verlag, Basel/Switzerland.

UNIFORMITY TEST ON $[0, 1]^d$

H_0 : \mathbf{U} have the uniform distribution on $[0, 1]^d$. It can be used the statistic

$$\omega_n^2 = n \int_{\mathbf{t} \in [0, 1]^d} (F_n(\mathbf{t}) - \mathbf{t})^2 d\mathbf{t}.$$

$$F_n(\mathbf{t}) = \frac{1}{n} \sum_{i=1}^n I_{\mathbf{t} < \mathbf{U}_i}$$

The empirical process $\xi_n(\mathbf{t}) = n^{1/2}(F_n(\mathbf{t}) - \mathbf{t})$ have the covariance function

$$K(\mathbf{s}, \mathbf{t}) = \prod_{i=1}^d \min(s_i, t_i) - \prod_{i=1}^d s_i t_i,$$

where $\mathbf{s} = (s_1, s_2, \dots, s_d)$, $\mathbf{t} = (t_1, t_2, \dots, t_d)$.

At first we will consider the covariance function

$$K_0(\mathbf{s}, \mathbf{t}) = \prod_{i=1}^d \min(s_i, t_i).$$

The corresponding covariance operator has the eigenfunctions

$$\phi_{ij\dots k}(\mathbf{t}) = 2^{m/2} \sin \pi(i - 1/2)t_1 \cdot \dots \cdot \sin \pi(k - 1/2)t_m$$

and the corresponding eigenvalues

$$\lambda_{ij\dots k} = [\pi^{2m}(i - 1/2)^2 \cdot \dots \cdot (k - 1/2)^2]^{-1}.$$

There $ij\dots k$ are all permutations of $1, 2, \dots, m$.

.

The eigenvalues, corresponding to $K(\mathbf{s}, \mathbf{t})$, are

1) numbers $2^{2m} / \pi^{2m} (2N + 1)$, where $N > 0$ with the multiplicities q_n such, that $q_n + 1$ is equal to numbers of ways of writing $2N + 1$ as a product of m positive integers.

2) simple eigenvalues, computing as the roots of the equation

$$\sum_{k=1}^{\infty} \frac{(q_k + 1) \lambda_k^2}{\lambda_k + \mu} = \frac{1}{2^m}$$

It was calculated the tables of considered statistic for $d = 2$ and $d = 3$.

Krivyakova, E.N., Martynov, G.V. and Tyurin Yu. N. (1977) on the distribution omega-square statistics in the multidimensional case, *Theory Probab.Appl.* 22 406-410

**CRAMÉR–VON MISES GAUSSIANTY TEST
FOR THE PROCESS IN HILBERT SPACE**

Martynov, G. V., (1979) Martynov, G. V., (1992)

Hypothesis to be tested

We will consider here the problems of the hypothesis testing that an observed random process $x(t)$ on the interval $[0, 1]$ is Gaussian. The hypothesis is

$H_{0x} :$ $x(t), t \in [0, 1],$ **is the Gaussian process**
with $Ex(t) = 0$ and $E(x(t)x(\tau)) = K_x(t, \tau), t, \tau \in [0, 1].$

The alternative H_{1x} to H_{0x} is the set of all another Gaussian processes on the interval $[0, 1]$. Hypothesis testing is performed using n observations $x_1(t), x_2(t), \dots, x_n(t)$ of the process $x(t)$.

Realizations of the process $x(t)$ belongs with probability 1 to the separable Hilbert space $\mathcal{H} = L^2([0, 1])$. As a basis for \mathcal{H} we choose the orthonormal basis formed by eigenfunctions $g_1(t), g_2(t), \dots$ of the integral equation

$$g(t) = \lambda \int_0^1 K_x(t, \tau) g(\tau) d\tau. \quad (4)$$

Denote by $\lambda_1, \lambda_2, \lambda_3, \dots$ the eigenvalues of the equation (4). The process $x(t)$ can be represented under hypothesis in the form of Karhunen-Loeve expansion

$$x(t) = \sum_{j=1}^{\infty} \frac{z_j g_j(t)}{\lambda_j}, \quad (5)$$

where $z_j, j = 1, 2, \dots$, are independent standard normal variables.

