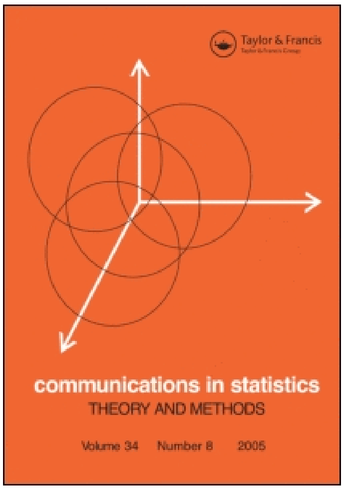

Communications in Statistics - Theory and Methods
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597238

Karhunen-Loève Basis in Goodness-of-Fit Tests Decomposition: An Evaluation

Aurea Grané [a]; Josep Fortiana [b]
[a] Statistics Department, Universidad Carlos III de Madrid, Madrid, Spain [b] Departament de Probabilitats, Lògica i Estadística, Universitat de Barcelona, Barcelona, Spain

Online Publication Date: 01 January 2008

To cite this Article Grané, Aurea and Fortiana, Josep(2008)'Karhunen-Loève Basis in Goodness-of-Fit Tests Decomposition: An Evaluation',Communications in Statistics - Theory and Methods,37:19,3144 — 3163
To link to this Article: DOI: 10.1080/03610920802065099
URL: http://dx.doi.org/10.1080/03610920802065099

Communications in Statistics—Theory and Methods, 37: 3144–3163, 2008

ISSN: 0361-0926 print/1532-415X online
DOI: 10.1080/03610920802065099

Goodness-of-Fit Tests

Karhunen–Loève Basis in Goodness-of-Fit Tests Decomposition: An Evaluation

AUREA GRANÉ[1] AND JOSEP FORTIANA[2]

[1]Statistics Department, Universidad Carlos III de Madrid, Madrid, Spain
[2]Departament de Probabilitats, Lògica i Estadística, Universitat de Barcelona, Barcelona, Spain

In a previous article (Grané and Fortiana, 2006), we studied a flexible class of goodness-of-fit tests associated with an orthogonal sequence, the Karhunen–Loève decomposition of a stochastic process derived from the null hypothesis. Generally speaking, these tests outperform Kolmogorov–Smirnov and Cramér–von Mises, but we registered several exceptions. In this work we investigate the cause of these anomalies and, more precisely, whether and when such poor behavior may be attributed to the orthogonal sequence itself, by replacing it with the Legendre polynomials, a commonly used basis for smooth tests. We find an easily computable formula for the Bahadur asymptotic relative efficiency, a helpful quantity in choosing an adequate basis.

Keywords Asymptotic relative efficiency; Goodness-of-fit; Orthonormal functions; Smooth tests.

Mathematics Subject Classification 62G10; 62G30; 62G20.

1. Introduction

In Fortiana and Grané (2003), we defined a sequence of statistics, $\{\beta_{nj}\}_{j\in\mathbb{N}}$, based on Hoeffding's maximum correlation. This quantity, $\rho^+(F_1, F_2)$, for two univariate probability distributions F_1 and F_2, is defined as the maximum of the correlation coefficients of all bivariant probability distributions having marginals F_1 and F_2. It is a measure of proximity between both marginals and, when applied to an empirical and a theoretical distribution, yields a goodness-of-fit test.

The sequence $\{\beta_{nj}\}_{j\in\mathbb{N}}$ appears when this test is decomposed along orthogonal axes, a construction analogous to that of the Cramér–von Mises statistic (Durbin and Knott, 1972, 1975), studied in a general setting by Stephens (1974). More precisely, let F_n be the empirical cumulative distribution function (cdf) of n iid

Received January 11, 2008; Accepted March 14, 2008
Address correspondence to Aurea Grané, Statistics Department, Universidad Carlos III de Madrid, c/ Madrid, 126, 28903 Getafe, Spain; E-mail: aurea.grane@uc3m.es

random variables and let F_n^- be its pseudo-inverse, then β_{nj} is the jth Fourier coefficient of F_n^- for the orthonormal (in $L^2[0, 1]$) sequence $\beta \equiv \{\beta_j(t)\}_{j\in\mathbb{N}}$, $t \in [0, 1]$, obtained from the eigenfunctions of the covariance kernel of a certain Bernoulli stochastic process associated with the [0, 1] uniform distribution (see Cuadras and Fortiana, 1993, 1995; Fortiana and Grané, 2003 for details). Henceforth, we will refer to this particular sequence of statistics as the Karhunen–Loève (KL) sequence. In Grané and Fortiana (2006), we studied a class of statistics, linear combinations of $\{\beta_{nj}\}_{j\geq 0}$ (where $\beta_{n0} \equiv 1$), with adjustable coefficients depending on the alternative distribution or family of distributions. We found that their power properties were remarkably good, but for several alternatives their behavior was rather poor.

In this article, we substitute $\phi \equiv \{\phi_j(t)\}_{j\geq 0}$, an orthonormal sequence in $L^2[0, 1]$, for β yielding the sequence $\{\Phi_{nj}\}_{j\geq 0}$ of statistics, as defined in Sec. 2. Sections 3 and 4 are parallel to the corresponding ones in Grané and Fortiana (2006), with the obvious modifications: power optimization is translated into an eigenvalue-type problem with quadratic forms, functions of the first two moments of the order statistic. Some simplifications of the KL case are not possible, however. As an illustration, we perform the actual computations for ϕ being the Legendre polynomials, comparing the power of the statistic obtained with this basis with that of the KL one. In Sec. 5, we find an easy computable formula for the Bahadur approximate slope, and we use the Bahadur asymptotic relative efficiency as a criterion to select a basis. Section 6 contains some practical issues.

2. Definition of the Statistic

Let F be a probability cdf with finite second order moment and let F_n be the empirical distribution function of n iid $\sim F$ random variables. Given an orthonormal sequence, $\phi \equiv \{\phi_j(t)\}_{j\geq 0}$, in $L^2[0, 1]$, we define

$$\Phi_{nj} \equiv \Phi_{nj}(F) = \int_0^1 F_n^-(t)\phi_j(t)dt, \quad j \geq 0,$$

where F_n^- is the pseudo-inverse of F_n. They are L-statistics, i.e., linear combinations of the order statistic $\mathbf{x} \equiv (x_{(1)}, \ldots, x_{(n)})$, since

$$\Phi_{nj} = \int_0^1 F_n^-(t)\phi_j(t)dt = \sum_{i=1}^n \int_{(i-1)/n}^{i/n} x_{(i)}\phi_j(t)dt = \sum_{i=1}^n a_{ij}x_{(i)},$$

where $a_{ij} = \int_{(i-1)/n}^{i/n} \phi_j(t)dt$. We consider the class of all linear combinations

$$T \equiv T(\lambda_0, \ldots, \lambda_p) = \sum_{j=0}^p \lambda_j \Phi_{nj}, \tag{1}$$

where $\lambda_0, \ldots, \lambda_p$ are real parameters. They are L-statistics, too,

$$T = \sum_{i=1}^n c_{ni}x_{(i)},$$

with coefficients

$$c_{ni} = \sum_{j=0}^{p} \lambda_j a_{ji}.$$

In matrix notation,

$$T \equiv T(\boldsymbol{\lambda}) = \mathbf{x}\mathbf{A}\boldsymbol{\lambda}, \tag{2}$$

where $\boldsymbol{\lambda} = (\lambda_0, \ldots, \lambda_p)'$, $\mathbf{A} = (a_{ij})$, $1 \le i \le n$, $0 \le j \le p$.

Given an alternative cdf F_1, we select $\boldsymbol{\lambda}$ to maximize power for testing $H_0 : F = F_U$, vs. $H_1 : F = F_1$, where F_U is the cdf of a [0, 1] uniform random variable. Clearly, the resulting test is less powerful than the optimal (Neyman–Pearson) one, but its distribution under the null hypothesis is easily computed, both for large samples, applying the asymptotic theory of L-statistics, and for small samples, with the exact distribution, as described in Fortiana and Grané (2003).

