1

Bahadur local asymptotic optimality for generalizations of the Cramér-von Mises and the Anderson-Darling statistics

J.-R. Pycke*

March 31, 2004

Abstract

We introduce a family of statistics generalizing that of Cramér - von Mises and Anderson-Darling. The latter are known to be locally Bahadur optimal in the case of location alternatives with the hyperbolic cosine and the logistic densities respectively. We generalize this property to the whole family of statistics and a corresponding family of densities.

Keywords and phrases: goodness of fit test, Cramér-von Mises statistic, Anderson-Darling statistic, Bahadur efficiency, Bahadur local optimality.
2000 Mathematics Subject Classification : 62H15, 62G20, 62G30.

1 Introduction

In this paper we extend to a new family of statistics some of the well-known properties of the celebrated Cramér-von Mises and Anderson-Darling statistics (discussed for example in [2]). The latter are commonly used in the following goodness of fit problem. Let $X_1, ..., X_n$ be independent observations on a random variable X with continuous cumulative distribution function (d.f.) $F(x) = P(X \leq x)$. Suppose that we wish to test the hypothesis

$$H_0 : \; F = F_0$$

where F_0 is a completly specified continuous d.f., against the general alternative $F \neq F_0$. For this purpose many tests of goodness of fit have been

*Jean-Renaud Pycke
Université Paris VI,
L.S.T.A, case 158, 175 rue du Chevaleret, 75013 Paris,
e-mail:pycke@ccr.jussieu.fr

introduced, often based on measures of discrepancy between the hypothesized d.f., F_0, and the sample d.f.

$$F_n(x) := \frac{\text{number of observations} \leq x}{n}, \; x \in \mathbb{R}.$$

A wide class of tests springs from the discrepancy measure given by the Cramér-von Mises family of quadratic statistics

$$\omega_{q,n}^2 := n \int_{-\infty}^{\infty} \big\{F_n(x) - F_0(x)\big\}^2 q(F_0(x)) dF_0(x), \quad (n \geq 1) \tag{1.1}$$

where $q : (0,1) \to [0,\infty)$ is a suitable weight-function. If we introduce the uniform empirical process defined on $[0,1]$ by

$$\mathbf{U}_n(t) := n^{-1/2} \sum_{i=1}^{n} \big(1_{\{F_0(X_i) \leq t\}} - t\big),$$

(1.1) may be rewritten

$$\omega_{q,n}^2 = \int_0^1 q(t) \mathbf{U}_n^2(t) dt. \tag{1.2}$$

When $q(t) = 1$ the statistic is the Cramér-von Mises statistic ; when $q(t) = [t(1-t)]^{-1}$ it is the Anderson-Darling statistic.

In section 2 we introduce a family (q_ν) of weight-functions indexed by $\nu \in (0,\infty)$ with $q_1(t) = 1$, $q_2(t) = [t(1-t)]^{-1}$ and such that q_ν is summable for $\nu \in (0,2)$. We define the corresponding Cramér-von Mises type statistics $\omega_{\nu,n}^2$. Proposition 2.4 gives their asymptotic distribution under the null hypothesis in the case $\nu \in [1,\infty)$. In section 3 we generalize the following property. It is well known (see [9] Chapter 6, Corollary 1 p. 225 and Theorem 6.3.5 p. 227) that the Cramér-von Mises and the Anderson-Darling statistics are locally asymptotically optimal (LAO) in the sense of Bahadur in the case of location alternatives for the hyperbolic cosine and the logistic distributions respectively. Theorem 3.18 extends this result to the statistics $\omega_{\nu,n}^2$ in the case $\nu \in [1,2]$, stating that $\omega_{\nu,n}^2$ is LAO in the case of location alternatives for the Gumbel generalized logistic distribution f_ν given by (3.21). For results and references about this density, the reader is refered to [8], Chapter 22, Section 11. We give in Section 4 some useful technical results.

Most results of the present paper are based upon those expounded in [10], in which a new family of explicit Karhunen-Loève expansions is given, including that of the Brownian bridge and the Anderson-Darling processes as particular cases. These results enabled us to state Proposition 2.3 which is the key property for the proofs of Proposition 2.4 and Theorem 3.1. Recent examples of interesting statistical applications arising from the explicit knowledge of Karhunen-Loève expansions can be found in [6] and [7].

