Communications in Statistics - Theory and Methods

An Adaptive Goodness-of-Fit Test

Aurea Grane a, Josep Fortiana b

a Dpto. Estadística, Universidad Carlos III Madrid. Madrid. Spain
b Department d’Estadística, Universitat de Barcelona. Barcelona. Spain

To link to this article: DOI: 10.1080/03610920600672286
URL: http://dx.doi.org/10.1080/03610920600672286

Full terms and conditions of use: http://www.informaworld.com/terms-and-conditions-of-access.pdf
This article maybe used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.
The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
© Taylor and Francis 2007
Goodness-of-Fit Tests

An Adaptive Goodness-of-Fit Test

AUREA GRANE1 AND JOSEP FORTIANA2

1Dpto. Estadística, Universidad Carlos III Madrid, Madrid, Spain
2Department d’Estadística, Universitat de Barcelona, Barcelona, Spain

In Fortiana and Grané (2003), we introduced the statistic $Q_n$, based on Hoeffding’s maximum correlation, as a general-purpose goodness-of-fit test of uniformity. It admits an expansion along a countable set of orthogonal axes, originating a sequence of statistics. Linear combinations of a given number $p$ of terms in this sequence have easy-to-compute probability distributions, either the exact ones for a finite sample or their normal asymptotic approximations for a large sample. In this article we develop an algorithm for tailoring a statistic within this class of linear combinations to test uniformity with optimal power against a specific alternative or family of alternatives.

Keywords Decomposition of test statistics; Goodness-of-fit; Maximum correlation; Tests of uniformity.

Mathematics Subject Classification 62G10; 62G30; 62E15.

1. Introduction

Most goodness-of-fit statistics can be regarded as measures of proximity between two distributions: empirical and hypothesized. The family of tests we are concerned with in this article is based on Hoeffding’s maximum correlation between cumulative distribution functions (cdf’s) $F_1$ and $F_2$, henceforth denoted by $\rho^+(F_1, F_2)$, defined as the maximum of the correlation coefficients of bivariate distributions having $F_1$ and $F_2$ as marginals. Since $\rho^+(F_1, F_2)$ equals 1 if and only if $F_1 = F_2$ (almost everywhere) up to a scale and location change, it is adequate to our purpose. In Fortiana and Grané (2003) we defined and studied the properties of

$$Q_n = \frac{s_n}{\sqrt{1/12}} \rho^+(F_n, F_U),$$

where $F_n$ is the empirical cdf for $n$ iid real-valued random variables, $s_n$ is the empirical standard deviation, and $F_U$ is the probability distribution function of a

Address correspondence to Aurea Grane, Dpto. Estadística, Universidad Carlos III Madrid, C/Madrid 126, 28903 Getafe, Madrid, Spain; E-mail: agrane@est-econ.uc3m.es
[0, 1] uniform random variable, as a test statistic for $H_0 : F = F_U$ (also for $H_0 : F = F_0$, where $F_0$ is a fully specified continuous probability distribution function, case 0 in Stephens, 1986). We found that $Q_n$ has good properties as a goodness-of-fit test: since it is an $L$-statistic, it is possible to find its exact distribution under the null hypothesis for small samples and, additionally, it is asymptotically normally distributed. The test based on $Q_n$ can advantageously replace those of Kolmogorov–Smirnov, Cramér-von Mises, and Anderson–Darling for a wide range of alternatives. We also proved the following identity:

$$Q_n = \frac{24\sqrt{2}}{\pi^2} \sum_{j=0}^{\infty} \frac{\beta_{n,2j+1}}{(2j+1)^2},$$

where the sequence $\{\beta_{nj}\}_{j=0}^\infty$ appears as a decomposition of this test, analogous to those studied by Anderson and Darling (1952, 1954), Durbin and Knott (1972, 1975), and Stephens (1974). Also, $\beta_{nj}$ is the $j$th Fourier coefficient of the pseudoinverse of the empirical distribution function $F_n^-$ for an orthonormal sequence, $\{\beta_j(t)\}_{j=0}^\infty$, in $L^2[0, 1]$ (see Cuadras and Fortiana, 1993; Fortiana and Cuadras, 1997; Fortiana and Grané, 2003 for details). In the present article, we seek to improve the performance of $Q_n$ for an alternative or family of alternatives by choosing a linear combination of a finite number $p$ of terms in the sequence $\{\beta_{nj}\}_{j=0}^\infty$ that maximizes power against the given alternative. Even if the best Neyman–Pearson power is not attained, the resulting test keeps the computational advantages of $Q_n$.

In Sec. 2 we define and list properties of the class of statistics we are interested in. In Sec. 3 we express the problem of power optimization in terms of certain quadratic forms depending on moments of the order statistic. As an example of family of alternatives we use the $[0, \theta]$-uniform distributions. Section 4 contains a method for finding the statistic in a generic case, including an algorithm to perform the actual computation, with some numerical illustrations. Finally, Sec. 5 includes power computations, as well as a general discussion of results.