3. Computation and Optimization of the Power Function

It is possible to use a one-sided test for a fixed alternative F_1, since it may be proved that when H_1 is true T tends to its upper tail if $var(F_1) > 1/12$ and to its lower tail otherwise. In general, however, for a family of alternatives we must consider the two-sided test. Henceforth, this will be our assumption. To test $H_0 : F = F_U$ against $H_1 : F = F_1$, a known cdf with support contained in [0, 1], we consider (1) where $\boldsymbol{\lambda}$ is to be determined. Its asymptotic distribution is normal, from the general theory of L-statistics (see, e.g., Stigler, 1974, or Chap. 19 of Shorack and Wellner, 1986). For a fixed significance level $\varepsilon \in (0, 1)$, we are looking for $c_1, c_2 \in \mathbb{R}$, such that:

$$P(T > c_1 \mid H_0) = \varepsilon/2, \quad P(T < c_2 \mid H_0) = \varepsilon/2.$$

We take c_1, c_2 symmetric with respect to $\mu_0 = E(T \mid H_0)$, that is, $c_1 = \mu_0 + c_{\varepsilon/2}\sigma_0$, $c_2 = \mu_0 - c_{\varepsilon/2}\sigma_0$, where $\sigma_0^2 = var(T_p \mid H_0)$ and $c_{\varepsilon/2}$ is the $(1-\varepsilon/2)\cdot 100$-percentile of the $N(0, 1)$ distribution. The power function $P(T > c_1 \mid H_1) + P(T < c_2|H_1)$ is asymptotically approximated by:

$$\Psi(\boldsymbol{\lambda}) = 1 - P_Z\left[\left(\frac{\mu_0 - \mu_1}{\sigma_1} - c_{\varepsilon/2}\frac{\sigma_0}{\sigma_1}, \frac{\mu_0 - \mu_1}{\sigma_1} + c_{\varepsilon/2}\frac{\sigma_0}{\sigma_1}\right)\right],$$

where $\mu_1 = E(T \mid H_1)$, $\sigma_1^2 = var(T \mid H_1)$, and $Z \sim N(0, 1)$. Due to the symmetry of this distribution, $\mu_0 - \mu_1$ can be replaced by $|\mu_0 - \mu_1|$, and then:

$$\Psi(\boldsymbol{\lambda}) = 1 - P_Z\left\{\left(\left[\frac{a(\boldsymbol{\lambda})}{c(\boldsymbol{\lambda})}\right]^{1/2} - \left[\frac{b(\boldsymbol{\lambda})}{c(\boldsymbol{\lambda})}\right]^{1/2}, \left[\frac{a(\boldsymbol{\lambda})}{c(\boldsymbol{\lambda})}\right]^{1/2} + \left[\frac{b(\boldsymbol{\lambda})}{c(\boldsymbol{\lambda})}\right]^{1/2}\right)\right\},$$

in terms of the following quadratic forms:

$$\begin{aligned} a(\boldsymbol{\lambda}) &= (\mu_0 - \mu_1)^2 = \boldsymbol{\lambda}'\mathbf{A}'(\mathbf{M}_0 - \mathbf{M}_1)'(\mathbf{M}_0 - \mathbf{M}_1)\mathbf{A}\boldsymbol{\lambda}, \\ b(\boldsymbol{\lambda}) &= c_{\varepsilon/2}^2\sigma_0^2 = \boldsymbol{\lambda}'\mathbf{A}'\boldsymbol{\Sigma}_0\mathbf{A}\boldsymbol{\lambda}, \\ c(\boldsymbol{\lambda}) &= \sigma_1^2 = \boldsymbol{\lambda}'\mathbf{A}'\boldsymbol{\Sigma}_1\mathbf{A}\boldsymbol{\lambda}, \end{aligned} \tag{3}$$

where $\mathbf{M}_i = E(\mathbf{x} \mid H_i)$, $\boldsymbol{\Sigma}_0 = c_{\varepsilon/2}^2 Var(\mathbf{x} \mid H_0)$, $\boldsymbol{\Sigma}_1 = Var(\mathbf{x} \mid H_1)$.

Since $\Psi(\boldsymbol{\lambda})$ remains invariant when $\boldsymbol{\lambda}$ is multiplied by an arbitrary constant, we assume $c(\boldsymbol{\lambda}) = 1$, and we compute the extremes of:

$$\Upsilon(\boldsymbol{\lambda}) = 1 - \Phi(a(\boldsymbol{\lambda})^{1/2} + b(\boldsymbol{\lambda})^{1/2}) + \Phi(a(\boldsymbol{\lambda})^{1/2} - b(\boldsymbol{\lambda})^{1/2}) + \xi(c(\boldsymbol{\lambda}) - 1), \tag{4}$$

where Φ is the standard normal distribution function and ξ is a Lagrange multiplier.

Degenerate Case. If $a(\boldsymbol{\lambda}) = 0$, the expectation of T is the same under both hypotheses, the power function is:

$$\Psi(\boldsymbol{\lambda}) = 1 - P_Z\left\{\left(-\left[\frac{b(\boldsymbol{\lambda})}{c(\boldsymbol{\lambda})}\right]^{1/2}, \left[\frac{b(\boldsymbol{\lambda})}{c(\boldsymbol{\lambda})}\right]^{1/2}\right)\right\},$$

with the constraint $c(\boldsymbol{\lambda}) = 1$, and (4) is written as:

$$\Upsilon(\boldsymbol{\lambda}) = 2 - 2\Phi\big(b(\boldsymbol{\lambda})^{1/2}\big) + \xi(c(\boldsymbol{\lambda}) - 1).$$

Equating to zero its gradient, we obtain an eigenvalue-type problem,

$$\beta(\boldsymbol{\lambda})\mathbf{A}'\boldsymbol{\Sigma}_0\mathbf{A}\boldsymbol{\lambda} = \xi\mathbf{A}'\boldsymbol{\Sigma}_1\mathbf{A}\boldsymbol{\lambda},$$

where $\beta(\boldsymbol{\lambda}) = 2b(\boldsymbol{\lambda})^{-1/2}\phi(b(\boldsymbol{\lambda})^{1/2})$, and ϕ is the standard normal probability density function (pdf).

General Case. If $a(\boldsymbol{\lambda}) \neq 0$, differentiating (4) and equating to zero we obtain:

$$\Big[\alpha(\boldsymbol{\lambda})\mathbf{A}'(\mathbf{M}_0 - \mathbf{M}_1)'(\mathbf{M}_0 - \mathbf{M}_1)\mathbf{A} + \beta(\boldsymbol{\lambda})\mathbf{A}'\boldsymbol{\Sigma}_0\mathbf{A}\Big]\boldsymbol{\lambda} = \xi\mathbf{A}'\boldsymbol{\Sigma}_1\mathbf{A}\boldsymbol{\lambda}, \tag{5}$$

where $\alpha(\boldsymbol{\lambda}) = a(\boldsymbol{\lambda})^{-1/2}(\phi_+(\boldsymbol{\lambda}) - \phi_-(\boldsymbol{\lambda}))$, $\beta(\boldsymbol{\lambda}) = b(\boldsymbol{\lambda})^{-1/2}(\phi_+(\boldsymbol{\lambda}) + \phi_-(\boldsymbol{\lambda}))$, $\phi_+(\boldsymbol{\lambda}) = \phi(a(\boldsymbol{\lambda})^{1/2} + b(\boldsymbol{\lambda})^{1/2})$, $\phi_-(\boldsymbol{\lambda}) = \phi(a(\boldsymbol{\lambda})^{1/2} - b(\boldsymbol{\lambda})^{1/2})$, and ξ has been redefined. The degenerate case appears when $\alpha(\boldsymbol{\lambda}) = 0$.

To compute $\boldsymbol{\lambda}$ set $u = (\mathbf{A}'\boldsymbol{\Sigma}_1\mathbf{A})^{1/2}\boldsymbol{\lambda}$, $\mathbf{G}(u) = \alpha(u)\mathbf{E} + \beta(u)\mathbf{F}$, where $\alpha(u)$, $\beta(u)$ are those defined above in (5), now in terms of the new variable u, and

$$\mathbf{E} = (\mathbf{A}'\boldsymbol{\Sigma}_1\mathbf{A})^{-1/2}\mathbf{A}'(\mathbf{M}_0 - \mathbf{M}_1)'(\mathbf{M}_0 - \mathbf{M}_1)\mathbf{A}(\mathbf{A}'\boldsymbol{\Sigma}_1\mathbf{A})^{-1/2},$$
$$\mathbf{F} = (\mathbf{A}'\boldsymbol{\Sigma}_1\mathbf{A})^{-1/2}(\mathbf{A}'\boldsymbol{\Sigma}_0\mathbf{A})(\mathbf{A}'\boldsymbol{\Sigma}_1\mathbf{A})^{-1/2}.$$

For a given u, we compute eigenvectors and eigenvalues of $\mathbf{G}(u)$. The new u will be the eigenvector for which $\Psi(u)$ is maximum. This process is iterated until stability. The last step is to recover and normalize $\boldsymbol{\lambda}$. The result is rather robust, leading to a single maximum with a small number of iterations for a widely diverse choice of the initial u. A Matlab program implementing this computation may be requested from the authors.