.

Let $h = (h^1, h^2, h^3, \dots)$ denote the coordinates of $x(t)$ in \mathcal{H} . Here

$$h^j = \frac{z_j}{\lambda_j} = \int_0^1 x(t)g_j(t)dt, \quad j = 1, 2, \dots \quad (6)$$

Analogously, the observations $x_i(t)$ of the random process $x(t)$ can be represented in \mathcal{H} as

$$h_i = (h^{i1}, h^{i2}, h^{i3}, \dots), \quad i = 1, 2, \dots, n,$$

where

$$h^{ij} = \int_0^1 x_i(t)g_j(t)dt, \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots \quad (7)$$

.

General formulation of the Gaussianity hypothesis

We will consider the probability space $(\mathcal{X}, \mathcal{B}, \nu)$, where \mathcal{X} is a separable Hilbert space of elementary events,

\mathcal{B} is the σ -algebra of Borel set on \mathcal{X} ,

ν is a probability measure.

Let we have n observations X^1, X^2, \dots, X^n of the random element X of $(\mathcal{X}, \mathcal{B}, \nu)$.

Let μ be a Gaussian measure on $(\mathcal{X}, \mathcal{B})$ defined by the mathematical expectation 0 and a covariance operator K with the kernel $K(z, w)$, $z, w \in \mathcal{X}$. The covariance function $K(z, w)$ supposed be known.

Let X be Gaussian random element with a characteristic function

$$Ee^{i(X,t)} = e^{-(Kt,t)/2}, \quad t \in \mathcal{X}. \quad (8)$$

We will test the hypothesis

$$H_{0\mu} : \nu = \mu$$

versus alternative

$$H_{1\mu} : \nu \text{ is a measure different from } \mu.$$

.

Let $e = (e_1, e_2, \dots)$ be the orthonormal basis of the eigenvectors of K and $\sigma_1^2, \sigma_2^2, \dots$ be the eigenvalues of K .

Let $x = (x_1, x_2, \dots)$ be the representation of x in the basis e .

Random element $X = (X_1, X_2, \dots)$ has the independent components with the distributions $N(0, \sigma_j^2)$, $j = 1, 2, \dots$

We transform the probability space $(\mathcal{X}, \mathcal{B}, \nu)$ to a probability space

$$([0, 1]^\infty, \mathcal{C}^\infty, \Gamma).$$

Here \mathcal{C} is the Borel set on $[0, 1]$ and Γ is the measure corresponding to ν .

We denote transformed random element X as $T = (T_1, T_2, \dots)$. Here $T_i = \Phi(X_j; 0, \sigma^2)$, $j = 1, 2, \dots$, $\Phi(\cdot; \mu, \sigma^2)$ is the normal distribution function with the mean μ and the variance σ^2 .

The observations X^i of X are transformed to $T^i = (T^{i1}, T^{i2}, T^{i3}, \dots)$, $i = 1, 2, \dots$, $T^{ij} = \Phi(X^{ij}; 0, \sigma^2)$, $i = 1, 2, \dots, n$, $j = 1, 2, \dots$. All T^i belong to $[0, 1]^\infty$.

Let Υ be the "uniform" measure on $([0, 1]^\infty, C^\infty)$. Now we will test the hypothesis

$$H_{0\Upsilon} : \Gamma = \Upsilon \tag{9}$$

versus alternative

$$H_{1\Upsilon} : \Gamma \neq \Upsilon. \tag{10}$$

.

EXAMPLE

We will test the hypothesis H_0 that the observed random process on $[0,1]$ is Gaussian with the zero mean and the covariance function

$$K_0(t, \tau) = \min(t, \tau).$$

This process (and all its observations) can be multiplied on

$$\frac{1}{\sqrt{t}}.$$

Resulting process has the unit variance. Its covariance function is

$$K(t, \tau) = \frac{\min(t, \tau)}{\sqrt{t\tau}}.$$

Corresponding covariance operator has the eigenvalues $\lambda_k = (z_{0,k}/2)^2$ and eigenfunctions

$$\varphi_k(t) = J_0(z_{0,k}t)/\sqrt{t}.$$

.