Let us recall the following basic facts about Karhunen-Loève expansions. For details, the reader is refered to [12], Chapter 5. If $\mathbf{X} = \{\mathbf{X}(t) : 0 < t < 1\}$ is a centered Gaussian process with continuous covariance function $K(s,t) = \mathbf{E}\mathbf{X}(s)\mathbf{X}(t)$ such that

$$\int_0^1 K(t,t)dt < \infty,$$

one has almost surely

$$\mathbf{X}(t) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\, \xi_k\ f_k(t) \quad \text{in } L^2(0,1) \tag{1.3}$$

where $\{\xi_k : k \geq 1\}$ denotes a sequence of independent $N(0,1)$ variables and the set $\{(\lambda_k, f_k) : k \geq 1\}$ has the following properties :

$$P1 : \forall k \geq 1,\ (\lambda_k, f_k) \in [0,\infty) \times L^2(0,1);$$
$$P2 : \text{the sequence } (\lambda_k) \text{ is decreasing;}$$
$$P3 : \forall k \geq 1,\ \int_0^1 K(s,.)f_k(s)ds = \lambda_k f_k(.);$$
$$P4 : \int_0^1 f_k(s)f_\ell(s)ds = \begin{cases} 1 & \text{if } k = \ell, \\ 0 & \text{if } k \neq \ell. \end{cases}$$

Consequently one has the equality

$$\int_0^1 \mathbf{X}^2(t)dt = \sum_{k=1}^{\infty} \lambda_k \xi_k^2, \tag{1.4}$$

and the characteristic function of the random variable on the left-hand side of (1.4) is given by

$$\exp\{iu \int_0^1 \mathbf{X}^2(t)dt\} = \prod_{k=1}^{\infty} \left(1 - 2iu\lambda_k\right)^{-1/2},\ (u \in \mathbb{R}). \tag{1.5}$$

In the sequel (1.3) will be refered to as the Karhunen-Loève (K-L) expansion of $\mathbf{X}$.

2 A family of statistics

To begin with we recall that the normalized incomplete beta function I is defined (see e.g. [1] formulas 6.2.2 and 26.5.1), for $\alpha,\ \beta > 0$ by

$$I(\alpha,\beta,x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^x t^{\alpha-1}(1-t)^{\beta-1}dt \qquad \text{for } 0 \leq x \leq 1. \tag{2.6}$$

For each $\alpha,\ \beta > 0$ the function $I(\alpha, \beta, .)$ is increasing on $[0, 1]$ and we have

(2.7) $I(\alpha, \beta, 0) = 0, \quad I(\alpha, \beta, 1 - x) = 1 - I(\beta, \alpha, x) \quad$ hence $I(\alpha, \beta, 1) = 1$.

Particular cases are

$$(2.8) \qquad I(\frac{1}{2}, \frac{1}{2}, x) = \frac{\arccos(1 - 2x)}{\pi} \quad \text{and} \quad I(1, 1, x) = x \quad \text{for } x \in [0, 1].$$

Definition 2.1. For each $\nu \in (0, \infty)$ we set

$$(2.9) \qquad \omega^2_{\nu,n} := \pi^{2\nu-3} \frac{2^{\nu-1}\Gamma^2(\frac{\nu}{2})}{\Gamma(\nu)} \int_0^1 \frac{\mathbf{U}_n^2\{I(\frac{\nu}{2}, \frac{\nu}{2}, \frac{1-\cos(\pi r)}{2})\}}{\sin^{\nu-1}(\pi r)} dr.$$

Hence $\omega^2_{1,n} = \int_0^1 \mathbf{U}_n^2(r)dr$ is the Cramér-von Mises statistic and

$$\omega^2_{2,n} = 2\pi \int_0^1 \frac{\mathbf{U}_n^2(\frac{1-\cos(\pi r)}{2})}{\sin(\pi r)} dr = \int_0^1 \frac{\mathbf{U}_n^2(t)}{t(1-t)} dt$$

is the Anderson-Darling statistic.

In order to express the statistic $\omega^2_{\nu,n}$ in the form (1.2), we introduce the following notations. For each $\nu > 0$, J_ν will denote the function uniquely defined on $[0, 1]$ by

$$t = I(\frac{\nu}{2}, \frac{\nu}{2}, \frac{1 - \cos(\pi r)}{2}) \iff J_\nu(t) = r \quad \text{for } r, t \in [0, 1].$$

In view of (2.8) it is readily checked that

$$(2.10) \qquad J_1(t) = t, \quad J_2(t) = \frac{1}{\pi} \arccos(1 - 2t).$$

For each $\nu \in (0, \infty)$, we define the weight-function q_ν by setting, for $t \in (0, 1)$,

$$(2.11) \qquad q_\nu(t) := \pi^{2\nu-4} \Big(\frac{2^{\nu-1}\Gamma^2(\frac{\nu}{2})}{\Gamma(\nu)}\Big)^2 \sin^{2(1-\nu)}\{\pi J_\nu(t)\}.$$

From (2.10) we see that, as claimed in the introduction,

$$q_1(t) = 1,\ q_2(t) = \frac{1}{t(1-t)}.$$

Furthermore, from Lemma 4.3 (applied with f constant), and from the combination of Lemma 4.2 and Lemma 4.3 (applied with $f(t) = t(1 - t)$), we obtain respectively

