UNIVERSIDAD CARLOS III DE MADRID working papers

Working Paper 06-27
Statistics and Econometrics Series 10
April 2006

Departamento de Estadística
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-49

KARHUNEN-LOÈVE BASIS IN GOODNESS-OF-FIT TESTS DECOMPOSITION: AN EVALUATION

Aurea Grané(1) and Josep Fortiana(2)

Abstract:
In a previous paper (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 behaviour 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: Goodness-of fit, orthonormal functions, smooth tests, asymptotic relative efficiency.

AMS subject classification: 62G10, 62G30, 62G20.

(1) A. Grané, Departamento de Estadística, Universidad Carlos III de Madrid, C/ Madrid 126, 28903 Getafe (Madrid), Spain, e-mail: agrane@est-econ.uc3m.es; (2) J. Fortiana, Departament de Probabilitat, Lògica i Estadística, Universitat de Barcelona, C/ Gran Via de les Corts Catalanes, 585, 08007 Barcelona, Spain, e-mail: fortiana@ub.edu . Work partially supported by Spanish grants TEC-2004-07853-C02-01 (Ministerio de Ciencia y Tecnología) and MTM2005-02370 (Ministerio de Edicación y Ciencia- FEDER).

Karhunen-Loève basis in goodness-of-fit tests decomposition: an evaluation

A. Grané[(1)] J. Fortiana[(2)]

(1) Departamento de Estadística. Universidad Carlos III de Madrid.

(2) Departament d'Estadística. Universitat de Barcelona.

Abstract

In a previous paper (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 behaviour 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: Goodness-of-fit, Orthonormal Functions, Smooth tests, Asymptotic Relative Efficiency.

AMS 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.

Work supported in part by TEC2004-07853-C02-01 (Ministerio de Ciencia y Tecnoloía) and MTM2005-02370 (Ministerio de Educación y Ciencia–FEDER)

Authors address: Departamento de Estadística, Universidad Carlos III de Madrid, c/ Madrid, 126, 28903–Getafe, Spain.

E–mail: A. Grané, `agrane@est-econ@uc3m.es`; J. Fortiana, `fortiana@ub.edu`.

Date: April 27, 2006.

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 cdf of n iid random variables and let F_n^- be its pseudoinverse, then β_{nj} is the j-th Fourier coefficient of F_n^- for the orthonormal (in $L^2[0,1]$) sequence $\{\beta_j(t)\}_{j\in\mathbb{N}}$, $t\in[0,1]$, obtained from the eigenfunctions of the covariance kernel of a certain Bernoulli stochastic process associated with the $[0,1]$ uniform distribution (see Fortiana and Grané 2003, Cuadras and Fortiana 1993 and 1995 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 $\beta\equiv\{\beta_{nj}\}_{j\geq 0}$ (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 behaviour was rather poor.
In this work 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 Section 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 $\phi=$ the Legendre polynomials, comparing the power of the statistic obtained with this basis with that of the KL one. In section 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 the concluding remarks.

2 Karhunen-Loève basis and its generalization

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 pseudoinverse 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

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

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

A bilateral test is appropriate in the absence of further information about F_1. Also, we take c_1, c_2 symmetric with respect to $\mu_0 = E(T|H_0)$, that is, $c_1 = \mu_0 + c_{\varepsilon/2}\,\sigma_0$, $c_2 = \mu_0 - c_{\varepsilon/2}\,\sigma_0$, where $\sigma_0^2 = \mathrm{var}(T_p|H_0)$ and $c_{\varepsilon/2}$ is the $(1-\varepsilon/2)\cdot 100$-percentile of the $N(0,1)$ distribution. The power function $P(T > c_1|H_1) + P(T < c_2|H_1)$ 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 $\mu_1 = E(T|H_1)$, $\sigma_1^2 = \mathrm{var}(T|H_1)$ 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:

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

where $\mathbf{M}_i = E(\mathbf{x}|H_i)$, $\boldsymbol{\Sigma}_i = Var(\mathbf{x}|H_i)$, $i = 0, 1$.
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

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

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

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

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 $\beta(\boldsymbol{\lambda}) = 2\,b(\boldsymbol{\lambda})^{-1/2}\,\phi\left(b(\boldsymbol{\lambda})^{1/2}\right)$, and ϕ is the standard normal pdf.
General case: If $a(\boldsymbol{\lambda}) \neq 0$, differentiating (4) and equating to zero we obtain:

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

where $\alpha(\boldsymbol{\lambda}) = a(\boldsymbol{\lambda})^{-1/2}\,(\phi_+(\boldsymbol{\lambda}) - \phi_-(\boldsymbol{\lambda}))$, $\beta(\boldsymbol{\lambda}) = b(\boldsymbol{\lambda})^{-1/2}\,(\phi_+(\boldsymbol{\lambda}) + \phi_-(\boldsymbol{\lambda}))$, $\phi_+(\boldsymbol{\lambda}) = \phi\left(a(\boldsymbol{\lambda})^{1/2} + b(\boldsymbol{\lambda})^{1/2}\right)$, $\phi_-(\boldsymbol{\lambda}) = \phi\left(a(\boldsymbol{\lambda})^{1/2} - b(\boldsymbol{\lambda})^{1/2}\right)$ 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

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

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: 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)

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

where

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

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 $\boldsymbol{\lambda}'\mathbf{A}'\mathbf{M}_0'\mathbf{M}_0\mathbf{A}\boldsymbol{\lambda}\big/\,\boldsymbol{\lambda}'\mathbf{A}'\boldsymbol{\Sigma}_0\mathbf{A}\boldsymbol{\lambda}$, 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 $\boldsymbol{\lambda}'\mathbf{A}'\boldsymbol{\Sigma}_0\mathbf{A}\boldsymbol{\lambda} = 1$.
In Grané and Fortiana (2006) we used the orthonormal basis β referred to above. Explicitly,

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

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:

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

and the recurrence relation is:

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

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).
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 1 shows the power curves for the tests based on these statistics. These curves have been plotted from 20 computed points, for each of which we have generated $N = 1000$ samples of size $n = 20$. We allowed θ to take values below and above 1, thus obtaining a two sided power curve.

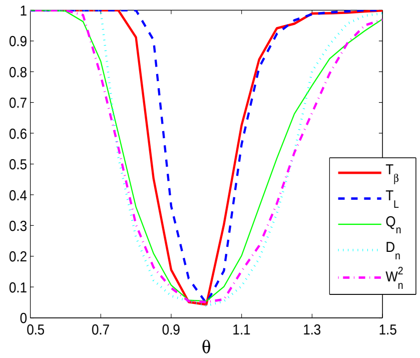
4 Generic alternatives

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,

Figure 1: Power functions for scale alternatives.



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). $\mathbf{M}_0$ and $\boldsymbol{\Sigma}_0$ are the same as in (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

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*

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

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

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

$$I_{jklm} = \int_0^1 \int_0^1 K(s,t) \phi_j(s) \psi_k'(s) \phi_l(t) \psi_m'(t) \, dtds,$$

where $K(s,t) = \min(s,t) - s\,t$ *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 convergences 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:

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

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

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

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

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

if $k \neq j$, and δ is Kronecker's delta. For a complete proof see Grané and Fortiana (2006).
Comparing the expression for $c(\boldsymbol{\lambda}) = \sigma_1^2 = \boldsymbol{\lambda}'\,\mathbf{A}\,\boldsymbol{\Sigma}\,\mathbf{A}\,\boldsymbol{\lambda}$ in (3) with (12), we see that the entries in $\mathbf{A}\,\boldsymbol{\Sigma}\,\mathbf{A}$ are either $\sigma_{n,jl}$ or the limit $\sigma_{jl} = \lim_{n\to\infty}\sigma_{n,jl}$. Some computational examples suggest that a better approximation is obtained with $\sigma_{n,jl}$.

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,

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

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,

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

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

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

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 determine coefficients γ_k of (7), for $0 \le k \le q = 5$. For sample size $n = 20$ and significance level $\varepsilon = 0.05$ we determine $p = 4$ coefficients $\boldsymbol{\lambda}$ for T_β (KL test statistic) and for T_ℓ (with both $\{\phi_j(t)\}_{j\ge 0}$ and $\{\psi_k(t)\}_{k\ge 0}$ the Legendre polynomials). Results for T_β and T_ℓ appear in Table 1 and Table 2, respectively.
Table 3, Table 4, Table 5 and Table 6 contain 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. These powers have been estimated from $N = 10000$ 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, we have included these results in Table 3 for comparison.