2. Definition of the Statistic

Let $F$ be a probability distribution function with finite second order moment and let $F_n$ be the empirical cdf of $n$ i.i.d. random variables with distribution $F$, $x_1, \ldots, x_n$. With the orthonormal (in $L^2[0, 1]$) system $\{\beta_0 = 1; \beta_j(t) = \sqrt{2} \cos(j\pi t), j \geq 1\}$, we define the statistics:

$$\beta_{nj} = \beta_{nj}(F) = \int_{0}^{1} F^{-n}_n(t) \beta_j(t) dt, \quad j \geq 0, \quad (1)$$

where $F^{-n}_n$ is the pseudoinverse of $F_n$, defined as $F^{-n}_n(t) = \inf \{x \in \mathbb{R} : F_n(x) \geq t\}$. This sequence was thoroughly studied in Fortiana and Grané (2003). Explicitly, in terms of the order statistic $\mathbf{x} = (x_{(1)}, \ldots, x_{(n)})'$,

$$\beta_{n0} = \bar{x}_n,$$

$$\beta_{nj} = \frac{\sqrt{2}}{j\pi} \sum_{i=1}^{n} \left( \sin \frac{i\pi}{n} - \sin \frac{(i-1)j\pi}{n} \right) x_{(i)}, \quad j \geq 1, \quad (2)$$
and their expectations under $H_0$ are given by

\[
E(\beta_{n0}) = \frac{1}{2},
\]

\[
E(\beta_{nj}) = \begin{cases} 
0, & \text{if } j \text{ is even}, \\
-\frac{\sqrt{2}}{j\pi} \frac{1}{n+1} \cot \left( \frac{j\pi}{2n} \right), & \text{if } j \text{ is odd}, \quad j \geq 1.
\end{cases}
\] (3)

In Fortiana and Grané (2003, Eq. (3), p. 118), the $\beta_{nj}$ statistics were defined in a slightly different way. Here we have included $\beta_{n0}$ and changed the sign of $\beta_{nj}$, for $j \geq 1$. We will consider all statistics of the form

\[
T = T(\lambda) = \sum_{j=0}^{\infty} \tilde{\lambda}_j \beta_{nj},
\] (4)

where $\{\tilde{\lambda}_j\} \in \ell^1_k$ is a sequence of real numbers. $T$ is an $L$-statistic,

\[
T = w'x = \sum_{i=1}^{n} w_i x_{(i)},
\]

with coefficients

\[
w_i = \frac{\tilde{\lambda}_0}{n} + \frac{\sqrt{2}}{\pi} \sum_{j=0}^{\infty} \frac{\tilde{\lambda}_j}{j} \left( \sin \frac{ij\pi}{n} - \sin \frac{(i-1)j\pi}{n} \right).
\]

The $\ell^1$ condition imposed on $\{\tilde{\lambda}_j\}$ ensures that each $w_i$, $0 \leq i \leq n$, is the sum of an absolutely convergent sequence. Moreover, due to the periodicity $j\beta_{nj} = (j+2n)\beta_{n,j+2n}$, $j \geq 1$, definition (4) is equivalent to

\[
T = T(\lambda) = \sum_{j=0}^{2n} \tilde{\lambda}_j \beta_{nj},
\] (5)

where

\[
\tilde{\lambda}_0 = \tilde{\lambda}_0, \quad \tilde{\lambda}_j \sum_{k=0}^{\infty} \tilde{\lambda}_{2nk+j} \frac{j}{j+2nk}, \quad \text{for } 1 \leq j \leq 2n.
\]

Note that this is an absolutely convergent sum. We will use $T_p \equiv T_p(L)$, the result of truncating (5) at $j = p$, where $L = (\tilde{\lambda}_0, \ldots, \tilde{\lambda}_p)'$. Since in practical situations $p$ will be much smaller than $n$, henceforth we assume $p < n$.

In matrix notation

\[
T_p(L) = w'x = L'TS_p'(I - N)'x,
\] (6)
where $L = (\lambda_1, \ldots, \lambda_p)'$, $\Gamma = \text{diag}(1/n, \sqrt{2}/\pi, \sqrt{2}/(2\pi), \ldots, \sqrt{2}/(p\pi))$, 

$$I - N = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ -1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & -1 & 1 \end{pmatrix}, \quad S_p = \begin{pmatrix} 1 & s_{11} & s_{12} & \cdots & s_{1p} \\ 2 & s_{21} & s_{22} & \cdots & s_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ n & s_{n1} & s_{n2} & \cdots & s_{np} \end{pmatrix},$$

where $s_{ij} = \sin(ij\pi/n)$, $1 \leq i \leq n$, $1 \leq j \leq p$.