GAUSSIANTY STATISTIC DEFINITION

For application of the Cramér-von Mises-type test for testing the hypothesis $H_{0\mu}$, it need to introduce the function $F^*(t)$ on $(0, 1)^\infty$ such, that it defines the measure Υ . In the finite dimensional case, as a variant of $F^*(t)$ can be chosen the obvious distribution function. It is not possible in the case considered. Instead, it can be proposed the function

$$\begin{aligned} F^*(t) &= P\{T_1 \leq Q_1(t_1), T_2 \leq Q_2(t_2), T_3 \leq Q_3(t_3) \dots\} \\ &= Q_1(t_1)Q_2(t_2)Q_3(t_3) \dots, \quad t = (t_1, t_2, t_3, \dots), \end{aligned} \quad (11)$$

in which functions $Q_i(t)$, $i = 1, 2, \dots$, increase monotonically

The convenient example of the distribution function is

$$F^*(t) = t_1^{r_1} t_2^{r_2} t_3^{r_3} \dots, \quad (12)$$

where $1 = r_1 \geq r_2 \geq r_3 \geq \dots$, and $r_i \rightarrow 0$. Powers r_i must tend to zero sufficiently rapidly.

Let $T^{(i)} = (T_{i1}, T_{i2}, \dots)$ be the observations of T . The distribution function is

$$F_n^*(t) = \frac{1}{n} \sum_{i=1}^{\infty} I_{T_{i1} \leq t_1^{r_1}, T_{i2} \leq t_2^{r_2}, \dots} = \frac{1}{n} \sum_{i=1}^n \prod_{j=1}^{\infty} I_{T_{ij} \leq t_j^{r_j}}. \quad (13)$$

The Cramér-von-Mises statistics can be written analogously to one and multi-dimensional cases

$$\omega_n^{*2} = n \int_{[0,1]^\infty} \left(F_n^*(t) - \prod_{j=1}^{\infty} t_j^{r_j} \right)^2 dt_1 dt_2 \dots \quad (14)$$

Corresponding "empirical process" is

$$\xi_n^*(t) = \sqrt{n} \left(F_n^*(t) - \prod_{j=1}^{\infty} t_j^{r_j} \right), \quad t \in [0, 1]^\infty,$$

or

$$\xi_n^*(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \left(\prod_{j=1}^{\infty} I_{T_{ij} < t_j^{r_j}} - \prod_{j=1}^{\infty} t_j^{r_j} \right), \quad t \in [0, 1]^\infty. \quad (15)$$

The covariance function of $\xi_n^*(t)$ is

$$K^*(s, t) = \prod_{j=1}^{\infty} \min(s_j^{r_j}, t_j^{r_j}) - \prod_{j=1}^{\infty} s_j^{r_j} t_j^{r_j}, \quad s, t \in [0, 1]^\infty. \quad (16)$$

.

The empirical process can be written as the normalized sum of the independent identically distributed random functions in the form

$$\xi_n^*(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^n g_i(t), \quad t \in [0, 1]^\infty, \quad (17)$$

where

$$g_i(t) = \prod_{j=1}^{\infty} I_{T_{ij} < t_j^{r_j}} - \prod_{j=1}^{\infty} t_j^{r_j}.$$

Under the condition that

$$\int_{[0,1]^\infty} K^*(s, s) ds < \infty \quad (18)$$

empirical process weakly converges in $L_2(\mathcal{X})$ to the Gaussian process with the covariance function $K^*(s, t)$.

.