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 $\{\phi_j\}_{j\ge 0}$ 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 $p_n(t) = P_{H_0}(T \ge t)$, and $\mu(\theta_0)$ and $\sigma_{\theta_0}^2$ are, respectively, the expectation and variance of T under H_0.

Table 1: Computations for statistic T_β for families A1, A2, A3 and A4.

Family	Fourier coeff.	weights	critical values
A1 $\alpha = 1/2$	$\gamma_0 = \frac{1}{3}$ $\gamma_k = (-1)^k \frac{2\sqrt{2}}{(k\,\pi)^2}$ $1 \leq k \leq q$	-0.801406 -0.426894 0.353300 -0.036017 0.222244	$c_1 = -0.159319$ $c_2 = -0.407395$
A2 $\beta = 2$	$\gamma_0 = 1/2$ $\gamma_1 = -0.197286$ $\gamma_2 = 0$ $\gamma_3 = -0.0448157$ $\gamma_4 = 0$ $\gamma_5 = -0.0197851$	0 -0.971767 0 -0.235944 0	$c_1 = 0.327609$ $c_2 = 0.215799$
A3 $\gamma = 0.15$	$\gamma_0 = 1/2$ $\gamma_k = 0,$ $1 \leq k \leq q,\ \ k$ even, $\gamma_k = -\frac{2\sqrt{2}}{(k\,\pi)^2}(1 - 2\,\gamma),$ $1 \leq k \leq q,\ \ k$ odd.	0 0.837951 0 0.545746 0	$c_1 = -0.200292$ $c_2 = -0.288662$
A4 $\delta = 0.05$	$\gamma_0 = 1/2$ $\gamma_k = 0,$ $1 \leq k \leq q,\ \ k$ even, $\gamma_k = -\frac{4\delta\sqrt{2}}{(k\pi)^2} +$ $\frac{(2\delta-1)\sqrt{2}}{k\pi}\sin(k\pi/2),$ $1 \leq k \leq q,\ \ k$ odd.	0 0.998779 0 -0.049396 0	$c_1 = -0.207086$ $c_2 = -0.334052$

Table 2: Computations for statistic T_ℓ for families A1, A2, A3 and A4.

Family	Fourier coeff.	weights	critical values
A1 $\alpha = 1/2$	$\gamma_0 = 1/3$ $\gamma_1 = 1/(2\sqrt{3})$ $\gamma_2 = 1/(6\sqrt{5})$ $\gamma_k = 0,\ k > 2$	0.866627 -0.389738 0.240569 0.136847 0.143045	$c_1 = 0.428548$ $c_2 = 0.224193$
A2 $\beta = 2$	$\gamma_0 = 1/2$ $\gamma_1 = 0.202073$ $\gamma_2 = 0$ $\gamma_3 = 0.027822$ $\gamma_4 = 0$ $\gamma_5 = 0.006362$	0.02226 -0.905230 0.058483 -0.412617 0.079936	$c_1 = -0.193508$ $c_2 = -0.279914$
A3 $\gamma = 0.15$	$\gamma_0 = 1/2$ $\gamma_1 = \frac{\sqrt{3}}{6}(1 - 2\gamma)$ $\gamma_k = 0 \quad k > 1.$	0.000101 0.717871 0.000330 0.696175 0.000592	$c_1 = 0.226837$ $c_2 = 0.165733$
A4 $\delta = 0.05$	$\gamma_0 = 1/2$ $\gamma_1 = 0.418579$ $\gamma_2 = 0$ $\gamma_3 = -0.148824$ $\gamma_4 = 0$ $\gamma_5 = 0.093280$	0.070022 0.970818 0.149801 -0.147607 0.091548	$c_1 = 0.363615$ $c_2 = 0.239237$

Table 3: Power of the test based on T_ℓ, T_β, Q_n, D_n, W_n^2 and the UMP test for the 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	1.0000
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	1.0000

Table 4: Power of the tests based on T_ℓ, T_β, Q_n, D_n and W_n^2 for the A2 family.