(2.12)

$$\forall \nu \in (0, 2) \ : \ \int_0^1 q_\nu(t)dt < \infty \quad \text{and} \quad \forall \nu > 0 \ : \ \int_0^1 t(1 - t)q_\nu(t)dt < \infty.$$

Proposition 2.1. *The statistic $\omega^2_{\nu,n}$ can be written as*

$$\omega^2_{\nu,n} = \int_0^1 q_\nu(t)\mathbf{U}^2_n(t)dt. \tag{2.13}$$

Proof. If we use Lemma 4.3 with $f = \mathbf{U}^2_n$ and multiply both sides of (4.29) by $\pi^{2\nu-3}\frac{2^{\nu-1}\Gamma^2(\frac{\nu}{2})}{\Gamma(\nu)}$, we obtain

$$\pi^{2\nu-3}\frac{2^{\nu-1}\Gamma^2(\frac{\nu}{2})}{\Gamma(\nu)}\frac{2^{\nu-1}\Gamma^2(\frac{\nu}{2})}{\pi\Gamma(\nu)}\int_0^1 \frac{\mathbf{U}^2_n(t)}{\sin^{2(\nu-1)}\{\pi J_\nu(t)\}}dt$$
$$= \pi^{2\nu-3}\frac{2^{\nu-1}\Gamma^2(\frac{\nu}{2})}{\Gamma(\nu)}\int_0^1 \frac{\mathbf{U}^2_n\big\{I(\frac{\nu}{2},\frac{\nu}{2},\frac{1-\cos(\pi r)}{2})\big\}}{\sin^{\nu-1}(\pi r)}dr,$$

which is the desired equality. □

Recall that a Brownian bridge is a centered Gaussian process on $[0,1]$ with covariance function $\mathbf{E}\mathbf{B}(s)\mathbf{B}(t) = \min(s,t) - st$ for $s,t \in [0,1]$

Proposition 2.2. *For each $\nu > 0$ we have, under the null hypothesis, the convergence in law*

$$\lim_{n\to\infty} \omega^2_{\nu,n} = \int_0^1 q_\nu(t)\mathbf{B}^2(t)dt \tag{2.14}$$

where $\mathbf{B}$ is a Brownian bridge.

Proof. This result is a consequence of the first assertion in (2.12) in combination with Theorem 3.3 p. 325 in [5]. □

The next proposition will enable us to compute the characteristic function of the random variable on the right-hand side of (2.14) in the case $\nu \geq 1$. For each $\nu \in [1,\infty)$ and $k \in \mathbb{N}^*$, let $f_{\nu,k}$ denote the function defined on $(0,1)$ by

$$f_{\nu,k}(t) := \alpha_{\nu,k}\sin^{1-\frac{\nu}{2}}\{\pi J_\nu(t)\}P^{-\frac{\nu}{2}}_{\frac{\nu}{2}+k-1}\big[\cos\{\pi J_\nu(t)\}\big],$$

where $\alpha_{\nu,k}$ is the positive real number such that $\int_0^1 f^2_{\nu,k}(t)dt = 1$, and $P^{-\frac{\nu}{2}}_{\frac{\nu}{2}+k-1}$ denotes a Legendre function of the first kind, hence satisfies

$$P^{-\frac{\nu}{2}}_{\frac{\nu}{2}+k-1}(x) = \frac{(-1)^{k-1}(1-x^2)^{-\frac{\nu}{4}}}{2^{\frac{\nu}{2}+k-1}\Gamma(\frac{\nu}{2}+k)}\frac{\mathrm{d}^{k-1}}{\mathrm{d}x^{k-1}}(1-x^2)^{\frac{\nu}{2}+k-1}, \quad \text{for } -1 < x < 1,$$

(see [11], formula (359) p. 184).

Proposition 2.3. *Suppose* $\nu \in [1, \infty)$. *The process* $\{\sqrt{q_\nu(t)}\ \mathbf{B}(t) : 0 < t < 1\}$ *admits a K-L expansion given by*

$$\frac{\pi^{\nu-2}2^{\nu-1}\Gamma(\frac{\nu}{2})^2}{\Gamma(\nu)} \frac{\mathbf{B}(t)}{\sin^{\nu-1}\{\pi J_\nu(t)\}} = \sum_{k=1}^{\infty} \Big\{\frac{\pi^{2\nu-4}}{k(k+\nu-1)}\Big\}^{1/2} \xi_k\ f_{\nu,k}(t). \tag{2.15}$$

Proof. It is proved in [10] that for each $\nu \in [1, \infty)$ a K-L expansion of form

$$\frac{2^{\nu-1}\Gamma(\frac{\nu}{2})^2}{\pi\Gamma(\nu)} \frac{\mathbf{B}(t)}{\sin^{\nu-1}\{\pi J_\nu(t)\}} = \sum_{k=1}^{\infty} \Big\{\frac{1}{\pi^2 k(k+\nu-1)}\Big\}^{1/2} \xi_k\ f_{\nu,k}(t)$$

is valid on $(0, 1)$. On multiplying the latter by $\pi^{\nu-1}$ we obtain the K-L expansion (2.15). □

The next corollary follows readily from the preceding proposition.