Then the inequality (18) goes into inequality

$$\int_{[0,1]^\infty} K^*(s, s) ds = \prod_{j=1}^{\infty} \frac{1}{r_j + 1} - \prod_{j=1}^{\infty} \frac{1}{2r_j + 1} < \infty.$$

This inequality holds for all positive sequences $\{r_j, j = 1, 2, \dots\}$. We are interested in the case when $1 \geq r_j \downarrow 0, j \rightarrow \infty$, and empirical process does non-degenerate, i.e.

$$\int_{[0,1]^\infty} K^*(s, s) ds > 0. \quad (19)$$

To fulfill this condition, is sufficient to find a sequence for which satisfies the inequality

$$\prod_{j=1}^{\infty} \frac{1}{r_j + 1} > 0. \quad (20)$$

As such a sequence can be selected the sequence $r_j = j^{-a}$, $a > 1$, $j = 1, 2, \dots$. The sequence r_j is not necessarily begins to decrease from 1.

.

It can be written

$$K^*(s, t) = K_0^*(s, t) - w(s)w(t),$$

where

$$K_0^*(s, t) = \prod_{j=1}^{\infty} K_{0i}^*(s_j, t_j) = \prod_{j=1}^{\infty} \min(s_j^{r_j}, t_j^{r_j}),$$

$$w(s) = \prod_{j=1}^{\infty} s_j^{r_j}, \quad s, t \in [0, 1]^{\infty}.$$

.

Monte Carlo results

The limiting distribution of the statistic Ω_n^2 was calculated also by the Monte-Carlo method. The Cramér-von Mises statistic can be represented as

$$\Omega_n^2 = n \int_{[0,1]^\infty} \left(\frac{1}{n} \sum_{i=1}^n \prod_{j=1}^{\infty} I_{T^{ij} < t_j^{r_i}} - \prod_{i=1}^{\infty} t_i^{r_i} \right)^2 dt.$$

Here we used the double-loop calculations by Monte Carlo method. Value of the statistic Ω_n^2 computed in the inner loop, while its distribution is modeled in the outer loop. The number of dimensions in the integral was 100, the number of Monte Carlo iterations to calculate the statistic Ω_n^2 was 500, and the number of Monte Carlo iterations for calculating the percentage points was 10000.

Here are a few estimated quantiles of the limiting distribution of Ω_n^2 , with $r_i = i^{-a(1-i^{-b})}$. When $a = 2.5$ and $b = 0.5$, the exact expectation is $E\Omega^2 \approx 0.1039$ and the estimated expectation is $\hat{E}\omega^2 \approx 0.104$. The corresponding percentage points are

$$P\{\Omega^2 \leq 0.90\} \approx 0.17 \quad \text{and} \quad P\{\Omega^2 \leq 0.95\} \approx 0.21.$$

When $a = 3$ and $b = 0.5$, the exact expectation is $E\Omega^2 \approx 0.1306$ and the estimated expectation is $\hat{E}\Omega^2 \approx 0.132$. The corresponding percentage points are

$$P\{\Omega^2 \leq 0.90\} \approx 0.22 \quad \text{and} \quad P\{\Omega^2 \leq 0.95\} \approx 0.28.$$

EIGENVALUES and EIGENFUNCTIONS of K_{0j}^*

Eigenvalues and eigenfunctions of

$$K_{0i}^*(s_j, t_j) = \min(s_j^{r_j}, t_j^{r_j})$$

⇓

Eigenvalues and eigenfunctions of

$$K_0^*(s, t) = \prod_{j=1}^{\infty} \min(s_j^{r_j}, t_j^{r_j})$$

⇓

Eigenvalues and eigenfunctions of

$$\begin{aligned} K^*(s, t) &= K_0^*(s, t) - w(s)w(t) \\ &= \prod_{j=1}^{\infty} \min(s_j^{r_j}, t_j^{r_j}) - \prod_{j=1}^{\infty} s_j^{r_j} t_j^{r_j}, \quad s, t \in [0, 1]^{\infty} \end{aligned}$$

Limit distribution of the test statistic under the null hypothesis**Eigenvalues and eigenfunctions of K_{0j}^***

First, we shall consider the elementary kernel $K_{0j}^*(t, \tau)$ in the one-dimensional case with arbitrary $0 < r \leq 1$

$$K_{0j}^*(t, \tau) = \min(t^r, \tau^r), \quad 0 \leq t, \tau \leq 1, \quad i = 1, 2, \dots$$