Example 3.1 (Scale Alternatives)**.** We consider an alternative distribution belonging to $U[0, \theta]$, the uniform on $[0, \theta]$ family, with $\theta > 0$. The expectation vector $\mathbf{M}_0$ and the covariance matrix $\boldsymbol{\Sigma}_0$ of the order statistic $\mathbf{x}$ obtained from n iid $\sim U[0, 1]$ random variables are (see, e.g., David, 1981):

$$\mathbf{M}_0 = \frac{1}{n+1}(1, 2, \ldots, n), \quad \boldsymbol{\Sigma}_0 = \left(v_{ij}\right)_{1\le i,j\le n}, \tag{6}$$

where

$$v_{ij} = \frac{1}{(n+2)(n+1)^2}[(n+1)\min\{i, j\} - ij].$$

The corresponding quantities for $U[0, \theta]$ are $\mathbf{M}_1 = \theta\mathbf{M}_0$, $\boldsymbol{\Sigma}_1 = \theta^2\boldsymbol{\Sigma}_0$. Then (3) is:

$$\begin{aligned} a(\boldsymbol{\lambda}) &= (1-\theta)^2\boldsymbol{\lambda}'\mathbf{A}'\mathbf{M}_0'\mathbf{M}_0\mathbf{A}\boldsymbol{\lambda}, \\ b(\boldsymbol{\lambda}) &= c_{\varepsilon/2}^2\boldsymbol{\lambda}'\mathbf{A}'\boldsymbol{\Sigma}_0\mathbf{A}\boldsymbol{\lambda}, \\ c(\boldsymbol{\lambda}) &= \theta^2\boldsymbol{\lambda}'\mathbf{A}'\boldsymbol{\Sigma}_0\mathbf{A}\boldsymbol{\lambda}. \end{aligned}$$

We have to maximize the power $\Psi(\boldsymbol{\lambda})$, equivalently, to minimize

$$P_Z\left(\frac{1-\theta}{\theta}\left(\frac{\boldsymbol{\lambda}'\mathbf{A}'\mathbf{M}_0'\mathbf{M}_0\mathbf{A}\boldsymbol{\lambda}}{\boldsymbol{\lambda}'\mathbf{A}'\boldsymbol{\Sigma}_0\mathbf{A}\boldsymbol{\lambda}}\right)^{1/2} - \frac{c_{\varepsilon/2}}{\theta}, \frac{1-\theta}{\theta}\left(\frac{\boldsymbol{\lambda}'\mathbf{A}'\mathbf{M}_0'\mathbf{M}_0\mathbf{A}\boldsymbol{\lambda}}{\boldsymbol{\lambda}'\mathbf{A}'\boldsymbol{\Sigma}_0\mathbf{A}\boldsymbol{\lambda}}\right)^{1/2} + \frac{c_{\varepsilon/2}}{\theta}\right),$$

or to maximize $\boldsymbol{\lambda}'\mathbf{A}'\mathbf{M}_0'\mathbf{M}_0\mathbf{A}\boldsymbol{\lambda}/\boldsymbol{\lambda}'\mathbf{A}'\boldsymbol{\Sigma}_0\mathbf{A}\boldsymbol{\lambda}$, both problems constrained to $\boldsymbol{\lambda}'\mathbf{A}'\boldsymbol{\Sigma}_0\mathbf{A}\boldsymbol{\lambda} = 1$. The solution is the (unique with non null eigenvalue) eigenvector of the generalized eigenvalue problem: $\mathbf{A}'\mathbf{M}_0'\mathbf{M}_0\mathbf{A}\boldsymbol{\lambda} = \xi\mathbf{A}'\boldsymbol{\Sigma}_0\mathbf{A}\boldsymbol{\lambda}$, normalized so that $\boldsymbol{\lambda}'\mathbf{A}'\boldsymbol{\Sigma}_0\mathbf{A}\boldsymbol{\lambda} = 1$. Note that $\boldsymbol{\lambda}$ does not depend on the scale parameter θ.

In Grané and Fortiana (2006), we used the orthonormal basis β introduced in Sec. 1. Explicitly:

$$\beta_0(t) = 1, \quad \beta_j(t) = \sqrt{2}\cos(j\pi t), \;\; j \ge 1, \;\; t \in (0, 1),$$

and we denoted by β_{nj} the resulting Φ_{nj} statistics. In this basis, formula (1) is $T_\beta = \sum_{j=0}^{p} \lambda_j\beta_{nj}$. For a sample of size $n = 20$, a significance level $\varepsilon = 0.05$ and $p = 4$, we obtain:

$$T_\beta = 0.3554\beta_{n0} - 0.4447\beta_{n1} + 0.4985\beta_{n2} - 0.4373\beta_{n3} + 0.4860\beta_{n4}.$$

In the context of *smooth-tests* (see, e.g., Rayner and Best, 1989, 1990), the sequence of Legendre polynomials is often used. After adapting them to the [0, 1] interval and standardizing them, the first two of them are:

$$\phi_0(t) = 1, \quad \phi_1(t) = \sqrt{3}(2t-1),$$

and the recurrence relation is:

$$\phi_{j+1}(t) = \frac{\sqrt{(2j+3)(2j+1)}}{j+1}(2t-1)\phi_j(t) - \frac{\sqrt{2j+3}}{\sqrt{2j-1}}\frac{j}{j+1}\phi_{j-1}(t), \quad j \ge 1.$$

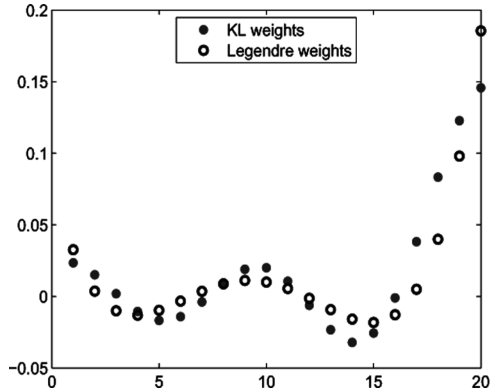

Figure 1. Comparison of the vectors of weights of $T(\lambda)$ for the observed order statistic ($n = 20$) for the scale alternatives.

Denoting by ℓ_{nj} the resulting Φ_{nj} statistics, formula (1) is $T_\ell = \sum_{j=0}^{p} \lambda_{nj}\ell_{nj}$. For a sample of size $n = 20$, a significance level of $\varepsilon = 0.05$ and $p = 4$, we obtain:

$$T_\ell = 0.3095\ell_{n0} + 0.4403\ell_{n1} + 0.5786\ell_{n2} + 0.4193\ell_{n3} + 0.4470\ell_{n4}.$$

In a practical situation, T_β and T_ℓ should be expressed directly in terms of the observed order statistic using (2). Figure 1 shows a comparison between the different vectors of weights of the order statistic, obtained applying formula (2) for T_β and T_ℓ.

We have compared T_β and T_ℓ with the Q_n statistic obtained in Fortiana and Grané (2003), with the Kolmogorov–Smirnov statistic D_n and with the Cramér–von Mises statistic W_n^2. Figure 2 shows the power curves at a 5% significance level for the tests based on these statistics. These curves have been plotted from 20 computed points, for each of which we have generated $N = 1{,}000$ samples of size $n = 20$.

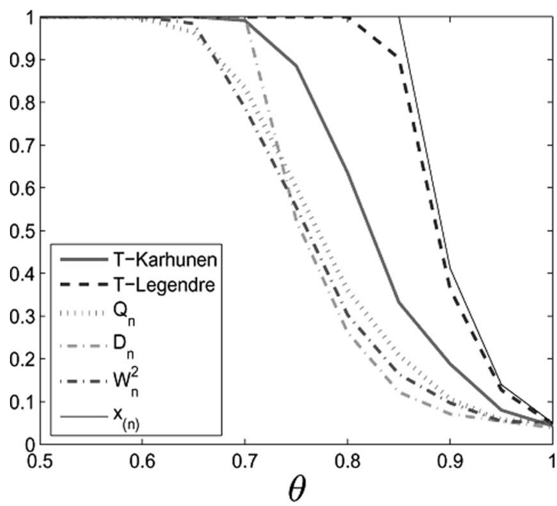

Figure 2. Power functions for scale alternatives.

We have also considered the natural test for testing uniformity on $[0, \theta]$, for $0 < \theta < 1$, based on the largest order statistic, $x_{(n)}$. Note that the performance of the T_ℓ statistic is rather similar to that of $x_{(n)}$.

4. Generic Alternatives

In this section, we develop an algorithm for locating the optimal $\boldsymbol{\lambda}$ in (2) for an alternative cdf F whose pseudoinverse has the form:

$$F^{-}(t) = \sum_{k=0}^{q} \gamma_k \psi_k(t), \tag{7}$$

where γ_k are real numbers and $\{\psi_k(t)\}_{k\geq 0}$ is an orthonormal sequence in $L^2[0, 1]$, possibly different from $\{\phi_j(t)\}_{j\geq 0}$.

Given an arbitrary F, the first q Fourier terms of F^- yield such an expression. In the present context, this is more natural than expanding F or the pdf, since the moments of the order statistics can be advantageously expressed in terms of F^-, e.g.,

$$\begin{aligned} E(x_{(i)}|H_1) &= i\binom{n}{i}\int_0^1 F^{-}(t)t^{i-1}(1-t)^{n-i}\,dt \\ &= i\binom{n}{i}\sum_{k=0}^{q}\gamma_k\int_0^1 \psi_k(t)t^{i-1}(1-t)^{n-i}\,dt. \end{aligned} \tag{8}$$

To solve (5) we must determine the quadratic forms $a(\boldsymbol{\lambda})$, $b(\boldsymbol{\lambda})$, $c(\boldsymbol{\lambda})$ in (3). The expectation vector and the covariance matrix of the order statistic under H_0 are given by (6) and (8) gives the entries in $\mathbf{M}_1$. In general, an exact $\boldsymbol{\Sigma}_1$ will not be available. Instead, we can determine $\mathbf{A}'\boldsymbol{\Sigma}_1\mathbf{A}$ from the asymptotic approximation given in the following proposition.