Corollary 2.1. For each $\nu \in [1, \infty)$, one has the equality in law

$$\int_0^1 q_\nu(t)\mathbf{B}^2(t)dt = \sum_{k=1}^{\infty} \frac{\pi^{2\nu-4}}{k(k+\nu-1)}\ \xi_k^2. \tag{2.16}$$

In view of (1.4), (1.5) and (2.16) we obtain the following result.

Proposition 2.4. *For each* $\nu \in [1, \infty)$, *under the null hypothesis, one has almost surely*

$$\lim_{n\to\infty} \exp\{iu\omega_{\nu,n}^2\} = \prod_{k=1}^{\infty} \Big(1 - \frac{2iu\pi^{2\nu-4}}{k(k+\nu-1)}\Big)^{-1/2},\ \ (u \in \mathbb{R}).$$

3 Bahadur local efficiency in the case of location alternatives

One way of measuring the efficiency of tests of fit is that introduced by Bahadur [3]. For a recent exposition of the concept of Bahadur efficiency and a comparison with other types of efficiencies we refer to [9]. We briefly recall the following basic facts. Suppose that the true d.f. corresponds to a probability measure that belongs to a parametric set $\{P_\theta : \theta \in \mathbb{R}\}$. The d.f. and the density function corresponding to P_θ will be denoted respectively by $F(\theta, .)$ and $f(\theta, .)$. We wish to test $H_0 : \theta = \theta_0$ against the alternative $H_1 : \theta \neq \theta_0$. Let s denote the sequence of observations $\{X_k : k \geq 1\}$ and $\{T_n\}$ a sequence of statistics based on s such that T_n depends on s only through $X_1, ..., X_n$, i.e. $T_n(s) = T_n(X_1, ..., X_n)$. Whithout loss of generality, it is assumed that the rejection region of H_0 is given by

$$\{s : T_n(s) \geq c\}$$

where $c \in \mathbb{R}$. Let ϕ_n denote the null distribution of T_n, that is

$$\phi_n(t) := P_{\theta_0}(T_n < t) \quad \text{for } t \in \mathbb{R}.$$

The level attained by T_n is defined to be

$$L_n(s) := 1 - \phi_n(T_n(s)).$$

If for each $\theta \neq \theta_0$ and a certain nonrandom positive function $c_T(\theta)$ the following convergence in P_θ-probability takes place :

$$\lim_{n \to \infty} n^{-1} \log L_n(s) = -\frac{1}{2} c_T(\theta),$$

then one says that the sequence of statistics $T = \{T_n\}$ has exact slope $c_T(\theta)$. A fundamental result in the Bahadur theory is that the local exact slope of any sequence of statistics $\{T_n\}$ satisfies the inequality

$$c_T(\theta) \leq 2K(\theta, \theta_0) \tag{3.17}$$

where $K(\theta, \theta')$ is the Kullback-Leibler information for two elements $f(\theta, .)$ and $f(\theta', .)$ of the familiy, defined by

$$K(\theta, \theta') := \int_{-\infty}^{\infty} \log \frac{f(\theta, x)}{f(\theta', x)} f(\theta, x) dx.$$

Statistics for which equality holds in (3.17) for each $\theta \neq \theta_0$ are said to be *asymptotically optimal*. A sequence of statistics $T = \{T_n\}$ is said to be *locally asymptotically optimal* (LAO) in the Bahadur sense if it satisfies the weaker condition

$$c_T(\theta) \sim 2K(\theta, \theta_0) \text{ as } \theta \to \theta_0. \tag{3.18}$$

Suppose moreover that for each fixed $x \in \mathbb{R}$ the function $\theta \mapsto \sqrt{f(\theta, x)}$ is differentiable with respect to θ and that the Fisher information defined by

$$I(\theta) := \int_{-\infty}^{\infty} \frac{\{\frac{\partial f(\theta,x)}{\partial x}\}^2}{f(\theta, x)} dx \tag{3.19}$$

is continuous. Then the Kullback-Leibler information can be shown to satisfy (see e.g. [4] Chapter 2, § 21, Theorem 2 and Remark 1 p. 205),

$$K_\nu(\theta_0, \theta) \sim \frac{I(\theta_0)}{2} \, (\theta - \theta_0)^2 \text{ as } \theta \to \theta_0.$$