Given an alternative cdf $F_1$, we select $L$ in order to maximize power for testing $H_0 : F = F_U$, vs. $H_1 : F = F_1$. Clearly, the resulting test is less powerful than the optimal (Neyman–Pearson), but on the other hand, 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).

From formula (2.9) in Ramallingam (1989), it is also possible to get explicit expressions for the moments under $H_0$:

$$\mu_r = E(T^r) = \frac{r!n!}{(n+r)!} \prod_{i=2}^{n} \frac{1}{b_i} \sum_{i=2}^{n} b_i^{r+1},$$

in particular,

$$\mu_1 = \frac{1}{n+1} \frac{\sum_{i=2}^{n} b_i^2}{\prod_{i=2}^{n} b_i},$$

where

$$b_i = \sum_{k=i}^{n} w_k = \frac{n-i+1}{n} \hat{\lambda}_0 - \frac{\sqrt{2}}{\pi} \sum_{j=1}^{p} \frac{\lambda_j}{j} \sin \left( i - 1 \right) \frac{j\pi}{n}.$$
The power,
\[ P(T_p(L) > c_1 | H_1) + P(T_p(L) < c_2 | H_1), \]
is asymptotically approximated by
\[ \Phi(L) = 1 - P_Z \left[ \left( \frac{\mu_0 - \mu_1}{\sigma_1} - c_{e/2} \frac{\sigma_0}{\sigma_1}, \frac{\mu_0 - \mu_1}{\sigma_1} + c_{e/2} \frac{\sigma_0}{\sigma_1} \right) \right], \]
where \( \mu_1 = E(T_p(L) | H_1) \), \( \sigma_1^2 = var(T_p(L) | H_1) \), and \( Z \sim N(0,1) \). Due to the symmetry of the normal distribution, \( \mu_0 - \mu_1 \) can be replaced with \( |\mu_0 - \mu_1| \), hence
\[ \Phi(L) = 1 - P_Z \left\{ \left( \left[ \frac{a(L)}{c(L)} \right]^{1/2} - \left[ \frac{b(L)}{c(L)} \right]^{1/2} \right), \left[ \frac{a(L)}{c(L)} \right]^{1/2} + \left[ \frac{b(L)}{c(L)} \right]^{1/2} \right\}, \]
in terms of the following quadratic forms:
\[ a(L) = (\mu_0 - \mu_1)^2 = L'AL, \quad b(L) = c_{e/2}^2 \sigma_0^2 = L'BL, \quad c(L) = \sigma_1^2 = L'CL, \]
where \( M_i = E(x | H_i) \), \( \Sigma_i = Var(x | H_i) \), \( i = 0,1 \), and \( D = (I - N)S_p \Gamma \). The expectation and variance of the order statistic under the null hypothesis are given by formula (10) below. See Sec. 4 to compute them under the alternative hypothesis.

As stated in the previous section, we want to find \( L \) such that the power is maximized. Since \( \Phi(L) \) is invariant when \( L \) is multiplied by an arbitrary constant, we assume \( c(L) = 1 \), thus we have to compute the extremes of
\[ \Theta(L) = 1 - \Phi(a(L)^{1/2} + b(L)^{1/2}) + \Phi(a(L)^{1/2} - b(L)^{1/2}) + \lambda(c(L) - 1), \]
where \( \Phi \) is the standard normal probability distribution function and \( \lambda \) is a Lagrange multiplier.

**Degenerate case.** If \( a(L) = 0 \), the expectation of \( T_p(L) \) is the same under both hypotheses, then the asymptotically approximated power function is
\[ \Phi(L) - 1 - P_Z \left\{ \left( \left[ \frac{b(L)}{c(L)} \right]^{1/2} - \left[ \frac{b(L)}{c(L)} \right]^{1/2} \right) \right\}, \]
with the restriction \( c(L) = 1 \), and (8) can be written as
\[ \Theta(L) = 2 - 2\Phi(b(L)^{1/2}) + \lambda(c(L) - 1). \]
Differentiating with respect to \( L \) and equating to zero, we obtain an eigenvalue-type problem:
\[ \beta(L)BL = \lambda CL, \]
where \( \beta(L) = b(L)^{-1/2} \hat{\phi}(b(L)^{1/2}) \), and \( \hat{\phi} \) is the standard normal probability density function.

**General case.** If \( a(L) \neq 0 \), then differentiating (8) with respect to \( L \) and equating to zero, we obtain the following eigenvalue-type problem:

\[
[z(L)A + \beta(L)B]L = \lambda CL,
\]

where

\[
\begin{align*}
z(L) &= a(L)^{-1/2}(\phi_+(L) - \phi_-(L)), \\
\beta(L) &= b(L)^{-1/2}(\phi_+(L) + \phi_-(L)), \\
\phi_+(L) &= \phi(a(L)^{1/2} + b(L)^{1/2}), \\
\phi_-(L) &= \phi(a(L)^{1/2} - b(L)^{1/2}).
\end{align*}
\]

The degenerate case appears for \( z(L) = 0 \).