β	T_ℓ	T_β	Q_n	D_n	W_n^2
0.25	0.9678	0.9447	0.9651	0.8071	0.8597
0.5	0.6600	0.6524	0.7238	0.2879	0.2840
0.75	0.1827	0.1893*	0.2203	0.0916	0.0905
2	0.8252	0.8045	0.7523	0.1288	0.1013
3	0.9979	0.9929	0.9955	0.4029	0.5107
4	1.0000	1.0000	1.0000	0.7361	0.8951

Table 5: Power of the test based on T_ℓ, T_β, Q_n,D_n and W_n^2 for the A3 family.

γ	T_ℓ	T_β	Q_n	D_n	W_n^2
0.05	0.2059	0.1761*	0.1016	0.0453	0.0387
0.10	0.7475	0.6049	0.3609	0.0451	0.0426
0.15	1.0000	0.9894	0.8244	0.0677	0.0669
0.25	1.0000	1.0000	1.0000	0.3775	0.6195
0.35	1.0000	1.0000	1.0000	1.0000	1.0000

Table 6: Power of the test based on T_ℓ, T_β, Q_n,D_n and W_n^2 for the 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

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 chapter 19 of Shorack and Wellner (1986)) which ensures the following convergence in law:

$$\sqrt{n}[T - \mu(\theta)] \xrightarrow[n\to\infty]{\mathcal{L}} N(0, \sigma_\theta^2),$$

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 $K(s,t) = \min(s,t) - s\,t$. 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 have to 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

$$S_{jk} = \int_0^1 \int_0^1 \phi_j(s)\, \phi_k(t)\, K(s,t)\, ds\, dt \tag{16}$$

$$= \int_0^1 \left((1-s)\phi_j(s) \int_0^s t\phi_k(t)dt \right) ds + \int_0^1 \left(s\phi_j(s) \int_s^1 (1-t)\phi_k(t)dt \right) ds,$$

and also we need to use the well-known result for large deviations of a standard normal random variable, described in p.851 of Shorack and Wellner (1986):

Lemma 5.1 *Let Z be a standard normal random variable, and consider the sequences $\lambda_n \to \infty$, $\delta_n \to \infty$, then:*

$$P(Z > \lambda_n) = \exp\left[-\frac{\lambda_n^2}{2}(1-\delta_n)\right], \quad n \to \infty.$$

In our case, we have that:

$$\begin{aligned} p_n(t) = P_{H_0}(T \geq t) &= P\left(\frac{T-\mu(\theta_0)}{\sigma_{\theta_0}/\sqrt{n}} \geq \frac{t-\mu(\theta_0)}{\sigma_{\theta_0}/\sqrt{n}}\right) \\ &= \exp\left\{-\frac{(t-\mu(\theta_0))^2}{2\sigma_{\theta_0}^2/n}(1-\delta_n)\right\}. \end{aligned}$$

And

$$\lim_{n\to\infty}\frac{1}{n}\log p_n(t) = \lim_{n\to\infty}\frac{1}{n}\left\{-\frac{(t-\mu(\theta_0))^2}{2\sigma_{\theta_0}^2}n\,(1-\delta_n)\right\} = -\frac{1}{2}\left(\frac{t-\mu(\theta_0)}{\sigma_{\theta_0}}\right)^2.$$

□

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

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

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

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

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

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

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 $c^\star(\theta) = \boldsymbol{\phi}' \, \mathbf{S}^{-1} \, \boldsymbol{\phi}$. □

Comparing two statistics: Bahadur ARE

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

$$\boldsymbol{\phi} = (\Phi_{\theta,0} - \Phi_{0,0}, \ldots, \Phi_{\theta,p} - \Phi_{0,p})', \quad \mathbf{S}_1 = \left(S^1_{jk}\right)_{0\le j,k\le p},$$

$$\boldsymbol{\psi} = (\Psi_{\theta,0} - \Psi_{0,0}, \ldots, \Psi_{\theta,p} - \Psi_{0,p})', \quad \mathbf{S}_2 = \left(S^2_{jk}\right)_{0\le j,k\le p}.$$

We have used the concept of Bahadur asymptotic efficiency to compare T_ℓ and T_β statistics. Figure 2, Figure 3 and Figure 4 show the Bahadur approximate slopes of T_ℓ and T_β for the A1, A2, A3 and A4 families of distributions introduced in section 4.