Consider the particular case of a location family, i.e. when

$$F(\theta,.) = F(0,\theta+.) \quad \text{for each } \theta \in \mathbb{R}.$$

In this case the Fisher information is easily seen to be a constant I, and it is readily checked that $K(\theta,0) = K(0,-\theta)$. Thus if we fix $\theta_0 = 0$, the condition for LAO (3.18) reduces to

$$c_T(\theta) \sim I(0)\theta^2 \text{ as } \theta \to 0. \tag{3.20}$$

For each specified $\nu \in (0,\infty)$ we consider the location family such that for each $\theta \in \mathbb{R}$,

$$F_\nu(\theta,x) = I(\frac{\nu}{2},\frac{\nu}{2},\frac{1+\tanh(\frac{x+\theta}{\nu})}{2}),\ (x \in \mathbb{R}).$$

It follows from (4.30) that the corresponding density function is given by

$$f_\nu(\theta,x) := \frac{\Gamma(\nu)}{\nu 2^{\nu-1}\Gamma^2(\frac{\nu}{2})}\frac{1}{\cosh^\nu(\frac{x+\theta}{\nu})}. \tag{3.21}$$

The value $\nu = 1$ gives the hyperbolic cosine distribution with

$$f_1(0,x) = \frac{1}{\pi\cosh x} \text{ and } F_1(0,x) = \frac{2}{\pi}\arctan e^x.$$

For $\nu = 2$ we obtain the logistic distribution characterized by

$$f_2(0,x) = \frac{e^x}{(1+e^x)^2} = \frac{1}{4\cosh^2(\frac{x}{2})} \text{ and } F_2(0,x) = \frac{1}{1+e^{-x}} = \frac{1+\tanh(\frac{x}{2})}{2}.$$

For each ν, the quantities I, $K(\theta)$ and $c_{\omega^2_{\nu,n}}(\theta)$ will be denoted by I_ν, K_ν and $c_\nu(\theta)$.

Proposition 3.1. *For each $\nu > 0$, the Fisher information satisfies*

$$I_\nu = \frac{1}{\nu+1}. \tag{3.22}$$

Proof. From (3.21) and (3.19) we see that

$$I_\nu = I_\nu(0) = \frac{\Gamma(\nu)}{\nu 2^{\nu-1}\Gamma^2(\frac{\nu}{2})}\int_{-\infty}^{\infty}\{-\nu\cosh^{-\nu-1}(\frac{x}{\nu})\cdot\frac{1}{\nu}\sinh(\frac{x}{\nu})\}^2\cosh^\nu(\frac{x}{\nu})dx$$
$$= \frac{\Gamma(\nu)}{2^{\nu-1}\Gamma^2(\frac{\nu}{2})}\int_{-\infty}^{\infty}\frac{\sinh^2 x}{\cosh^{\nu+2} x}dx$$

and the result is a consequence of (4.31). □

In order to compute $c_\nu(\theta)$ in the case $\nu \in (0,2)$, we use the following result (see [9] Section 2.6, Table 2 p.76 and the remarks following relation (2.6.2) p. 77-78). When $T = \{T_{q,n}\}$ is a weighted Cramér-von Mises type statistic, that is to say of the form

$$T_{q,n} := \int_0^1 q(t)\mathbf{U}_n^2(t)dt$$

where $q : (0,1) \to \mathbb{R}^+$ is a measurable function, then under the summability condition

$$\int_0^1 q(t)dt < \infty, \tag{3.23}$$

the local exact slope of $\{T_{q,n}\}$ is given by

$$c_q(\theta) = a(q) \int_{-\infty}^{\infty} \big\{F(\theta, y) - F(\theta_0, y)\big\}^2 q\big(F(\theta_0, y)\big) dF(\theta,y)dy, \tag{3.24}$$

with

$$a(q) := -\lim_{x\to\infty} \frac{2}{x^2} \log\Big(\int_0^1 q(t)\mathbf{B}^2(t)dt \geq x^2\Big). \tag{3.25}$$

Proposition 3.2. *Suppose $\nu \in (0,2)$. As $\theta \to 0$ the local exact slope $c_\nu(\theta)$ satisfies*

$$c_\nu(\theta) \sim \frac{\theta^2}{\lambda_1(\nu)} \int_{-\infty}^{\infty} \Big\{\frac{\partial F_\nu(\theta, y)}{\partial \theta}\Big|_{\theta=0}\Big\}^2 \; q\{F_\nu(0,y)\} \; f_\nu(0,y)dy \tag{3.26}$$

$$= \frac{\theta^2}{\nu+1} \frac{\pi^{2\nu-4}}{\lambda_1(\nu)\nu} \tag{3.27}$$

where $\lambda_1(\nu)$ is the first eigenvalue in the K-L expansion of the process $\{\sqrt{q_\nu(t)}\,\mathbf{B}(t) : 0 < t < 1\}$.