In order to compute \( L \) we change to the new variable \( u = C^{1/2}L \), and let \( E = C^{-1/2}AC^{-1/2}, F = C^{-1/2}BC^{-1/2} \), and \( G(u) = z(u)E + \beta(u)F \), where \( z(u) \), \( \beta(u) \) denote the quantities \( z(L) \), \( \beta(L) \) as defined in (9) in terms of the new variable \( u \). Given an initial \( u \), we compute the set of eigenvectors/eigenvalues of \( G(u) \). The new \( u \) will be the eigenvector such that \( \Psi(u) \) is maximum. This process is iterated to stationarity. The last step is to recover and normalize \( L \). 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.

### 3.1. Example: Scale Alternatives

We consider an alternative distribution belonging to the family \( U[0, \theta] \), uniform on \( [0, \theta] \), with \( \theta > 0 \). The vector of expectations \( \mathbf{M}_0 \) and the matrix of covariances \( \Sigma_0 \) of the order statistic \( x_{(1)}, \ldots, x_{(n)} \) obtained from \( n \) random variables iid ~ \( U[0, 1] \) are

\[
\begin{align*}
\mathbf{M}_0 &= \frac{1}{n+1}(1, 2, \ldots, n)', \\
\Sigma_0 &= (v_{ij})_{1 \leq i, j \leq n},
\end{align*}
\]

where

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

(see, e.g., David, 1981) and those of the order statistic obtained from \( n \) iid ~ \( U[0, \theta] \) random variables are \( \mathbf{M}_i = \theta\mathbf{M}_0, \Sigma_1 = \theta^2\Sigma_0 \). Hence (7) becomes

\[
\begin{align*}
a(L) &= (1 - \theta)^2 L'AL, \quad \text{where } A = D'M_0M_0D, \\
b(L) &= c_{\theta/2}^2 L'BBL, \quad \text{where } B = D'S_0D, \\
c(L) &= \theta^2 L'BBL.
\end{align*}
\]
We have to maximize
\[ \Psi(L) = 1 - P_z \left( \frac{1 - \theta}{\theta} \left( \frac{L'AL}{L'BL} \right)^{1/2} - \frac{c_{\varepsilon/2}}{\varepsilon} \right) \]
under the restriction \( L'BL = 1 \), which is equivalent to maximizing the quotient \( L'AL/L'BL \), and also to finding the eigenvector of maximum eigenvalue in \( AL = \lambda BL \) with the restriction \( L'BL = 1 \), i.e.,
\[ D'M_0M_0DL = \lambda D'S_0DL \]
with the restriction \( L'D'S_0DL = 1 \). Note that the left-hand side matrix has unit rank, hence there is only one eigenvector with a non null eigenvalue. Additionally, this solution does not depend on the parameter \( \theta \).

As a numerical example, for a sample size \( n = 20 \), a significance level \( \varepsilon = 0.05 \) and \( p = 4 \),
\[ T_p = 0.3554\beta_{n0} - 0.4447\beta_{n1} + 0.4985\beta_{n2} - 0.4373\beta_{n3} + 0.4860\beta_{n4}. \]
In a practical situation, \( T_p \) should be expressed directly in terms of the observed order statistic using (6). The critical values of the test based on \( T_p \) are
\[ c_1 = 0.33974, \quad c_2 = 0.26608. \]

We have compared \( T_p \) with the statistic \( Q_n \) obtained in Fortiana and Grané (2003), with the Kolmogorov–Smirnov statistic \( D_n \) and the Cramér-von Mises statistic \( W_n^2 \). Figure 1 shows the power curves for the tests based on these statistics. These curves are plotted from 20 computed points, for each of which we have generated \( N = 1000 \).

![Figure 1. Power functions for scale alternatives.](image)
samples of size \( n = 20 \), and we have estimated the power as the relative frequency of values in the critical region. We allowed \( \theta \) to take values below and above 1, thus obtaining a two-sided power curve.