Proposition 4.1. *Let T be the statistic defined in* (1) *and* (2), *where* $\mathbf{x}$ *is the order statistic from* n iid *random variables with cdf* (7). *We have the following convergences in law*:

$$\sqrt{n}[T-\mu] \xrightarrow[n\to\infty]{\mathscr{L}} N(0, \sigma_1^2), \tag{9}$$

$$\sqrt{n}\frac{[T-\mu]}{\sigma_n} \xrightarrow[n\to\infty]{\mathscr{L}} N(0, 1), \tag{10}$$

where

$$\mu = \sum_{j=0}^{p}\sum_{k=0}^{q}\lambda_j\gamma_k\int_0^1 \phi_j(t)\psi_k(t)dt, \tag{11}$$

$$\sigma_1^2 = \lim_{n\to\infty}\sigma_n^2, \quad \sigma_n^2 = \sum_{j=0}^{p}\sum_{l=0}^{p}\lambda_j\lambda_l\sigma_{n,jl}, \tag{12}$$

$$\sigma_{n,jl} = \sum_{k=0}^{q}\sum_{m=0}^{q}\gamma_k\gamma_m I_{jklm},$$

where

$$I_{jklm} = \int_0^1 \int_0^1 K(s,t)\phi_j(s)\psi'_k(s)\phi_l(t)\psi'_m(t)dt\,ds,$$

where $K(s,t) = \min(s,t) - st$ *and* $\psi'_k(t)$ *denotes the derivative of* $\psi_k(t)$.

Proof. The T statistic of (1) can be written as:

$$T = \frac{1}{n}\sum_{i=1}^{p} J(i/n)x_{(i)}, \tag{13}$$

where

$$J(i/n) = \sum_{j=0}^{n} n\lambda_j a_{ij},$$

$$a_{ij} = \int_{(i-1)/n}^{i/n} \phi_j(u)du = b_j(i/n) - b_j((i-1)/n),$$

and $b_j(i/n) = \int_0^{i/n} \phi_j(u)du$. Using these expressions the $J(i/n)$ coefficients are:

$$\sum_{j=0}^{p} n\lambda_j a_{ij} = \sum_{j=0}^{p} \lambda_j \frac{b_j(i/n) - b_j((i-1)/n)}{1/n} = \sum_{j=0}^{p} \lambda_j B_j(i/n),$$

where

$$B_j(i/n) = \frac{b_j(i/n) - b_j((i-1)/n)}{1/n},$$

verifying that B_j tends to ϕ_j when n tends to infinity. We can use the asymptotic approximation

$$J(t) \approx \sum_{j=0}^{p} \lambda_j \phi_j(t), \quad t \in (0,1).$$

Since $J(t)$ is a continuous and bounded a.s. (F^-) function, we can compute the asymptotic expectation of T, under H_1, as:

$$\mu = \int_0^1 J(t)F^-(t)dt = \sum_{j=0}^{p}\sum_{k=0}^{q} \lambda_j \gamma_k \int_0^1 \phi_j(t)\psi_k(t)dt$$

and also its asymptotic variance as:

$$\sigma_1^2 = \int_0^1 \int_0^1 J(s)J(t)K(s,t)dF^-(s)dF^-(t),$$

where $K(s,t) = \min(s,t) - st$, see, e.g., Shorack and Wellner (1986). Substituting the expressions for function J and for the derivative of F^- formulas (11) and (12) are obtained.

The convergence of (9) and (10) are obtained applying the general theory for L-statistics described in Shorack and Wellner (1986). □

These expressions can be simplified when both $\{\phi_j(t)\}_{j\geq 0}$ and $\{\psi_k(t)\}_{k\geq 0}$ are the KL (trigonometric) basis $\{1, \sqrt{2}\cos(j\pi t)\}_{j\geq 1}$. This is due to the fact that $\phi_j(t)$ and $\psi'_k(t)$ in I_{jklm} can be expressed in terms of eigenfunctions of $K(s, t)$. In this case, the expression of σ_n^2 is:

$$\sigma_n^2 = \sum_{j=1}^{p}\sum_{l=1}^{p} \lambda_j \lambda_l \sigma_{n,jl},$$

$$\sigma_{n,jl} = 4\pi^2 a_{nj} a_{nl} \sum_{k=1}^{q}\sum_{m=1}^{q} km\gamma_k\gamma_m I_{jklm},$$

where $a_{nj} = -\sqrt{2}(2n/(j\pi))\sin(j\pi/(2n))$,

$$I_{jklm} = \frac{1}{(4\pi)^2}\left\{\frac{1}{(k+j)^2}\left[\delta_{m-l,k+j} + \delta_{m+l,k+j}\right]\right\}, \quad \text{if } k = j,$$

$$I_{jklm} = \frac{1}{(4\pi)^2}\left\{\frac{\delta_{m-l,k-j} + \delta_{m+l,k-j}}{(k-j)^2} + \frac{\delta_{m-l,k+j} + \delta_{m+l,k+j}}{(k+j)^2}\right\},$$

if $k \neq j$, and δ is Kronecker's delta. For a complete proof see Grané and Fortiana (2006).

Comparing the expression for $c(\boldsymbol{\lambda}) = \sigma_1^2 = \boldsymbol{\lambda}'\mathbf{A}'\boldsymbol{\Sigma}_1\mathbf{A}\boldsymbol{\lambda}$ in (3) with (12), we see that the entries in $\mathbf{A}\boldsymbol{\Sigma}_1\mathbf{A}'$ are the limit $\sigma_{jl} = \lim_{n\to\infty}\sigma_{n,jl}$, but some computational examples suggest that a better approximation is obtained with $\sigma_{n,jl}$.

4.1. *Some Examples*

To illustrate the method we have chosen four parametric families of alternative distributions with support on [0, 1]. We have chosen them so that either the mean or the variance differs from those of the null hypothesis, $U[0, 1]$, which in each case is obtained for a value of the parameter. They are defined by the following probability distribution functions:

A1. Lehmann alternatives,

$$F_\alpha(x) = x^\alpha, \quad 0 \leq x \leq 1, \ \alpha > 0;$$

A2. symmetric (with respect to 1/2) distributions having U-shaped pdf, for $\beta \in (0, 1)$, or wedge-shaped pdf, for $\beta > 1$,

$$F_\beta(x) = \begin{cases} \frac{1}{2}(2x)^\beta, & 0 \leq x \leq 1/2, \\ 1 - \frac{1}{2}(2(1-x))^\beta, & 1/2 \leq x \leq 1; \end{cases}$$

A3. compressed uniform alternatives,

$$F_\gamma(x) = \begin{cases} 0, & 0 \le x \le \gamma, \\ \dfrac{x-\gamma}{1-2\gamma}, & \gamma \le x \le 1-\gamma, \quad 0 \le \gamma \le \dfrac{1}{2}, \\ 1, & 1-\gamma \le x \le 1; \end{cases}$$

Table 1
Power comparisons for A1–A4 families

A1 family						
α	T_ℓ	T_β	Q_n	D_n	W_n^2	UMP
0.25	0.9962	0.9980	0.4411	0.9970	0.9973	1
0.5	0.7615	0.7492	0.1203	0.6550	0.7211	0.9259
0.75	0.2096	0.1973	0.0764	0.1830	0.1987	0.3918
2	0.8792	0.8347	0.3984	0.6730	0.7708	0.9185
3	0.9982	0.9872	0.8779	0.9910	0.9955	0.9998
4	1.0000	0.9991	0.9900	1.0000	1.0000	1
A2 family						
β	T_ℓ	T_β	Q_n	D_n	W_n^2	UMP*
0.25	0.9678	0.9447	0.9651	0.8071	0.8597	1
0.5	0.6600	0.6524	0.7238	0.2879	0.2840	0.9238
0.75	0.1827	0.1893	0.2203	0.0916	0.0905	0.3973
2	0.8252	0.8045	0.7523	0.1288	0.1013	0.9252
3	0.9979	0.9929	0.9955	0.4029	0.5107	0.9997
4	1.0000	1.0000	1.0000	0.7361	0.8951	1
A3 family						
γ	T_ℓ	T_β	Q_n	D_n	W_n^2	
0.05	0.9639	0.9619	0.9585	1.0000	1.0000	
0.15	0.9130	0.8905	0.9309	1.0000	1.0000	
0.25	0.7749	0.7951	0.7736	0.8817	0.7533	
0.35	0.4351	0.3934	0.3097	0.3321	0.1964	
0.45	0.1257	0.0745	0.0697	0.0931	0.0752	
A4 family						
δ	T_ℓ	T_β	Q_n	D_n	W_n^2	
0.05	0.9639	0.9619	0.9585	1.0000	1.0000	
0.15	0.9130	0.8905	0.9309	1.0000	1.0000	
0.25	0.7749	0.7951	0.7736	0.8817	0.7533	
0.35	0.4351	0.3934	0.3097	0.3321	0.1964	
0.45	0.1257	0.0745	0.0697	0.0931	0.0752	

*Critical values have been obtained via Monte Carlo.