The eigenfunctions and eigenvalues of $K_{0j}^*(t, \tau)$ can be found from the Fredholm integral equation

$$\varphi_r(t) = \lambda_r \int_0^1 \min(t^r, \tau^r) \varphi_r(\tau) d\tau, \quad t \in [0, 1]. \quad (21)$$

Let $\varphi_r(t) = h(t^r)$ and substitute $\varphi_r(t)$ into equation (21). It can be transformed to the equation

$$h_r''(x) + \rho_r x^{\frac{1}{r}-1} h_r(x) = 0, \quad h_r(0) = 0, \quad h_r'(1) = 0. \quad (22)$$

Let $\mu = r/(1+r)$. The eigenvalues of the equation (21) are

$$\lambda_{r,k} = \rho_{r,k} r = r \left(\frac{z_{\mu-1,k}}{2\mu} \right)^2, \quad k = 1, 2, \dots \quad (23)$$

r varies from 1 to zero while μ varies from 1/2 to zero.

.

Corresponding eigenfunctions are

$$\tilde{\varphi}_{r,k}(t) = t^{\mu/(2(1-\mu))} J_{\mu} \left(z_{\mu-1,k} t^{1/(2(1-\mu))} \right), \quad k = 1, 2, \dots \quad (24)$$

The squared normalizing divisor for $\tilde{\varphi}_{r,k}(t)$ is

$$D_{r,k}^2 = \frac{(\mu - 1) z_{\mu-1,k}^{2\mu} \Gamma(2 + \mu) \Gamma(1 + 2\mu)}{\sqrt{\pi}} \times {}_1F_2 \left(\frac{1}{2} + \mu; 2 + \mu, 1 + 2\mu; -z_{\mu-1,k}^2 \right).$$

Here, ${}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q, z)$ is the generalized hypergeometric function defined by the series

$${}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q, z) = 1 + \sum_{m=1}^{\infty} \frac{(a_1)_m \dots (a_p)_m}{(b_1)_m \dots (b_q)_m m!} t^m, \quad (25)$$

and $(a)_m = a(a+1) \dots (a+m-1)$, where $(a)_0 = 1$, is the Pochhammer symbol, $|t| < 1$. We shall use the notation

$$\varphi_{r,k}(t) = \tilde{\varphi}_{r,k}(t) / D_{r,k}$$

.

.

Eigenvalues and eigenfunctions of K_0^*

Further, we obtain the eigenvalues and eigenfunctions of the kernel $K_0^*(t, \tau)$, $t, \tau \in [0, 1]^\infty$. In this case, the Fredholm equation becomes

$$\varphi_r(t) = \lambda \int_{[0,1]^\infty} \prod_{j=1}^{\infty} \min(t_j^{r_j}, \tau_j^{r_j}) \varphi_r(\tau) d\tau, \quad t, \tau \in [0, 1]^\infty. \quad (26)$$

Here, r_j , $j = 1, 2, \dots$, is a prescribed sequence of numbers monotonically decreasing from 1 to 0 and satisfying condition (20). Denote by $\alpha_{j,k} = 1/\lambda_{r_j,k}$, $k = 1, 2, 3, \dots$, the characteristic numbers for equation (21). Let $A_j = \{\alpha_{j,1}, \alpha_{j,1}, \alpha_{j,1}, \dots\}$ be the set of characteristic numbers for equation (21) corresponding to the value $r = r_j$. Then the set of the characteristic numbers for equation (26) consists of all possible products

$$\alpha_{k_1, k_2, k_3, \dots} = \prod_{j=1}^{\infty} \alpha_{j, k_j}, \quad (27)$$

where the (k_1, k_2, k_3, \dots) are all permutation of all positive natural numbers. The eigenfunction corresponding to $\alpha_{k_1, k_2, k_3, \dots}$ is

$$\varphi_{k_1, k_2, k_3, \dots}(t) = \prod_{j=1}^{\infty} \varphi_{r_j, k_j}(t). \quad (28)$$

We propose (hope) that the characteristic numbers $\alpha_{k_1, k_2, k_3, \dots}$ have multiplicities equal to 1. We arrange all the characteristic numbers in descending order, by giving them the serial numbers $\alpha_1 > \alpha_2 > \alpha_3 > \dots$.