Proof. Since (3.23) is satisfied for $\nu \in (0,2)$, it is legitimate to use (3.24). When $q = q_\nu$ we note $a(q) = a_\nu$. Firstly we infer from Lemma 4.1 that

$$q\{F_\nu(0,y)\} = \pi^{2\nu-4}\Big(\frac{2^{\nu-1}\Gamma^2(\frac{\nu}{2})}{\Gamma(\nu)}\Big)^2 \cosh^{-2(1-\nu)} y. \tag{3.28}$$

Secondly from (4.32) we see that

$$\frac{\partial F_\nu(\theta, y)}{\partial \theta} = \frac{2^{1-\nu}}{\nu} \frac{\Gamma(\nu)}{\Gamma^2(\frac{\nu}{2})} \cosh^{-\nu}\Big(\frac{x+\theta}{\nu}\Big).$$

Consequently for each $\nu > 0$ we have

$$\begin{aligned}\{\frac{F(\theta,y)-F(0,y)}{\theta}\}^2 f_\nu(0,y) &\leq \big|\sup_{y\in\mathbb{R}}\sup_{\theta\in\mathbb{R}} \frac{\partial F_\nu(\theta,y)}{\partial\theta}\big|^2 f_\nu(0,y) \\ &= \frac{2^{2-2\nu}}{\nu^2}\{\frac{\Gamma(\nu)}{\Gamma^2(\frac{\nu}{2})}\}^2 f_\nu(0,y).\end{aligned}$$

From this inequality and (3.28) we deduce the existence of $M > 0$ such that for each $\theta \in \mathbb{R}$,

$$\{\frac{F(\theta,y)-F(0,y)}{\theta}\}^2 q\{F_\nu(0,y)\} f_\nu(0,y) \leq M\cosh^{-2(1-\nu)-\nu}(\frac{y}{\nu}) = M\cosh^{-2+\nu}(\frac{y}{\nu}).$$

This means that we may use the theorem of Lebesgue on dominated convergence in (3.24), as $\theta \to 0$, to obtain

$$c_\nu(\theta) \sim a(q)\theta^2 \int_{-\infty}^{\infty} \{\frac{\partial F_\nu(\theta,y)}{\partial\theta}\big|_{\theta=0}\}^2 \; q\{F_\nu(0,y)\}\; f_\nu(0,y)dy.$$

Next we have

$$\begin{aligned}&\int_{-\infty}^{\infty} \{\frac{\partial F_\nu(\theta,y)}{\partial\theta}\big|_{\theta=0}\}^2 \; q\{F_\nu(0,y)\}\; f_\nu(0,y)dy \\ &\qquad = \frac{\pi^{2\nu-4}}{\nu^2}\frac{\Gamma(\nu)}{\nu 2^{\nu-1}\Gamma^2(\frac{\nu}{2})}\int_{-\infty}^{\infty}\frac{dx}{\cosh^{-2\nu-2(1-\nu)-\nu}(\frac{x}{\nu})} \\ &\qquad\qquad = \frac{\pi^{2\nu-4}\Gamma(\nu)}{\nu^2 2^{\nu-1}\Gamma^2(\frac{\nu}{2})}\int_{-\infty}^{\infty}\frac{dx}{\cosh^{\nu+2}x} = \frac{\pi^{2\nu-4}}{\nu(\nu+1)},\end{aligned}$$

where the last equality follows from Lemma 4.4. Thus in order to establish (3.26) and (3.27) there remains only to calculate $a(q)$ thanks to (3.25) in the case $q = q_\nu$. But it is shown in [13] that a centered Gaussian process $\{\mathbf{X}(t) : 0 < t < 1\}$ for which the representation (1.4) holds satisfies

$$\lim_{x\to\infty}\frac{2}{x^2}\log \mathbf{P}\Big(\int_0^1 \mathbf{X}^2(t)dt \geq x^2\Big) = \frac{1}{\lambda_1}.$$

And λ_1 is nothing else but the first eigenvalue of the K-L expansion (1.3) of $\mathbf{X}$. This completes the proof. □

Theorem 3.1. *For each $\nu \in [1,2]$ the sequence of statistics $\{\omega^2_{\nu,n}\}$ is locally asymtotically optimal in the sense of Bahadur, for testing*

$$H_0(\nu) \;:\; F = F_\nu(0,.)$$

against the location alternative $H_1(\nu) \;:\; F = F_\nu(\theta,.)$ with $\theta \neq 0$.