A4. a bimodal locally uniform distribution, with probability mass concentrated near both extremes, 0 and 1,

$$F_\delta(x) = \begin{cases} x/(2\delta), & 0 \le x \le \delta, \\ \frac{1}{2}, & \delta \le x \le 1-\delta, \quad 0 < \delta \le 1/2, \\ 1-(1-x)/(2\delta), & 1-\delta \le x \le 1. \end{cases}$$

As examples of construction of the test for generic alternatives, we have considered the families above for several values of the parameters. For each alternative we computed coefficients γ_k of (7), for $0 \le k \le q = 5$. For sample size $n = 20$ and significance level $\varepsilon = 0.05$, we have determined the vector $\boldsymbol{\lambda}$ of coefficients for $p = 4$ for the adaptive statistics T_β (KL test statistic) and T_ℓ (with both $\{\phi_j(t)\}_{j\ge0}$ and $\{\psi_k(t)\}_{k\ge0}$ the Legendre polynomials). Note that the vector $\boldsymbol{\lambda}$ of coefficients depends on the parameter values. We have prepared a set of Matlab programs implementing these tests for A1–A4 families of alternatives, which can be easily extended for any alternative of the form (7). Table 1 contains the power comparisons of the test based on T_ℓ with the tests based on T_β, Q_n, D_n, and W_n^2.

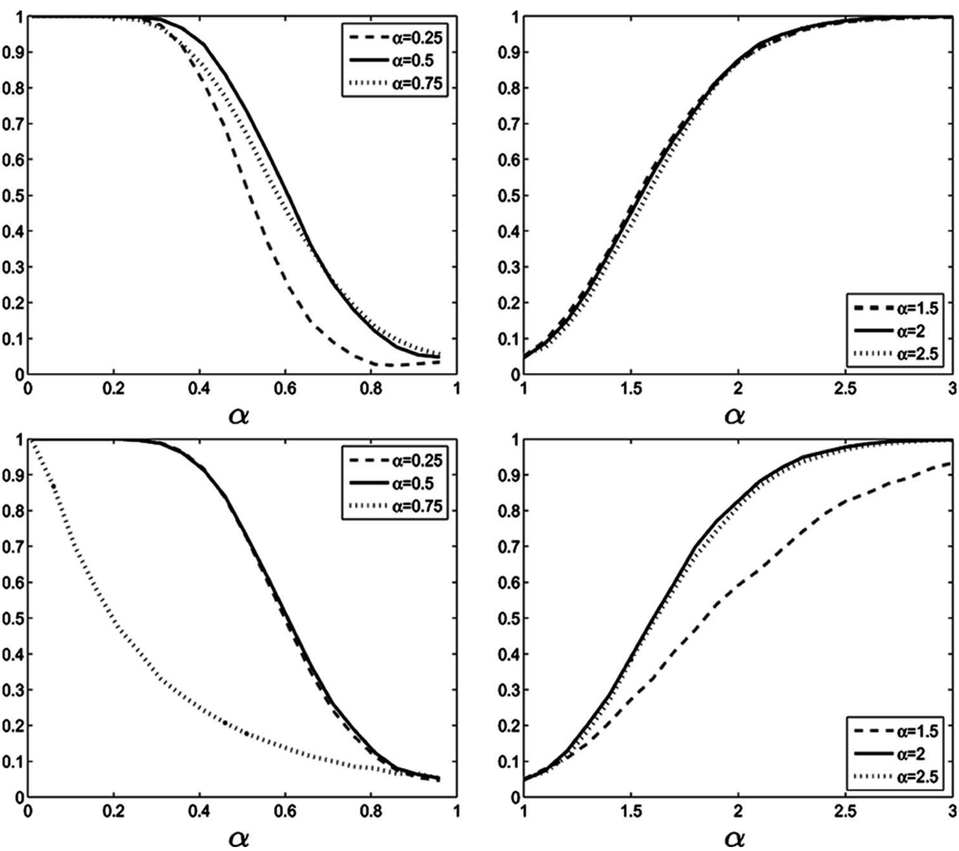

Figure 3. Power of the test based on T_ℓ (top) and T_β (bottom) for the A1 family.

These powers have been estimated from $N = 10,000$ samples of size $n = 20$ as the relative frequency of values of the statistic in the critical region. Since the UMP test is easy to compute for the A1 family and for the A2 family, the critical value of the UMP test can be reasonably easy obtained via a Monte Carlo study, we have included these results in Table 1 for comparison.

In order to compare the performance of the two orthonormal systems in detecting alternatives coming from A1–A4 families, we have fixed a particular set of values of the parameter (six values for A1 and A2 families, three values for A3 family, and four values for A4 family), we have constructed the adaptive statistics for these particular values and we have computed the corresponding power functions. These power functions can be used to construct an *envelop curve* which can be used as a guide function for power comparisons. Figures 3–5 contain these power functions. It should be noted the bad performance of the KL system when the values of the parameter are close to the null hypothesis. In these cases the statistic constructed is under the expectative, as was pointed out in Grané and Fortiana (2006), and the T_β statistic only succeeds in detecting A4 alternative for small values of the parameter. On the other hand, it is remarkable the good performance of the T_ℓ statistic in A1 and A2 families, especially the ones constructed, for values of the

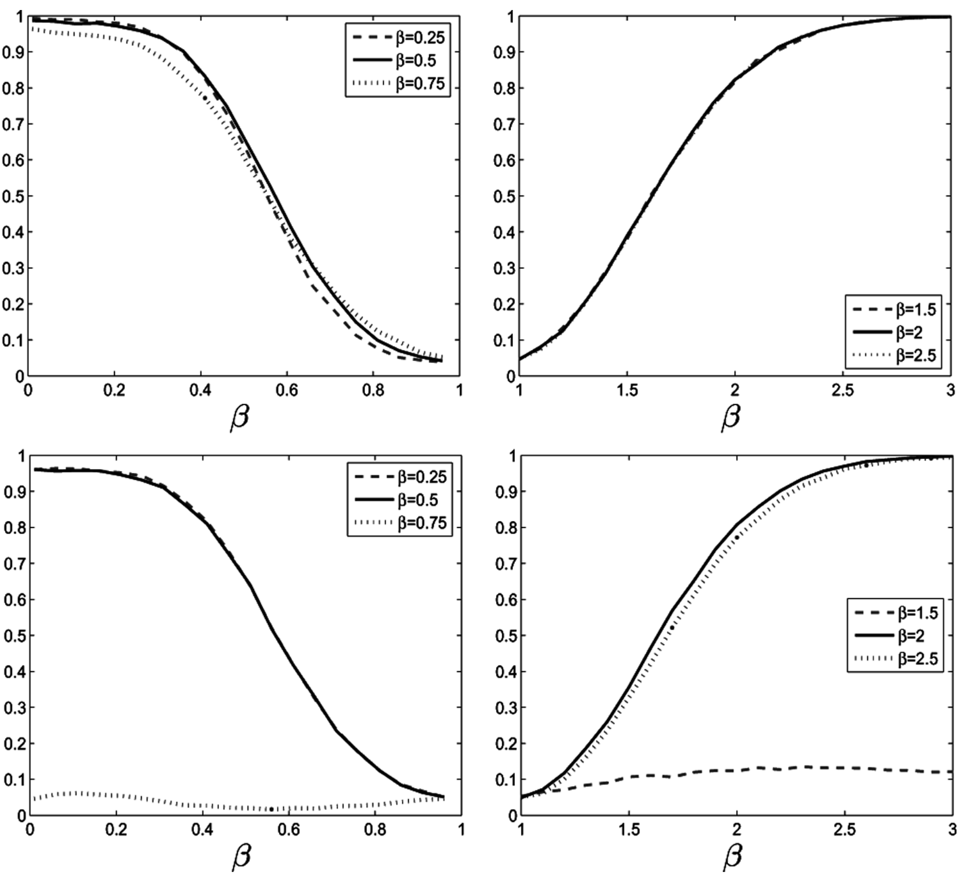

Figure 4. Power of the test based on T_ℓ (top) and T_β (bottom) for the A2 family.