The following theorem gives us an idea about the behavior of those characteristic numbers.

THEOREM. *If $r \rightarrow 0$, then the limit form of the Fredholm equation (21) has a unique root equal to one.*

PROOF. The limit form of the equation (21) is

$$\varphi(t) = \lambda \int_0^1 \varphi(\tau) d\tau, \quad t \in [0, 1]. \quad (29)$$

Hence, $\varphi(t)$ is a constant. It can be proved that, in the form (24), $\varphi(t) \equiv 1$. Hence, equation (21) in this case has the only eigenvalue $\lambda = 1$. This assertion follows from Parseval's equality because here, formally, $K(t, t) = 1$ and integrates to 1.

It follows from this theorem that $\lambda_{r_j,1} \rightarrow 1$, but $\lambda_{r_j,k} \rightarrow \infty$ ($\alpha_{j,k} \rightarrow 0$), $k = 2, 3, \dots$, as $j \rightarrow \infty$.

The behavior of the characteristic numbers when r_j tends to zero is shown in Table 1. Here and below, we take $r_i = i^{-a(1-i^{-b})}$.

j	r_j	$\alpha_{j,1}$	$\alpha_{j,2}$	$\alpha_{j,3}$	$\alpha_{j,4}$	$\alpha_{j,5}$...
1	1	0.4053	0.0450	0.0162	0.0083	0.0050	...
2	0.60197	0.5298	0.0463	0.0160	0.0081	0.0048	...
3	0.31323	0.6829	0.0400	0.0132	0.0065	0.0039	...
4	0.17678	0.7917	0.0303	0.0097	0.0047	0.0028	...
5	0.10816	0.8610	0.0219	0.0068	0.0033	0.0019	...
10	0.01952	0.9716	0.0050	0.0015	0.0007	0.0004	...
25	0.00160	0.9976	0.0004	0.0001	0.0001	0.0000	...
50	0.00023	0.9997	0.0001	0.0000	0.0000	0.0000	...
99	0.00003	1.0000	0.0000	0.0000	0.0000	0.0000	...
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
∞	0	1	0	0	0	0	...

We can compute the largest characteristic number

$$\alpha_1 = \prod_{j=1}^{\infty} \alpha_{j,1} \approx 0.0642. \quad (30)$$

It is also obvious that the second largest characteristic number is

$$\alpha_2 = \alpha_1 \alpha_{1,2} / \alpha_{1,1} \approx 0.007137. \quad (31)$$

The corresponding eigenfunctions are

$$\varphi_1(t) = \prod_{j=1}^{\infty} \varphi_{r_j,1}(t) \quad \text{and} \quad \varphi_2(t) = \varphi_1(t) \varphi_{r_1,2}(t) / \varphi_{r_1,1}(t).$$

Notice that the sequence k_1, k_2, k_3, \dots in (27) may contain only a finite number of terms different from 1. Denote $v_{j,k} = \alpha_{j,k} / \alpha_{j,1}$, $j \geq 2$. Then the m -th characteristic number can be represented as

$$\alpha_m = \alpha_1 v_{m,k_{m,1}} v_{m,k_{m,2}} \dots v_{m,k_{m,s_m}}, \quad (32)$$

where $v_{m,k_{m,1}} v_{m,k_{m,2}} \dots v_{m,k_{m,s_m}}$ is the m -th value among the finite products

$$v_{j,k_1} v_{j,k_2} \dots v_{j,k_s} \quad j \geq 2, \quad 1 \leq k_1 < k_2 < \dots < k_s \leq \infty, \quad s \geq 1,$$

ranked in descending order.

.