4. Generic Alternatives

If for a given alternative distribution we know how to compute \( M_1 \) and \( \Sigma_1 \) in Eq. (7), we can proceed as described in Sec. 3. Here we develop an algorithm to compute the \( T_p \) statistic for an alternative cdf \( F \) such that its pseudoinverse \( F^- \) admits the expression:

\[
F^- (t) = \gamma_0 + \sqrt{2} \sum_{k=1}^{q} \gamma_k \cos(k \pi t),
\]

(11)

where \( \gamma_k \) are real numbers. An arbitrary cdf can be approximated by taking the first \( q \) terms of the corresponding Fourier series. The pseudoinverse \( F^- \) appears in the formulas of the moments of \( L \)-statistics, (18) and (19) see below, hence it seems natural to expand \( F^- \) instead of \( F \) itself or its density. \( M_0 \) and \( \Sigma_0 \) in (7) are the same as in (10). The entries in \( M_1 \) are given by

\[
E(x_{(i)} | H_t) = i \binom{n}{i} \int_0^1 F^- (t) t^{-1} (1 - t) n^{-i} dt
\]

\[
= i \binom{n}{i} \gamma_0 \text{Beta}(i, n - i + 1) + i \binom{n}{i} \sqrt{2} \sum_{k=1}^{q} \frac{q}{k} \int_0^1 \cos(k \pi t) t^{-1} (1 - t) n^{-i} dt
\]

\[
= \gamma_0 + \sqrt{2} \sum_{k=1}^{q} \gamma_k F_{2,3} \left( \left\{ \frac{i + 1}{2}, \frac{i}{2} \right\}; \left\{ \frac{1}{2}, \frac{n + 1}{2}, \frac{n + 2}{2} \right\}; -\frac{k^2 \pi^2}{4} \right), 1 \leq i \leq n,
\]

where

\[
F_{p,q}(a; b; z) = \sum_{k=0}^{\infty} \frac{(a_1) \cdots (a_p) z^k}{(b_1) \cdots (b_q) k!}
\]

is the generalized hypergeometric function with parameters \( a = \{a_1, \ldots, a_p\} \), \( b = \{b_1, \ldots, b_q\} \), for \( p \geq 0, q \geq 1, \) and \( (r)_k \) denotes the Pochhammer symbol, that is

\[
(r)_k = r(r + 1) \cdots (r + k - 1) = \Gamma(r + k)/\Gamma(r),
\]

see Wolfram (1996, Sec. 3.2.10). In general, an exact formula to compute \( \Sigma_1 \) will not be available, instead we can determine matrix \( C \) from the following asymptotic approximation.

**Proposition 4.1.** Let \( T_p = T_p (L) \) be the statistic defined in (6), in which the order statistic has been obtained from \( n \) iid random variables with distribution (11). We have the following convergences in law

\[
\sqrt{n} [T_p - \mu_n] \xrightarrow{\mathcal{D}} N(0, \sigma_n^2), \quad n \to \infty
\]

(12)

\[
\sqrt{n} \frac{[T_p - \mu_n]}{\sigma_n} \xrightarrow{\mathcal{D}} N(0, 1), \quad n \to \infty
\]

(13)
An Adaptive Goodness-of-Fit Test

where

\[ \mu_n = \hat{\lambda}_0 \gamma_0 + \sum_{j=1}^{\min(p,q)} \hat{\lambda}_j \gamma_j \frac{2n}{j \pi} \sin \left( \frac{j \pi}{2n} \right) \]  \hspace{1cm} (14)

\[ \sigma_1^2 = \lim_{n \to \infty} \sigma_n^2, \quad \sigma_n^2 = \sum_{j=0}^{p} \sum_{j=0}^{q} \lambda_j \lambda_j \sigma_{n,j} \sigma_{n,j} \quad \sigma_{n,j} = 2 \pi^2 \xi_j \xi_j \sum_{k=1}^{q} \sum_{m=1}^{q} \lambda_k \lambda_m I_{jkm}, \]  \hspace{1cm} (15)

where \( \xi_{n0} = 1 \), and \( \xi_{nj} = \sqrt{2} (2n/(jn)) \sin(jn/(2n)) \), for \( 1 \leq j \leq p \).

\[ I_{jkm} = \frac{1}{8 \pi^2} \left\{ \frac{1}{(k+j)^2} \left[ \delta_{m-l,k+j} + \delta_{m+l,k+j} \right] \right\}, \quad \text{for } k = j, \]

\[ I_{jkm} = \frac{1}{8 \pi^2} \left\{ \frac{1}{(k-j)^2} \left[ \delta_{m-l,k-j} + \delta_{m+l,k-j} \right] + \frac{1}{(k+j)^2} \left[ \delta_{m-l,k+j} + \delta_{m+l,k+j} \right] \right\}, \]

for \( k \neq j \), and \( \delta \) is Kronecker's delta.