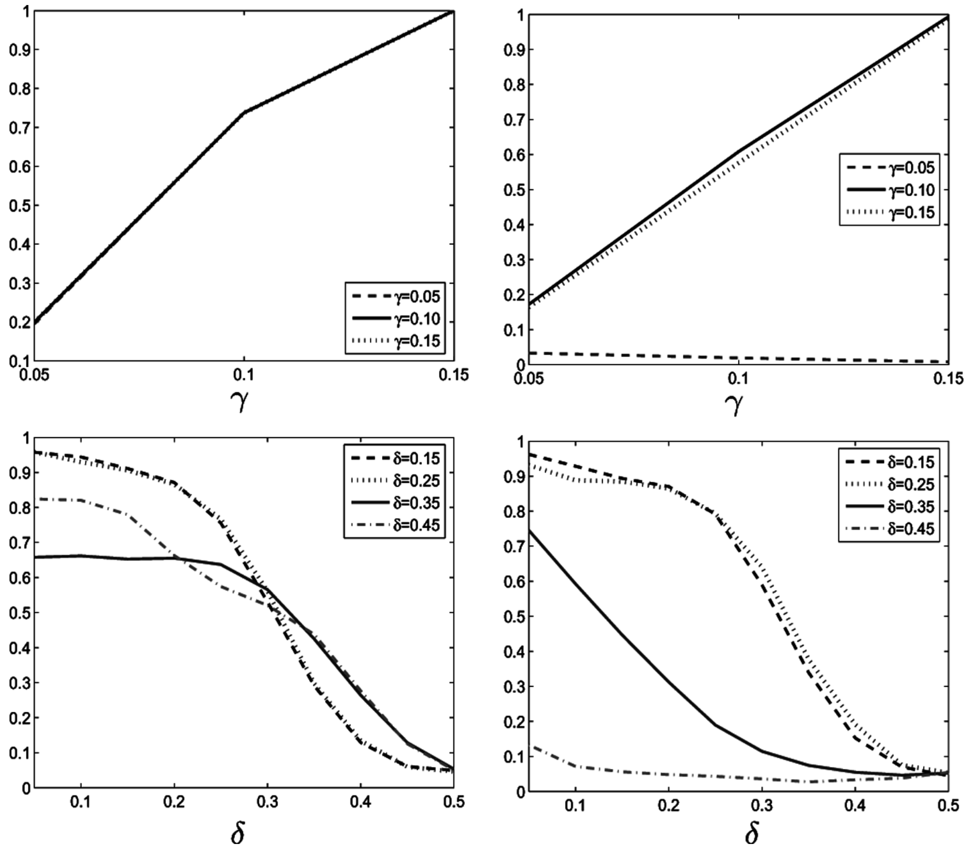

Figure 5. Power of the test based on T_ℓ (left) and T_β (right) for the A3 family (top) and for the A4 family (bottom).

parameter smaller than one, by taking $\alpha = 1/2$ for the A1 family ($\beta = 1/2$ for the A2 family) and, for values of the parameter greater than one, by taking $\alpha = 2$ for the A1 family ($\beta = 2$ for the A2 family). This suggests that the Legendre system is preferable to the KL system at least for A1–A3 alternatives.

5. Bahadur Approximate Slope

Let us consider the family of alternative distributions depending on a parameter θ, such that its cdf is F_θ, and let F_{θ_0} be the cdf of the [0, 1] uniform random variable.

Proposition 5.1. *Let T be the statistic defined in* (1) *and in* (13), *and let* $\{\phi_j\}_{j\geq 0}$ *be an orthonormal sequence in* $L^2[0, 1]$. *Then we have the following convergences*:

$$T \underset{n\to\infty}{\longrightarrow} \mu(\theta) = \sum_{j=0}^{p} \lambda_j \Phi_{\theta,j}, \tag{14}$$

where $\Phi_{\theta,j} = \int_0^1 F_\theta^-(t)\phi_j(t)dt$, *and*

$$\frac{1}{n}\log p_n(t) \underset{n\to\infty}{\longrightarrow} -\frac{1}{2}\left(\frac{t-\mu(\theta_0)}{\sigma_{\theta_0}}\right)^2, \tag{15}$$

where $p_n(t) = P_{H_0}(T \geq t)$, *and* $\mu(\theta_0)$ *and* $\sigma_{\theta_0}^2$ *are, respectively, the expectation and variance of* T *under* H_0.

Proof. The T statistic can be written as:

$$T = \int_0^1 J(t)F_n^-(t)dt,$$

where

$$J(t) = \sum_{j=0}^p \lambda_j \phi_j(t), \quad t \in (0, 1).$$

Convergence (14) is obtained from the general theory of L-statistics (see Theorem 3 in Ch. 19 of Shorack and Wellner, 1986) which ensures the following convergence in law:

$$\sqrt{n}[T-\mu(\theta)] \underset{n\to\infty}{\overset{\mathscr{L}}{\longrightarrow}} N(0, \sigma_\theta^2),$$

where

$$\mu(\theta) = \int_0^1 J(t)F_\theta^-(t)dt, \quad \sigma_\theta^2 = \int_0^1\int_0^1 J(s)J(t)K(s,t)dF_\theta^-(s)dF_\theta^-(t),$$

and $K(s, t) = \min(s, t) - st$. So substituting the expression of $J(t)$ in $\mu(\theta)$, we have that:

$$\mu(\theta) = \sum_{j=0}^p \lambda_j \int_0^1 F_\theta^-(t)\phi_j(t)dt = \sum_{j=0}^p \lambda_j \Phi_{\theta,j},$$

where $\Phi_{\theta,j} = \int_0^1 F_\theta^-(t)\phi_j(t)dt$.

To prove convergence (15) we compute the expectation and variance of T under H_0:

$$\mu(\theta_0) = \int_0^1 J(t)F_{\theta_0}^-(t)dt = \sum_{j=0}^p \lambda_j \int_0^1 t\phi_j(t)dt = \sum_{j=0}^p \lambda_j \Phi_{0,j},$$

where $\Phi_{0,j} = \int_0^1 t\phi_j(t)dt$, (note that, since $\phi_0 = 1$, $\Phi_{\theta,0} = E(F_\theta)$ and $\Phi_{0,0} = E(F_{\theta_0})$) and

$$\sigma_{\theta_0}^2 = \int_0^1\int_0^1 J(s)J(t)K(s,t)ds\,dt = \sum_{j=0}^p\sum_{k=0}^p \lambda_j\lambda_k S_{jk},$$

where

$$S_{jk} = \int_0^1 \int_0^1 \phi_j(s)\phi_k(t)K(s,t)ds\,dt$$
$$= \int_0^1 \left((1-s)\phi_j(s)\int_0^s t\phi_k(t)dt\right)ds + \int_0^1 \left(s\phi_j(s)\int_s^1 (1-t)\phi_k(t)dt\right)ds, \quad (16)$$

and we apply the well-known result for large deviations of a standard normal random variable, described in p. 851 of Shorack and Wellner (1986). □

Proposition 5.2. *Let T be the statistic defined in* (1) *and in* (13), *and let $\{\phi_j\}_{j\geq 0}$ be an orthonormal sequence in $L^2[0, 1]$. Then*:

(i) *The Bahadur approximate slope of T for the F_θ family of distributions is given by*:

$$c^\star(\theta) = \frac{\boldsymbol{\lambda}'\boldsymbol{\phi}\boldsymbol{\phi}'\boldsymbol{\lambda}}{\boldsymbol{\lambda}'\mathbf{S}\boldsymbol{\lambda}},$$

where $\boldsymbol{\lambda} = (\lambda_0, \ldots, \lambda_p)'$, $\boldsymbol{\phi} = (\Phi_{\theta,0} - \Phi_{0,0}, \ldots, \Phi_{\theta,p} - \Phi_{0,p})'$ *and matrix* $\mathbf{S} = (S_{jk})_{0\leq j,k\leq p}$ *defined in* (16).

(ii) *For a fixed value of θ, the maximum of the Bahadur approximate slope of T for F_θ is $c^\star(\theta) = \boldsymbol{\phi}'\mathbf{S}^{-1}\boldsymbol{\phi}$.*

Proof. Part (i) is obtained applying Theorem 1.2.2. of Nikitin (1995). The Bahadur approximate slope of T for F_θ is:

$$c^\star(\theta) = \left(\frac{\mu(\theta) - \mu(\theta_0)}{\sigma_{\theta_0}}\right)^2, \quad (17)$$

which is a quocient of two quadratic forms, since the numerator and denominator of (17) can be written in the following way:

$$(\mu(\theta) - \mu(\theta_0))^2 = \left(\sum_{j=0}^{p} \lambda_j(\Phi_{\theta,j} - \Phi_{0,j})\right)^2 = \boldsymbol{\lambda}'\boldsymbol{\phi}\boldsymbol{\phi}'\boldsymbol{\lambda},$$

where $\boldsymbol{\lambda} = (\lambda_0, \ldots, \lambda_p)'$, $\boldsymbol{\phi} = (\Phi_{\theta,0} - \Phi_{0,0}, \ldots, \Phi_{\theta,p} - \Phi_{0,p})'$ and

$$\sigma_{\theta_0}^2 = \sum_{j=0}^{p}\sum_{k=0}^{p} \lambda_j\lambda_k S_{jk} = \boldsymbol{\lambda}'\mathbf{S}\boldsymbol{\lambda},$$

where $\mathbf{S} = (S_{jk})_{0\leq j,k\leq p}$ and

$$S_{jk} = \int_0^1 \left((1-s)\phi_j(s)\int_0^s t\phi_k(t)dt\right)ds + \int_0^1 \left(s\phi_j(s)\int_s^1 (1-t)\phi_k(t)dt\right)ds.$$

Note that $c^\star(\theta)$ depends on θ through vector $\boldsymbol{\phi}$.