TABLE 2. Symbolic table for computing the first eigenvalues and characteristic functions of $K_0^*(s, t)$

jk	1 → 5	6 → 10	11 → 15	16 → 20	21 → 25	26 → 30
1	1 2 1 1 3	1 1 1 4 1	1 1 5 1 1	2 1 1 1 6	1 1 2 1 1	7 1 1 1 1
↓	1 1 2 1 1	1 3 1 1 1	1 4 1 1 1	2 1 5 1 1	1 1 1 6 1	1 1 1 1 2
5	1 1 1 2 1	1 1 1 1 3	1 1 1 1 1	1 4 1 1 1	1 1 2 1 1	1 5 1 1 2
	1 1 1 1 1	2 1 1 1 1	1 1 1 1 3	1 1 1 1 1	1 1 1 1 4	1 1 1 1 1
	1 1 1 1 1	1 1 2 1 1	1 1 1 1 1	1 1 1 1 1	3 1 1 1 1	1 1 1 1 1
6	1 1 1 1 1	1 1 1 1 1	2 1 1 1 1	1 1 1 1 1	1 1 1 1 1	1 1 3 1 1
↓	1 1 1 1 1	1 1 1 1 1	1 1 1 2 1	1 1 1 1 1	1 1 1 1 1	1 1 1 1 1
10	1 1 1 1 1	1 1 1 1 1	1 1 1 1 1	1 1 1 2 1	1 1 1 1 1	1 1 1 1 1
	1 1 1 1 1	1 1 1 1 1	1 1 1 1 1	1 1 1 1 1	1 2 1 1 1	1 1 1 1 1
	1 1 1 1 1	1 1 1 1 1	1 1 1 1 1	1 1 1 1 1	1 1 1 1 1	1 1 1 2 1
11	1 1 1 1 1	1 1 1 1 1	1 1 1 1 1	1 1 1 1 1	1 1 1 1 1	1 1 1 1 1
↓	1 1 1 1 1	1 1 1 1 1	1 1 1 1 1	1 1 1 1 1	1 1 1 1 1	1 1 1 1 1
15	1 1 1 1 1	1 1 1 1 1	1 1 1 1 1	1 1 1 1 1	1 1 1 1 1	1 1 1 1 1
	1 1 1 1 1	1 1 1 1 1	1 1 1 1 1	1 1 1 1 1	1 1 1 1 1	1 1 1 1 1
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 2 represents the realisation of the formula (32) for the considered sequence r_i . Denote its elements as s_{jk} .

Column k in the table corresponds to the k -th characteristic number α_k of K_0^* in descending order by the formula

$$\alpha_k = \prod_{j=1}^{\infty} \alpha_{j, s_{jk}}.$$

Table 3 shows some real values of the characteristic numbers in accordance with Table 2.

jk	1 → 5					6 → 10				
1 ↓ 5	$\alpha_{1,1}$	$\alpha_{1,2}$	$\alpha_{1,1}$	$\alpha_{1,1}$	$\alpha_{1,3}$	$\alpha_{1,1}$	$\alpha_{1,1}$	$\alpha_{1,1}$	$\alpha_{1,4}$	$\alpha_{1,1}$
	$\alpha_{2,1}$	$\alpha_{2,1}$	$\alpha_{2,2}$	$\alpha_{2,1}$	$\alpha_{2,1}$	$\alpha_{2,1}$	$\alpha_{2,3}$	$\alpha_{2,1}$	$\alpha_{2,1}$	$\alpha_{2,1}$
	$\alpha_{3,1}$	$\alpha_{3,1}$	$\alpha_{3,1}$	$\alpha_{3,2}$	$\alpha_{3,1}$	$\alpha_{3,1}$	$\alpha_{3,1}$	$\alpha_{3,1}$	$\alpha_{3,1}$	$\alpha_{3,3}$
	$\alpha_{4,1}$	$\alpha_{4,1}$	$\alpha_{4,1}$	$\alpha_{4,1}$	$\alpha_{4,1}$	$\alpha_{4,2}$	$\alpha_{4,1}$	$\alpha_{4,1}$	$\alpha_{4,1}$	$\alpha_{4,1}$
	$\alpha_{5,1}$	$\alpha_{5,1}$	$\alpha_{5,1}$	$\alpha_{5,1}$	$\alpha_{5,1}$	$\alpha_{5,1}$	$\alpha_{5,1}$	$\alpha_{5,2}$	$\alpha_{5,1}$	$\alpha_{5,1}$