Proof. The result is already known to be true in the case $\nu = 2$. For $0 < \nu < 2$, we first use Proposition 2.3, which implies $\lambda_1(\nu) = \pi^{2\nu-4}/\nu$. Then in view of (3.22) and (3.27), we see that the LAO condition (3.20) is satisfied. □

4 Useful technical results

Lemma 4.1. *For each* $x \in \mathbb{R}$,

$$\sin^2\Big(\pi J_\nu\big\{I(\frac{\nu}{2},\frac{\nu}{2},\frac{1+\tanh x}{2})\big\}\Big) = \cosh^{-2} x.$$

Proof. Let $r \in \mathbb{R}$ be such that $\tanh x = -\cos(\pi r)$, hence $\cosh^{-2} x = \sin^2(\pi r)$. Then by definition of J_ν one has

$$J_\nu\big\{I(\frac{\nu}{2},\frac{\nu}{2},\frac{1+\tanh x}{2})\big\}\Big) = r.$$

Combining these two equalities, we obtain the claimed result. □

Lemma 4.2. *Let* $f(t) = t(1-t)$. *For each* $\nu > 0$,

$$f\big\{I(\frac{\nu}{2},\frac{\nu}{2},\frac{1-\cos(\pi r)}{2})\big\} \sim \frac{\pi^\nu \Gamma(\nu)}{\nu 2^{\nu-1}\Gamma^2(\frac{\nu}{2})}\sin^\nu(\pi r) \quad \textit{as } r \to 0^+ \textit{ or as } r \to 1^-.$$

Proof. Formula 26.5.4 in [1] implies that for $\alpha, \beta > 0$, one has

$$I(\alpha,\beta,x) \sim \frac{\Gamma(\alpha+\beta)}{\alpha\Gamma(\alpha)\Gamma(\beta)}x^\alpha \quad \text{as } x \to 0^+,$$

which in view of the second equality of (2.7) implies in turn

$$I(\frac{\nu}{2},\frac{\nu}{2},x) \sim \frac{\Gamma(\nu)}{(\frac{\nu}{2})\Gamma^2(\frac{\nu}{2})}x^{\nu/2} \quad \text{as } x \to 0^+$$
$$\text{and} \quad 1 - I(\frac{\nu}{2},\frac{\nu}{2},x) \sim \frac{\Gamma(\nu)}{(\frac{\nu}{2})\Gamma^2(\frac{\nu}{2})}(1-x)^{\nu/2} \text{ as } x \to 1^-.$$

Thus

$$f\big(I(\frac{\nu}{2},\frac{\nu}{2},x)\big) \sim \frac{\Gamma(\nu)}{(\frac{\nu}{2})\Gamma^2(\frac{\nu}{2})}\big[x(1-x)\big]^{\nu/2} \text{ as } x \to 0^+ \text{ or as } x \to 1^-,$$

and the result follows since

$$\big[\frac{1-\cos(\pi r)}{2}\cdot\frac{1+\cos(\pi r)}{2}\big]^{\nu/2} \sim \big(\frac{\sin(\pi r)}{2}\big)^\nu \quad \text{as } r \to 0^+ \text{ or as } r \to 1^-.$$

□

Lemma 4.3. *If* f *is measurable on* $(0,1)$ *and* $\nu \in (0,\infty)$, *then*

$$\frac{2^{\nu-1}\Gamma^2(\frac{\nu}{2})}{\pi\Gamma(\nu)}\int_0^1 \frac{f(t)}{\sin^{2(\nu-1)}\{\pi J_\nu(t)\}}dt = \int_0^1 \frac{f\big\{I(\frac{\nu}{2},\frac{\nu}{2},\frac{1-\cos(\pi r)}{2})\big\}}{\sin^{\nu-1}(\pi r)}dr. \tag{4.29}$$

Proof. The change of variable $t = I(\frac{\nu}{2}, \frac{\nu}{2}, \frac{1-\cos(\pi r)}{2})$ gives $J_\nu(t) = r$, and in view of (2.6),

$$dt = \frac{\Gamma(\nu)}{\Gamma^2(\frac{\nu}{2})}\{\frac{1-\cos(\pi r)}{2}\}^{\frac{\nu}{2}-1}\{\frac{1+\cos(\pi r)}{2}\}^{\frac{\nu}{2}-1}(\frac{\pi\sin(\pi r)}{2}) = \frac{\Gamma(\nu)\sin^{\nu-1}(\pi r)}{2^{\nu-1}\Gamma^2(\frac{\nu}{2})}$$

hence

$$\frac{2^{\nu-1}\Gamma^2(\frac{\nu}{2})}{\pi\Gamma(\nu)}\frac{dt}{\sin^{2(\nu-1)}\{\pi r\}} = \frac{dr}{\sin^{\nu-1}(\pi r)}$$