Part (ii): for a fixed value of θ, the maximum of $c^\star(\theta)$ is attained for the eigenvector $\boldsymbol{\lambda}$ of maximum eigenvalue in

$$\boldsymbol{\phi}\boldsymbol{\phi}'\boldsymbol{\lambda} = \xi\mathbf{S}\boldsymbol{\lambda}, \quad \text{with the constraint } \boldsymbol{\lambda}'\mathbf{S}\boldsymbol{\lambda} = 1.$$

Setting $\boldsymbol{\lambda} = \mathbf{S}^{-1/2}\mathbf{u}$, we have that:

$$\mathbf{S}^{-1/2}\boldsymbol{\phi}\boldsymbol{\phi}'\mathbf{S}^{-1/2}\mathbf{u} = \xi\mathbf{u},$$

with the constraint $\mathbf{u}'\mathbf{u} = 1$, whose solution is the (unique with non null eigenvalue) eigenvector $\mathbf{u} = \mathbf{S}^{-1/2}\boldsymbol{\phi}$ with $\xi = \|\mathbf{u}\|^2$ its eigenvalue. Finally, we recover $\boldsymbol{\lambda} = \mathbf{S}^{-1}\boldsymbol{\phi}$ and the maximum Bahadur approximate slope of T for F_θ, for a fixed value of θ, is $c^\star(\theta) = \boldsymbol{\phi}'\mathbf{S}^{-1}\boldsymbol{\phi}$. □

5.1. *Comparing Two Statistics: Bahadur ARE*

Let $\{\phi_j(t)\}_{j\geq 0}$, $\{\psi_j(t)\}_{j\geq 0}$ be two orthonormal bases in $L^2[0,1]$ with $\phi_0(t) = \psi_0(t) = 1$. Let us consider T_1 and T_2 two statistics constructed in the following way:

$$T_1 = \sum_{j=0}^{p} \lambda_{1,j}\Phi_{nj}, \quad T_2 = \sum_{j=0}^{p} \lambda_{2,j}\Psi_{nj},$$

where $\Phi_{nj} = \int_0^1 F_n^-(t)\phi_j(t)dt$, $\Psi_{nj} = \int_0^1 F_n^-(t)\psi_j(t)dt$, $j \geq 0$.

Let $c_1^\star(\theta)$ and $c_2^\star(\theta)$ be the corresponding Bahadur approximate slopes of T_1 and T_2 for the F_θ family of distributions. For a fixed value of θ, let $\boldsymbol{\lambda}_1 = (\lambda_{1,0}, \ldots, \lambda_{1,p})'$ and $\boldsymbol{\lambda}_2 = (\lambda_{2,0}, \ldots, \lambda_{2,p})'$ the eigenvectors that, respectively, maximize $c_1^\star(\theta)$ and $c_2^\star(\theta)$.

We will say that T_1 is asymptotically more efficient (in the Bahadur sense) than T_2 if:

$$\frac{c_1^\star(\theta)}{c_2^\star(\theta)} > 1$$

or equivalently, if:

$$\boldsymbol{\phi}'\mathbf{S}_1^{-1}\boldsymbol{\phi} > \boldsymbol{\psi}'\mathbf{S}_2^{-1}\boldsymbol{\psi},$$

where

$$\boldsymbol{\phi} = (\Phi_{\theta,0} - \Phi_{0,0}, \ldots, \Phi_{\theta,p} - \Phi_{0,p})', \quad \mathbf{S}_1 = (S_{jk}^1)_{0\leq j,k\leq p},$$
$$\boldsymbol{\psi} = (\Psi_{\theta,0} - \Psi_{0,0}, \ldots, \Psi_{\theta,p} - \Psi_{0,p})', \quad \mathbf{S}_2 = (S_{jk}^2)_{0\leq j,k\leq p}.$$

We have used the concept of Bahadur asymptotic efficiency to compare T_ℓ and T_β statistics. Figure 6 shows the Bahadur approximate slopes of T_ℓ and T_β for the A1–A4 families of distributions introduced in Sec. 4. In order to compare T_ℓ and T_β statistics in terms of their Bahadur approximate slopes, we have constructed them for A1–A4 families taking $p = 3, 4, 5, 7$, $q = 2, 3, 4, 5$ and $n = 20$. T_ℓ has been constructed using the Legendre system in both expressions (2) and (7), while T_β has been constructed using the KL system. For $p = 3, 5, 7$, the Bahadur approximate slopes of the two statistics present the same behavior described for $p = 4$ (see Fig. 6), therefore there will be no great changes in terms of Bahadur asymptotic relative efficiency.

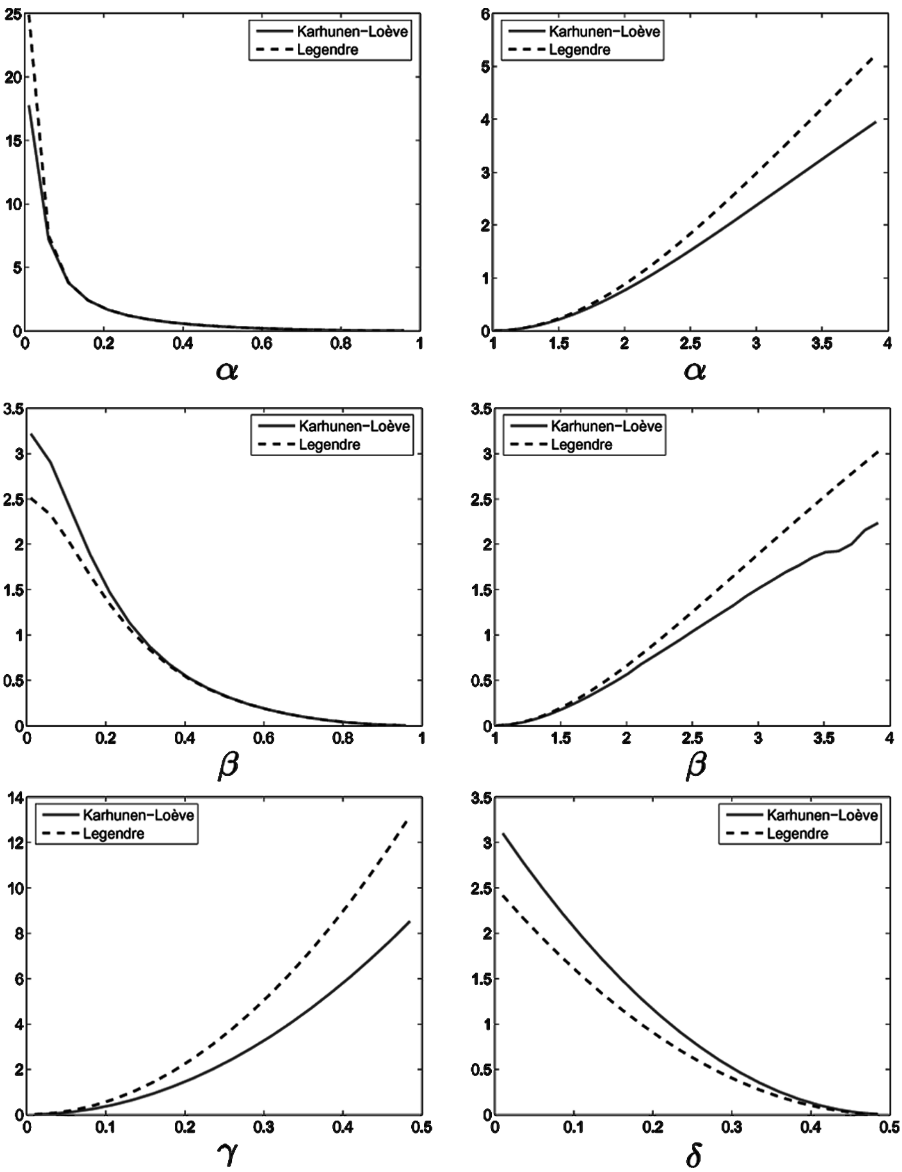

Figure 6. Bahadur approximate slope of T_ℓ and T_β (taking $p = 4$ and $q = 5$) for the A1 alternative (top), A2 alternative (middle), A3 alternative (bottom left), and A4 alternative (bottom right).

As a general comment, it can be said that T_β is preferable to T_ℓ in detecting the A4 family and, only for very small values of q, it is still preferable in detecting A1 and A2 alternatives. But for values of $q \geq 3$, which will be the usual case in practice, T_ℓ statistic performs better for A1, A2, and A3 families. This same behavior has been observed in Sec. 4.

Table 2
Weights of the T-statistic to test uniformity against some patterns. Sample size $n = 20$, $p = 4$.

pdf pattern	Legendre weights					Critical values
right asymmetry	0.866627	−0.389738	0.240569	0.136847	0.143045	$c_{inf} = 0.224193$ $c_{sup} = 0.428548$
left asymmetry	−0.572223	0.373264	−0.547693	0.114750	−0.469147	$c_{inf} = -0.236881$ $c_{sup} = -0.131240$
U-shaped	0.054602	0.955571	0.135623	0.209128	0.147560	$c_{inf} = 0.238348$ $c_{sup} = 0.340117$
wedge-sh.	0.022257	−0.905230	0.058483	−0.412617	0.079936	$c_{inf} = -0.279914$ $c_{sup} = -0.193508$
unimodal compressed	−0.000026	0.720722	0.000092	0.693224	0.000282	$c_{inf} = 0.166335$ $c_{sup} = 0.227677$
	Karhunen–Loève weights					
bimodal both extremes	0	−0.984738	0	0.174041	0	$c_{inf} = 0.197346$ $c_{sup} = 0.328723$

6. Practical Issues

A possible setting for applying the above results is when our data belongs to a given family F_θ with support in [0, 1], and we need to test the null hypothesis $H_0 : \theta = \theta_0$ vs. a fixed alternative $H_1 : \theta = \theta_1$ (known).