**Proof.** From (2) and (5),

\[ T_p = \frac{1}{n} \sum_{i=1}^{n} \left[ \hat{\lambda}_0 + \sum_{j=1}^{p} \hat{\lambda}_j \frac{2n}{j \pi} \sin \left( \frac{j \pi}{2n} \right) \sqrt{2} \cos \left( \frac{2i-1}{2} \frac{j \pi}{n} \right) \right] x(i) \]

and, defining,

\[ J_{nj}(t) = \sum_{j=0}^{p} \hat{\lambda}_j \xi_{nj} \cos(j \pi t), \]  \hspace{1cm} (16)

where

\[ \xi_{n0} = 1, \quad \xi_{nj} = \sqrt{2} \frac{2n}{j \pi} \sin \left( \frac{j \pi}{2n} \right), \quad 1 \leq j \leq p, \]  \hspace{1cm} (17)

the statistic, \( T_p \), can be written as

\[ T_p = \frac{1}{n} \sum_{i=1}^{n} J_{nj}(t) \left( \frac{1 - 1/2}{n} \right) x(i), \]

where \( J_{nj} \) is a continuous and bounded a.e. \( (F^-) \) function. The natural centering constant \( \mu_n \), (see p. 661 of Shorack and Wellner, 1986), defined as

\[ \mu_n = \int_{0}^{1} J_{nj}(t) F^-(t) dt, \]  \hspace{1cm} (18)

is equal to

\[ \hat{\lambda}_0 \gamma_0 + \hat{\lambda}_0 \sqrt{2} \sum_{k=1}^{q} \gamma_k \int_{0}^{1} \cos(k \pi t) dt + \gamma_0 \sum_{j=1}^{p} \hat{\lambda}_j \xi_{nj} \int_{0}^{1} \cos(j \pi t) dt \]

\[ + \sqrt{2} \sum_{j=1}^{p} \hat{\lambda}_j \xi_{nj} \sum_{k=1}^{q} \gamma_k \int_{0}^{1} \cos(j \pi t) \cos(k \pi t) dt, \]
which is (14). Similarly,

\[
\sigma_n^2 = \int_0^1 \int_0^1 J_n(s)J_n(t)[\min(s, t) - st]dF^-(s)dF^-(t)
\]

\[
= 2\pi^2 \sum_{j=0}^p \sum_{k=1}^p \sum_{l=0}^q \sum_{m=1}^q \lambda_j \lambda_i \zeta_n \zeta_n \kappa m^n_{k,l,m} I_{jklm},
\]

(19)

where

\[
I_{jklm} = \int_0^1 \int_0^1 K(s, t) \cos(j\pi s) \sin(k\pi s) \cos(l\pi t) \sin(m\pi t) dt ds,
\]

where \(K(s, t) = \min(s, t) - st\), \(0 \leq s, t \leq 1\). Defining

\[
\sigma_{n,jt} = 2\pi^2 \zeta_n \zeta_n \sum_{k=1}^q \sum_{m=1}^q \kappa m^n_{k,l,m} I_{jklm},
\]

we have that

\[
\sigma_n^2 = \sum_{j=0}^p \sum_{l=0}^p \lambda_j \lambda_i \sigma_{n,jt}.
\]

Observing that the eigenvalues and orthonormalized eigenfunctions of the integral operator with kernel \(K(s, t)\) are, respectively,

\[
\mu_j = \frac{1}{(jn)^2}, \quad f_j(t) = \sqrt{2} \sin(j\pi t), \quad 0 \leq t \leq 1, \ j \in \mathbb{N},
\]

(see pp. 213–214 of Shorack and Wellner, 1986), we obtain

\[
I_{jklm} = \frac{1}{4} \int_0^1 \int_0^1 K(s, t) \frac{1}{\sqrt{2}} [f_{k-j}(s) + f_{k+j}(s)] \frac{1}{\sqrt{2}} [f_{m-l}(t) + f_{m+l}(t)] dt ds
\]

\[
= \frac{1}{8} \int_0^1 [f_{m-l}(t) + f_{m+l}(t)] \int_0^1 K(s, t) [f_{k-j}(s) + f_{k+j}(s)] ds dt.
\]

For \(k \neq j\),

\[
I_{jklm} = \frac{1}{8} \int_0^1 [f_{m-l}(t) + f_{m+l}(t)] \left[ \frac{1}{(k-j)^2 \pi^2} f_{k-j}(t) + \frac{1}{(k+j)^2 \pi^2} f_{k+j}(t) \right] dt
\]

\[
= \frac{1}{8\pi^2} \left\{ \frac{1}{(k-j)^2} [\delta_{m-l,k-j} + \delta_{m+l,k-j}] + \frac{1}{(k+j)^2} [\delta_{m-l,k+j} + \delta_{m+l,k+j}] \right\},
\]
and for \( k = j \),

\[
I_{jklm} = \frac{1}{8\pi^2} \left\{ \frac{1}{(k + j)^2} \left[ \delta_{m-l,k+j} + \delta_{m+l,k+j} \right] \right\},
\]

where \( \delta \) is Kronecker’s delta. To compute \( \sigma_1^2 = \lim_{n \to \infty} \sigma_n^2 \), substitute

\[
J_n(t) = \lim_{n \to \infty} J_{n,t}(t) = \lambda_0 + \sqrt{2} \sum_{j=1}^{p} \hat{\lambda}_j \cos(j\pi t),
\]

for \( J_{n,t} \) inside the integral in (19), which can be done because \( J_{n,t} \) is a continuous and bounded a.e. \((F^-)\) function. Then,

\[
\sigma_1^2 = \sum_{j=0}^{p} \sum_{l=0}^{p} \hat{\lambda}_j \hat{\lambda}_l \sigma_{jl},
\]

where

\[
\sigma_{jl} = \lim_{n \to \infty} \sigma_{n,jl} = 4\pi^2 \sum_{k=1}^{q} \sum_{m=1}^{q} km \gamma_k \gamma_m I_{jkm}.
\]