Eigenvalues of K^*

The characteristic numbers α_i^* of $K^*(s, t)$ are solutions of the equation

$$\sum_{k=1}^{\infty} \frac{C_k^2}{\alpha_k - \alpha} = 1,$$

where

$$C_k = \prod_{j=1}^{\infty} C_{j,k}, \quad k = 1, 2, \dots$$

and

$$C_{j,k} = \int_0^1 w_j(t) \varphi_{j,k}(t) dt = \int_0^1 t^{\mu_j/1 - \mu_j} \varphi_{j,k}(t) dt, \quad j = 1, 2, \dots$$

From this, it can be derived that

$$C_{r,k}^2 = \left(2^{-\mu} (\mu - 1) \Gamma(2 + \mu) z_{\mu-1,k}^{\mu} {}_0F_1 \left(; 2 + \mu; -\frac{1}{4} z_{\mu-1,k}^2 \right) \right)^2 / D_{r,k}^2,$$

where ${}_0F_1$ is the *generalized hypergeometric function*

$${}_0F_1(; a; z) = \sum_{k=0}^{\infty} \frac{z^k}{(a)_k k!}.$$

.

Smirnov formula

In the previous section, we obtained the eigenvalues $\lambda_k^* = 1/\alpha_k^*$ for the covariance function $K^*(s, t)$. For calculation of the limiting distribution function of Ω^2 we can use the Smirnov formula, namely, for $t > 0$,

$$P\left(\Omega^2 > t\right) = \frac{1}{\pi} \sum_{k=1}^{\infty} (-1)^{k+1} \int_{\lambda_{2k-1}^*}^{\lambda_{2k}^*} \frac{e^{-tu/2} du}{u \sqrt{\left| \prod_{k=1}^{\infty} \left\{ 1 - \frac{u}{\lambda_k^*} \right\} \right|}}.$$

The Smirnov formula is designed for distinct eigenvalues.

.

The limiting distribution of Ω_n^2

It was obtained one thousand values $\alpha_i^* = 1/\lambda_i^*$. The distribution of the statistic Ω_n^2 is approximated by a finite quadratic form

$$Q_{100} = \sum_{k=1}^{100} \frac{z_k^2}{\lambda_k^*},$$

where z_i are independent identically distributed random variables with standard normal distribution. The residue

$$\sum_{k=101}^{\infty} \frac{z_k^2}{\lambda_k^*}$$

of the full quadratic form is replaced by its expectation 0.0296327. As a result, it was obtained the following percent points of the statistic Ω^2 :

$$P\{\Omega^2 \leq 0.90\} \approx 0.16450, \quad P\{\Omega^2 \leq 0.95\} \approx 0.20371,$$

$$P\{\Omega^2 \leq 0.99\} \approx 0.30039, \quad P\{\Omega^2 \leq 0.995\} \approx 0.34350,$$

$$\{\Omega^2 \leq 0.999\} \approx 0.44563.$$

.

BIBLIOGRAPHY

Deheuvels, P., Martynov, G. (2003) Karhunen-Loève expansions for weighted Wiener processes and Brownian bridges via Bessel functions., *Progress in Probability*, Birkhäuser, Basel/Switzerland, 55, 57–93.

Darling, D. A. (1955). The Cramér-Smirnov test in the parametric case. *Ann. Math. Statist.* 26 1–20.

Martynov, G. V. (1975). Computation of distribution function of quadratic forms of normally distributed random variables. *Theor. Probab. Appl.* 20 782–793.

Krivjakova E. N., Martynov G. V., and Tjurin J. N. (1977) The distribution of the omega square statistics in the multivariate case. "*Theory of Probability and its Applications.*", V. 22, N 2, 415–420.

.

Martynov, G. V., (1979) The omega square tests, Nauka, Moscow, 80pp., (in Russian).

Martynov, G. V., (1992) Statistical tests based on empirical processes and related questions, J. Soviet. Math, 61, 2195–2271.

Van der Vaart, A. W., Wellner, J. A. (1996) Weak Convergence and Empirical Processes. Springer-Verlag, 508 pp.

END