Figure 2: Bahadur approximate slope of T_ℓ and T_β ($p = 4$) for the A1 alternative

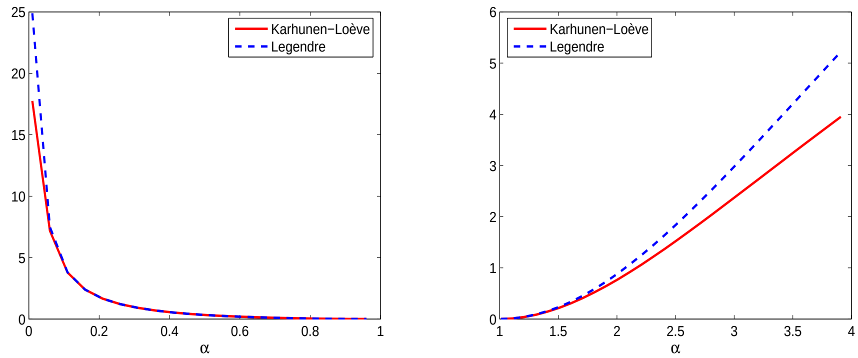
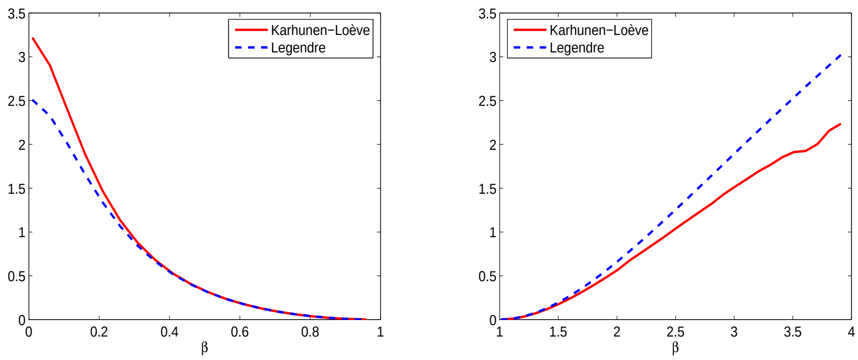
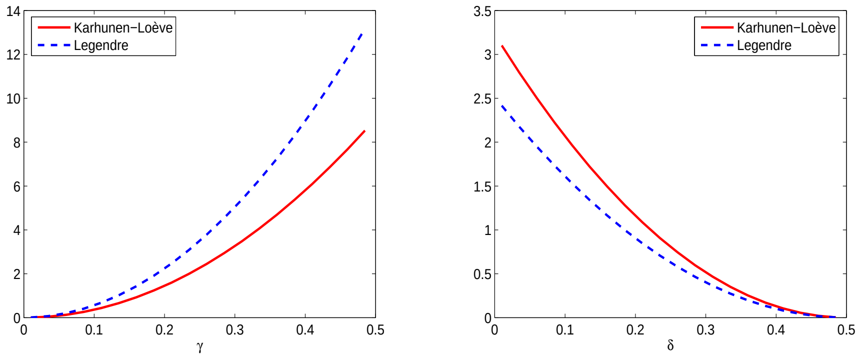


Figure 3: Bahadur approximate slope of T_ℓ and T_β ($p = 4$) for the A2 alternative



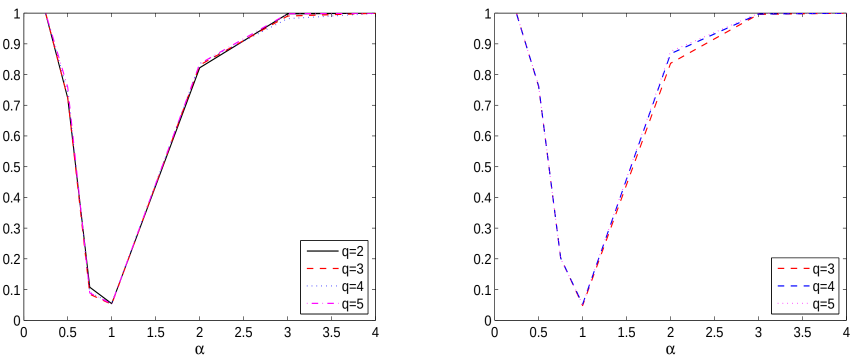
In order to compare T_ℓ and T_β statistics in terms of their power values and in terms of their Bahadur approximate slopes, we have constructed them