The convergence of (12) is obtained applying Theorem 1, pp. 664–665 of Shorack and Wellner (1986). The convergence of (13) is immediate from (12) and the fact that \( \sigma_1^2 = \lim_{n \to \infty} \sigma_n^2 \).

Comparing the expression for \( c(L) = \sigma_1^2 = L'CL \) in (7) with (15), we see that the entries in \( C \) 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 \( \sigma_{n,jl} \).

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, the uniform distribution, which in each case is obtained for a particular value of the parameter. They are defined by the following probability distribution function.

A1. Lehmann alternatives,

\[ F_{\alpha}(x) = x^\alpha, \quad 0 \leq x \leq 1, \quad \alpha > 0. \]

A2. Centered distributions having \( U \)-shaped probability density function, for \( \beta \in (0, 1) \), or wedge-shaped probability density function, 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,

\[ F_\gamma(x) = \frac{x - \gamma}{1 - 2\gamma}, \quad \gamma \leq x \leq 1 - \gamma, \quad 0 \leq \gamma \leq \frac{1}{2}. \]

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

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

As examples of construction of the test for generic alternatives, we have considered the families above for some values of the parameters. For each alternative we determine coefficients \( c_k \) of (11), for \( 0 \leq k \leq q \). Applying the algorithm for a sample size of \( n = 20 \) and a significance level \( \epsilon = 0.05 \) we determine

<table>
<thead>
<tr>
<th>Family</th>
<th>Fourier coeff.</th>
<th>Weights</th>
<th>Critical values</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( x = 1/2 )</td>
<td>( \gamma_0 = \frac{1}{3} )</td>
<td>-0.426894</td>
<td>( c_1 = -0.159319 )</td>
</tr>
<tr>
<td></td>
<td>( \gamma_k = (-1)^k \frac{2\sqrt{2}}{(k\pi)^2} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( 1 \leq k \leq q )</td>
<td>0.353300</td>
<td>( c_2 = -0.407395 )</td>
</tr>
<tr>
<td></td>
<td>( \gamma_5 = -0.0197851 )</td>
<td>-0.036017</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \gamma_0 = 1/2 )</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td>( \beta = 2 )</td>
<td>( \gamma_1 = -0.197286 )</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>( \gamma_2 = 0 )</td>
<td>-0.971767</td>
<td>( c_1 = 0.327609 )</td>
</tr>
<tr>
<td></td>
<td>( \gamma_3 = -0.0448157 )</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \gamma_4 = 0 )</td>
<td>-0.235944</td>
<td>( c_2 = 0.215799 )</td>
</tr>
<tr>
<td></td>
<td>( \gamma_5 = -0.0197851 )</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>A3</td>
<td>( \gamma = 0.15 )</td>
<td>( \gamma_0 = 1/2 )</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>( 1 \leq k \leq q, k \text{ even} )</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \gamma_k = -\frac{2\sqrt{7}}{(k\pi)^2}(1 - 2\gamma), )</td>
<td>0.545746</td>
<td>( c_2 = -0.288662 )</td>
</tr>
<tr>
<td></td>
<td>( 1 \leq k \leq q, k \text{ odd} )</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>A4</td>
<td>( \delta = 0.05 )</td>
<td>( \gamma_0 = 1/2 )</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>( \gamma_k = 0 )</td>
<td>( 1 \leq k \leq q, k \text{ even} )</td>
<td>0.998779</td>
</tr>
<tr>
<td></td>
<td>( \gamma_k = -\frac{4\sqrt{7}}{(k\pi)^2} + \frac{(2\delta - 1)\sqrt{2}}{k\pi} \text{ sin}(k\pi/2), )</td>
<td>0</td>
<td>( c_2 = -0.334052 )</td>
</tr>
<tr>
<td></td>
<td>( 1 \leq k \leq q, k \text{ odd} )</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
### Table 2
Power of the test based on $T_p$, $Q_n$, $D_n$, $W^2_n$, and the UMP test for the A1 family