Suppose that the data is possibly coming from one of the five following patterns: right asymmetric pdf, left asymmetric pdf, U-shaped pdf, wedge-shaped pdf, or compressed unimodal pdf concentrated around 1/2. We propose to use the

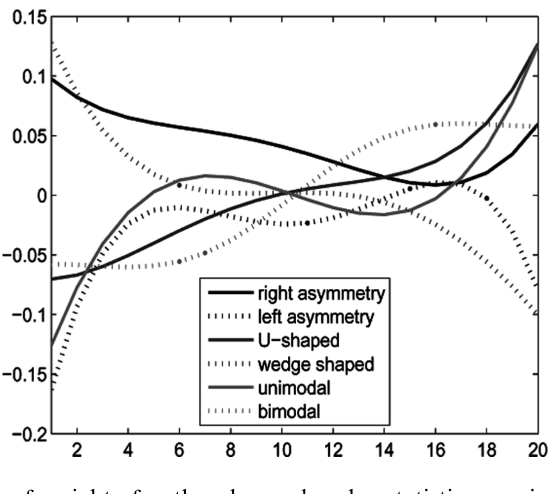

Figure 7. Vectors of weights for the observed order statistic associated to the adaptive statistics described in Table 2.

Legendre orthonormal system in constructing the statistic

$$T = \sum_{j=0}^{p} \lambda_j \Phi_{nj}, \tag{18}$$

and weights $\{\lambda_j\}_{0 \le j \le p}$ described in Table 2 to test uniformity against each pattern. These weights have been obtained for $\theta_1 = 1/2$ for A1 alternative if the histogram

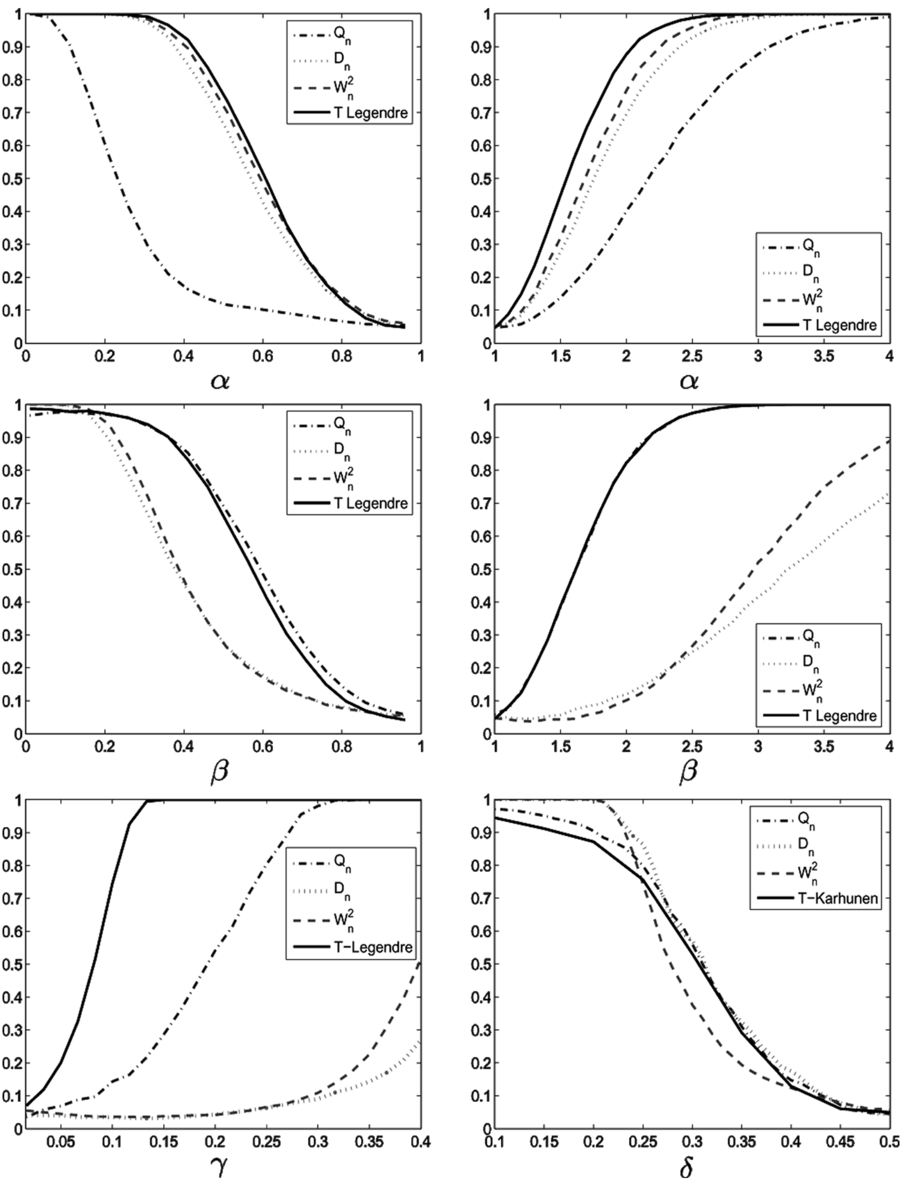

Figure 8. Comparison of the power of the test based on T, Q_n, D_n, and W_n^2 for the A1 family (top), the A2 family (middle), the A3 family (bottom left), and the A4 family (bottom right). See Table 2.

is right asymmetric, $\theta_1 = 1/2$ for A2 alternative if the histogram is U-shaped, $\theta_1 = 2$ for A1 alternative if the histogram is left asymmetric, $\theta_1 = 2$ for A2 alternative if the histogram is wedge-shaped, and $\theta_1 = 0.10$ for A3 alternative if the histogram is unimodal concentrated around 1/2.

If the data is coming from a bimodal pdf with mass concentrated near both extremes, we propose to use the Karhunen–Loève orthonormal system in constructing the statistic (18) and weights $\{\lambda_j\}_{0\leq j\leq p}$ described in Table 2, obtained for $\theta_1 = 0.15$ for A4 alternative. Figure 7 contains the different vectors of weights of the order statistic, associated to the adaptive statistics described in Table 2. We have represented them by solid and dotted lines (and not points) just for better comparison.

Figure 8 contains the power of the 5% significance level test based on the T-statistic described in Table 2 for the A1–A4 families of alternatives. We have compared them with the power of the tests based in Q_n, D_n, and W_n^2 statistics. These powers have been estimated from $N = 10{,}000$ samples of size $n = 20$ as the relative frequency of values of the statistic in the critical region. Note the remarkably good performance of the adaptive statistics.

Acknowledgment

Special thanks to Pedro Delicado for his helpful comments. Research project partially supported by Spanish grants MTM2006-09920 and SEJ2007-64500 (Spanish Ministry of Education and Science and FEDER).

References

Cuadras, C. M., Fortiana, J. (1993). Continuous metric scaling and prediction. In: Cuadras, C. M., Rao, C. R., eds. *Multivariate Analysis, Future Directions 2*. (North–Holland), Amsterdam: Elsevier Science Publishers B. V., pp. 47–66.

Cuadras, C. M., Fortiana, J. (1995). A continuous metric scaling solution for a random variable. *J. Multivariate Anal.* 52:1–14.

David, H. A. (1981). *Order Statistics*. 2nd ed. New York: John Willey & Sons, Inc.

Durbin, J., Knott, M. (1972). Components of Cramér–Von Mises statistics. I. *J. Roy. Statist. Soc. B* 34:290–307.

Durbin, J., Knott, M. (1975). Components of Cramér–Von Mises statistics. II. *J. Roy. Statist. Soc. B* 37:216–237.

Fortiana, J., Grané, A. (2003). Goodness-of-fit tests based on maximum correlations and their orthogonal decompositions. *J. Roy. Statist. Soc. B* 65:1–12.

Grané, A., Fortiana, J. (2006). An adaptive goodness-of-fit test. *Commun. Statist. A. Theor. Meth* 35(6):1141–1155.

Nikitin, Y. (1995). *Asymptotic Efficiency of Nonparametric Tests*. New York: Cambridge University Press.

Rayner, J. C. W., Best, D. J. (1989). *Smooth Tests of Goodness of Fit*. New York: Oxford University Press.

Rayner, J. W. C., Best, D. J. (1990). Smooth tests of goodness of fit: an overview. *Int. Statist. Rev.* 58:9–17.

Shorack, G. R., Wellner, J. A. (1986). *Empirical Processes with Applications to Statistics*. New York: John Wiley & Sons.

Stephens, M. A. (1974). Components of goodness–of–fit statistics. *Annales de l'Institut Henri Poincaré, Section B* 10:37–54.

Stigler, S. M. (1974). Linear functions of order statistics with smooth weight functions. *Ann. Statist.* 2:676–693. Correction in: 7:466, 1979.