Figure 4: Bahadur approximate slope of T_ℓ and T_β ($p = 4$) for the A3 alternative (on the left) and for A4 alternative (on the right)



for each family of distributions (A1, A2, A3, A4), taking $p = 3, 4, 5, 7$, $q = 2, 3, 4, 5$ and $n = 20$. The powers have been estimated from $N = 10000$ samples of size $n = 20$.
For $p = 3, 5, 7$, the Bahadur approximate slopes of the two statistics present the same behaviour described for $p = 4$ (see Figure 2, Figure 3, Figure 4), therefor there will be no great changes in terms of Bahadur asymptotic relative efficiency.
For better comparison we have plotted the power values of T_ℓ and T_β for $q = 2, 3, 4, 5$. Figure 5, Figure 6, Figure 7 and Figure 8 contain these plots.

Figure 5: Power of T_β (on the left) and T_ℓ (on the right) for the A1 alternative, for $q = 2, 3, 4, 5$ and values of the parameter $\alpha = 0.25, 0.5, 0.75, 1, 2, 3, 4$.



As a general comment it can be said that for $q = 2$, T_β is preferable to T_ℓ, except for the A3 family. But, in general, when $q \geq 3$, T_ℓ performs better.

Figure 6: Power of T_β (on the left) and T_ℓ (on the right) for the A2 alternative, for $q = 2, 3, 4, 5$ and values of the parameter $\beta = 0.25, 0.5, 0.75, 1, 2, 3, 4$.

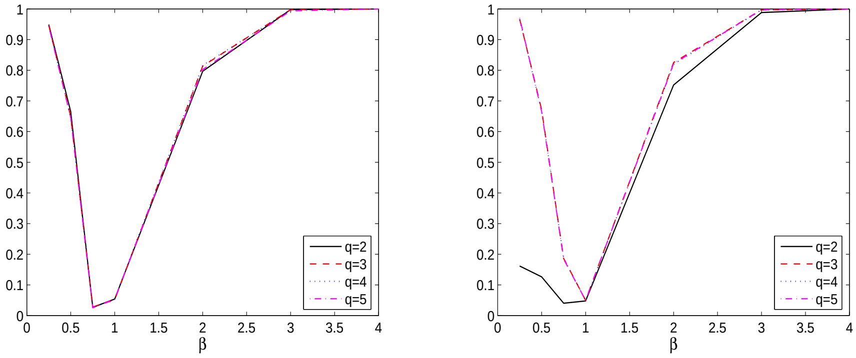
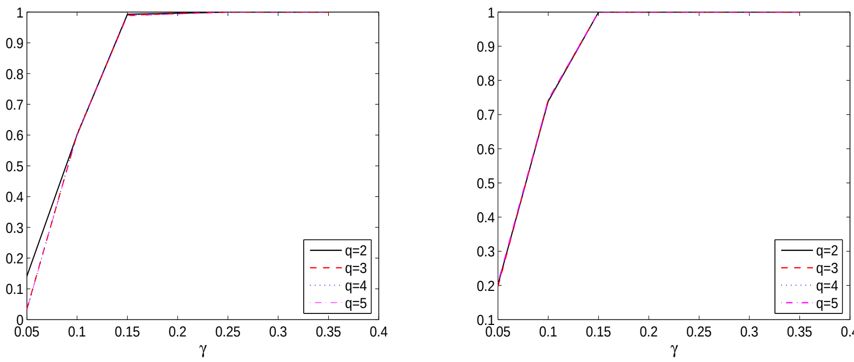
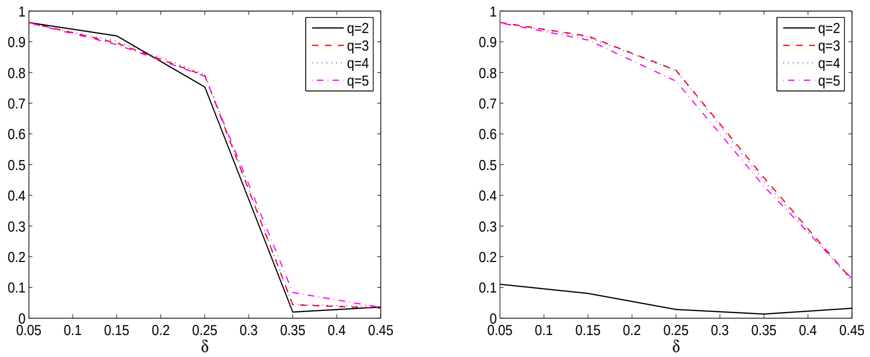