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$T_p$</th>
<th>$Q_n$</th>
<th>$D_n$</th>
<th>$W^2_n$</th>
<th>UMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.9980</td>
<td>0.4411</td>
<td>0.9970</td>
<td>0.9973</td>
<td>1.0000</td>
</tr>
<tr>
<td>0.5</td>
<td>0.7492</td>
<td>0.1203</td>
<td>0.6550</td>
<td>0.7211</td>
<td>0.9259</td>
</tr>
<tr>
<td>0.75</td>
<td>0.1973*</td>
<td>0.0764</td>
<td>0.1830</td>
<td>0.1987</td>
<td>0.3918</td>
</tr>
<tr>
<td>2</td>
<td>0.8347</td>
<td>0.3984</td>
<td>0.6730</td>
<td>0.7708</td>
<td>0.9185</td>
</tr>
<tr>
<td>3</td>
<td>0.9872</td>
<td>0.8779</td>
<td>0.9910</td>
<td>0.9955</td>
<td>0.9998</td>
</tr>
<tr>
<td>4</td>
<td>0.9991</td>
<td>0.9900</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

### Table 3
Power of the tests based on $T_p$, $Q_n$, $D_n$, and $W^2_n$ for the A2 family

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>$T_p$</th>
<th>$Q_n$</th>
<th>$D_n$</th>
<th>$W^2_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.9447</td>
<td>0.9651</td>
<td>0.8071</td>
<td>0.8597</td>
</tr>
<tr>
<td>0.5</td>
<td>0.6524</td>
<td>0.7238</td>
<td>0.2879</td>
<td>0.2840</td>
</tr>
<tr>
<td>0.75</td>
<td>0.1893*</td>
<td>0.2203</td>
<td>0.0916</td>
<td>0.0905</td>
</tr>
<tr>
<td>2</td>
<td>0.8045</td>
<td>0.7523</td>
<td>0.1288</td>
<td>0.1013</td>
</tr>
<tr>
<td>3</td>
<td>0.9929</td>
<td>0.9955</td>
<td>0.4029</td>
<td>0.5107</td>
</tr>
<tr>
<td>4</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.7361</td>
<td>0.8951</td>
</tr>
</tbody>
</table>

### Table 4
Power of the test based on $T_p$, $Q_n$, $D_n$, and $W^2_n$ for the A3 family

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>$T_p$</th>
<th>$Q_n$</th>
<th>$D_n$</th>
<th>$W^2_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>0.1761*</td>
<td>0.1016</td>
<td>0.0453</td>
<td>0.0387</td>
</tr>
<tr>
<td>0.10</td>
<td>0.6049</td>
<td>0.3609</td>
<td>0.0451</td>
<td>0.0426</td>
</tr>
<tr>
<td>0.15</td>
<td>0.9894</td>
<td>0.8244</td>
<td>0.0677</td>
<td>0.0669</td>
</tr>
<tr>
<td>0.25</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.3775</td>
<td>0.6195</td>
</tr>
<tr>
<td>0.35</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

### Table 5
Power of the test based on $T_p$, $Q_n$, $D_n$, and $W^2_n$ for the A4 family

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>$T_p$</th>
<th>$Q_n$</th>
<th>$D_n$</th>
<th>$W^2_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>0.9619</td>
<td>0.9585</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>0.15</td>
<td>0.8905</td>
<td>0.9309</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>0.25</td>
<td>0.7951</td>
<td>0.7736</td>
<td>0.8817</td>
<td>0.7533</td>
</tr>
<tr>
<td>0.35</td>
<td>0.3934*</td>
<td>0.3097</td>
<td>0.3321</td>
<td>0.1964</td>
</tr>
<tr>
<td>0.45</td>
<td>0.0745*</td>
<td>0.0697</td>
<td>0.0931</td>
<td>0.0752</td>
</tr>
</tbody>
</table>
the vector of coefficients \( L \), which are the components weights, for the test based on \( T_p \), for \( p = 4 \) and \( q = 5 \). We also compute the critical regions for each case. The results are shown in Table 1.

5. Discussion and Concluding Remarks

Tables 2–5 contain the power comparisons of the test based on \( T_p \) with the tests based on \( Q_n \), \( D_n \), and \( W_n^2 \). 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, we have included these results in Table 2 for comparison.

For some values of the parameter in the A2 and A4 families, the power of the test based on \( Q_n \) exceeds that of \( T_p \). According to theory, this should not happen. \( Q_n \) belongs to the class of statistics from which \( T_p \) has been extracted, via an optimization. After checking several possible causes of inaccuracy, we can discard both a too small \( p \) and a too small \( q \), since doubling them does not improve substantially final results. Our best guess as to the cause of this anomaly is that \( T_p \) is obtained as the exact solution to an approximate problem, namely the optimization of an asymptotic power. It appears that these phenomenon appears if the expectation of the alternative distribution coincides with that of the null hypothesis and it becomes more significant for those parameter values for which the variance of the alternative distribution is close to that of the null hypothesis. An asterisk denotes entries in the tables having this problem.

Acknowledgment

The work is supported in part by 2001SGR-00067 (Generalitat de Catalunya).

References