and the result follows immediately. □

Lemma 4.4. *For each* $\nu > 0$ *one has*

$$\int_{-\infty}^{x}\frac{dy}{\cosh^\nu y} = \frac{2^{\nu-1}\Gamma^2(\frac{\nu}{2})}{\Gamma(\nu)}I(\frac{\nu}{2}, \frac{\nu}{2}, \frac{1+\tanh x}{2}) \tag{4.30}$$

and

$$\int_{-\infty}^{\infty}\frac{dy}{\cosh^{\nu+2} y} = \nu\int_{-\infty}^{\infty}\frac{\sinh^2 y}{\cosh^{\nu+2} y}dy = \frac{\nu 2^{\nu-1}\Gamma^2(\frac{\nu}{2})}{(\nu+1)\Gamma(\nu)}. \tag{4.31}$$

Furthermore

$$\frac{\partial}{\partial\theta}I(\frac{\nu}{2}, \frac{\nu}{2}, \frac{1+\tanh(\frac{x+\theta}{\nu})}{2}) = \frac{2^{1-\nu}}{\nu}\frac{\Gamma(\nu)}{\Gamma^2(\frac{\nu}{2})}\cosh^{-\nu}(\frac{x+\theta}{\nu}). \tag{4.32}$$

Proof. The change of variables

$$t = \frac{1+\tanh y}{2}, \quad \text{hence } 4t(1-t) = \frac{1}{\cosh^2 y} \quad \text{and } 2dt = \frac{dy}{\cosh^2 y}$$

gives

$$\int_{-\infty}^{x}\frac{dy}{\cosh^\nu y} = \int_0^{\frac{1+\tanh x}{2}}\left[4t(1-t)\right]^{\frac{\nu-2}{2}}\cdot 2dt$$

which, on using (2.6), leads to (4.30). The latter, when ν is changed into $\nu+2$ and as $x \to \infty$, yields

$$\int_{-\infty}^{\infty}\frac{dy}{\cosh^{\nu+2} y} = \frac{2^{\nu+1}\Gamma^2(\frac{\nu}{2}+1)}{\Gamma(\nu+2)} = \frac{2^{\nu+1}(\frac{\nu}{2})^2\Gamma^2(\frac{\nu}{2})}{(\nu+1)\nu\Gamma(\nu+2)},$$

which proves the equality between the left-hand and the right-hand terms of (4.31). Next, formula 4.5.86 in [1] implies

$$-\frac{1}{\nu}\frac{\sinh y}{\cosh^{\nu+1} y} + \frac{1}{\nu}\int\frac{dy}{\cosh^{\nu+2} y} = \int\frac{\sinh^2 y}{\cosh^{\nu+2} y}dy$$

from which we easily deduce the equality between the first and the second terms of (4.31). Finally, on differentiating both sides of (4.30) with respect to x, one easily obtains (4.32). □

References

[1] Milton Abramowitz and Irene A. Stegun, editors. *Handbook of mathematical functions with formulas, graphs, and mathematical tables.* Dover Publications Inc., New York, 1992. Reprint of the 1972 edition.

[2] T. W. Anderson and D. A. Darling. Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *Ann. Math. Statistics*, 23:193–212, 1952.

[3] R. R. Bahadur. *Some limit theorems in statistics.* Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1971. Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, No. 4.

[4] A. Borovkov. *Statistique mathématique.* Editions Mir, Moscou, 1987.

[5] Miklós Csörgő and Lajos Horváth. *Weighted approximations in probability and statistics.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Ltd., Chichester, 1993. With a foreword by David Kendall.

[6] Paul Deheuvels and Guennady V. Martynov. A Karhunen-Loève decomposition of a Gaussian process generated by independent pairs of exponential random variables. Preprint, 2003.

[7] N. Henze and Ya. Yu. Nikitin. A new approach to goodness-of-fit testing based on the integrated empirical process. *J. Nonparametr. Statist.*, 12(3):391–416, 2000.

[8] Norman L. Johnson, Samuel Kotz, and N. Balakrishnan. *Continuous univariate distributions. Vol. 2.* Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons Inc., New York, 1995.

[9] Yakov Nikitin. *Asymptotic efficiency of nonparametric tests.* Cambridge University Press, Cambridge, 1995.

[10] J.-R. Pycke. A family of Karhunen-Loève expansions related to the n-dimensional Euclidean sphere and real projective space. Preprint, 2003.

[11] Louis Robin. *Fonctions sphériques de Legendre et fonctions sphéroïdales. Tome II.* Gauthier-Villars, Paris, 1958.

[12] Galen R. Shorack and Jon A. Wellner. *Empirical processes with applications to statistics.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986.

[13] V. M. Zolotarev. Concerning a certain probability problem. *Teor. Verojatnost. i Primenen.*, 6:219–222, 1961.