Figure 7: Power of T_β (on the left) and T_ℓ (on the right) for the A3 alternative, for $q = 2, 3, 4, 5$ and values of the parameter $\beta = 0.05, 0.10, 0.15, 0.25, 0.35$.



More precisely:
For the A1 family T_ℓ is preferable to T_β, in terms of power and in terms of Bahadur approximate slopes. Both statistics are, in general, better than Q_n, D_n and W_n^2, but, obviously, the UMP-test is the best.
For the A2 family and when the parameter $\beta > 1$, T_ℓ is the best. For very small values of the parameter, T_β is more efficient than T_ℓ. But for $0.2 < \beta < 1$, T_ℓ has more power than T_β, although they are equally efficient.
For the A3 family T_ℓ is the best and for the A4 family T_β is more efficient than T_ℓ but, in general, T_ℓ has more power. A possible explanation of this fact is that the pseudo-inverse of the A4 family, $F_\delta^-(t)$, is a discontinuous function. In fact, for A4, D_n is the most powerful statistic.

Figure 8: Power of T_β (on the left) and T_ℓ (on the right) for the A4 alternative, for $q = 2, 3, 4, 5$ and values of the parameter $\beta = 0.05, 0.15, 0.25, 0.35, 0.45$.



6 Practical implementation and concluding remarks

Given two orthonormal basis in $L^2[0,1]$, $\{\phi_j\}_{j\geq 0}$ and $\{\psi_j\}_{j\geq 0}$, and a family of distributions F_θ, such that F_{θ_0} is the cdf of a $[0,1]$ uniform random variable, we want to construct the statistic

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

which has maximum power for testing $H_0 : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta_0$.
We recommend:

1. Select the orthonormal basis in which the T statistic is more efficient, in the Bahadur sense, for F_θ alternative.
2. For this orthonormal basis, select q such that the Fourier expansion of order q is a good approximation of $F_\theta^-(t)$. In our examples, $q = 4$ or $q = 5$ were sufficient.
3. Take $p = 4$ to construct the T statistic and apply the algorithm described in section 3 to find coefficients $\lambda_0, \ldots, \lambda_p$.

References

Cuadras, C. M. and J. Fortiana (1993). Continuous metric scaling and prediction. In C. M. Cuadras and C. R. Rao (Eds.), *Multivariate Anal-*

ysis, Future Directions 2, pp. 47–66. Elsevier Science Publishers B. V. (North–Holland), Amsterdam.

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

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

Durbin, J. and M. Knott (1972). Components of Cramér–Von Mises statistics. I. *Journal of the Royal Statistical Society B 34*, 290–307.

Durbin, J. and M. Knott (1975). Components of Cramér–Von Mises statistics. II. *Journal of the Royal Statistical Society B 37*, 216–237.

Fortiana, J. and A. Grané (2003). Goodness-of-fit tests based on maximum correlations and their orthogonal decompositions. *Journal of the Royal Statistical Society B 65*, 1–12.

Grané, A. and J. Fortiana (2006). An adaptive goodness-of-fit test. *Communications in Statistics A. Theory and Methods 35 (6)*.

Inglot, T., W. Kallenberg, and T. Ledwina (1994). Power aproximations to and power comparison of smooth goodness-of-fit tests. *Scandinavian Journal of Statistics 21*, 131–145.

Nikitin, Y. (1995). *Asymptotic efficiency of nonparametric tests*. New York: Cambridge University Press.

Rayner, J. C. W. and D. J. Best (1989). *Smooth tests of goodness of fit*. New York: Oxford University Press.

Rayner, J. W. C. and D. J. Best (1990). Smooth tests of goodness of fit: an overview. *International Statistical Review 58*, 9–17.

Shorack, G. R. and J. A. Wellner (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. *The Annals of Statistics 2*, 676–693. Correction in: Vol. 7 (1979), pag. 466.