Journal of Statistical Planning and Inference 108 (2002) 85–97

journal of statistical planning and inference

www.elsevier.com/locate/jspi

A scale-free goodness-of-fit statistic for the exponential distribution based on maximum correlations ☆

J. Fortiana *, A. Grané

Departament d'Estadística, Universitat de Barcelona, Gran Via de les Corts Catalanes 585, 08007 Barcelona, Spain

Abstract

We propose a goodness-of-fit statistic for testing exponentiality based on Hoeffding's maximum correlation. We study its small and large sample properties, we obtain its exact distribution, tables of exact and asymptotic critical values, and some power curves. We compare this statistic with the Gini statistic, the Shapiro–Wilk statistic for exponentiality and the Stephens' modification of the latter.

MSC: 62G10; 62G30; 62E15

Keywords: Goodness-of-fit; Maximum correlation; L-statistics; Exponentiality

1. Introduction

Let F_1 and F_2 be two cdf's having second-order moments. The *Hoeffding maximum correlation* $\rho^+(F_1,F_2)$ is defined as the correlation coefficient corresponding to the bivariate cdf $H^+(x,y)=\min\{F_1(x),F_2(y)\}$, the *upper Fréchet bound* of F_1 and F_2, i.e., the upper bound of the Fréchet class $\mathscr{F}(F_1,F_2)$ of bivariate cdf's with marginals F_1 and F_2, ordered according to their correlation coefficients. H^+ is a singular distribution, having support on the one-dimensional set $\{(x,y)\in\mathbb{R}^2\colon F_1(x)=F_2(y)\}$, and

$$\rho^+(F_1,F_2)=\left(\int_0^1 F_1^-(p)F_2^-(p)\,\mathrm{d}p-\mu_1\,\mu_2\right)\Big/\,\sigma_1\,\sigma_2, \tag{1}$$

☆ Work supported in part by CGYCIT PB96-1004-C02-01 and 1997SGR-00183 (Generalitat de Catalunya).

* Corresponding author.

E-mail addresses: fortiana@cerber.mat.ub.es (J. Fortiana), aurea@porthos.bio.ub.es (A. Grané).

0378-3758/02/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0378-3758(02)00272-0

where F_i^- is the left-continuous pseudoinverse of F_i, $\mu_i = E(F_i)$ and $\sigma_i^2 = \mathrm{var}(F_i)$, $i = 1, 2$ (see, e.g., Cambanis et al., 1976). Here the notation $E(F)$ represents the expected value of any random variable whose cdf is F, and analogously for $\mathrm{var}(F)$. $\rho^+(F_1, F_2)$ is a measure of agreement between F_1 and F_2, since it equals 1 iff $F_1 = F_2$ (a.e). Cuadras and Fortiana (1993) proposed the statistic $\rho^+(F_n, F_0)$ as a qualitative measure of goodness-of-fit of an iid sample, with empirical cdf F_n, to a given distribution F_0. In this paper, we implement this idea in the form of a specific test and study its properties.

Given n random variables iid $\sim F$, we will test

$$\mathrm{H}_0\colon F = F_0 \equiv \mathrm{Exp}(0, \beta), \quad \beta \in \mathbb{R}_+,$$

where $\mathrm{Exp}(0, \beta)$ is the exponential distribution with location parameter $\alpha = 0$ and unknown scale parameter $\beta > 0$, $F_0(x) = 1 - \exp(-x/\beta)$, $x \geqslant 0$. To this end, we propose the statistic

$$Q_n = \frac{s_n}{\bar{y}_n} \rho^+(F_n, F_0), \tag{2}$$

where $\bar{y}_n$, s_n^2 and F_n are the empirical mean, variance and cdf, respectively, of the iid random variables $y_1, \ldots, y_n$. The motivation for this definition is that, as shown below (see Lemma 2.1), we have the equality

$$Q_n = \frac{\sum_{j=1}^n l_{nj}\, y_{(j)}}{\sum_{j=1}^n y_{(j)}}, \tag{3}$$

where $(y_{(1)}, \ldots, y_{(n)})$ are the order statistics of the sample, which enables us to borrow results from the theory of L-statistics.

Our first aim is to obtain the critical values of Q_n. For small sample sizes, we will use its exact distribution under H_0, which is derived in Section 2. For large sample sizes, we will use the asymptotic distribution under H_0 of the auxiliary function

$$L_n^a = s_n \rho^+(F_n, F_0) - a\bar{y}_n \tag{4}$$

which is obtained in Section 3. Next, in Section 4, we compute the power of the test based on Q_n and compare it to those of the Shapiro–Wilk test of exponentiality, the test based on the Stephens' modification of the Shapiro–Wilk statistic, the test based on the Gini statistic and the test based on another statistic also derived from Hoeffding's maximum correlation (See Fortiana and Grané, 1999), designed for testing goodness-of-fit to the two-parameter exponential distribution, denoted here by Q_n^*.

2. Exact distribution of Q_n under H_0

Let $y_1, \ldots, y_n$ be n random variables iid $\sim F_0 = \mathrm{Exp}(0, \beta)$, with empirical distribution function F_n, and let $y_{(1)}, \ldots, y_{(n)}$ be the order statistics.

Lemma 2.1. *The statistic* $T_n = s_n \rho^+(F_n, F_0)$ *is an L-statistic,*

$$T_n = \frac{1}{n} \sum_{j=1}^{n} l_{nj}\, y_{(j)},$$

where

$$l_{nj} = (n-j)\log(n-j) - (n-j+1)\log(n-j+1) + \log(n), \quad 1 \leqslant j \leqslant n, \tag{5}$$

and $0 \log 0 \equiv 0$.

Proof. Using (1) with F_n and F_0, ρ^+ can be written as

$$\rho^+(F_n, F_0) = \frac{1}{\beta s_n} \left(\int_0^1 F_n^-(p) F_0^-(p)\,\mathrm{d}p - \beta \bar{y}_n \right),$$

since the mean and the standard deviation of the exponential distribution is β. The integral in the numerator of ρ^+

$$\begin{aligned}\int_0^1 F_n^-(p) F_0^-(p)\,\mathrm{d}p &= \sum_{i=0}^{n-1} \int_{i/n}^{(i+1)/n} -\beta \log(1-p)\, y_{(i+1)}\,\mathrm{d}p \\ &= -\frac{\beta}{n} \sum_{i=0}^{n-1} y_{(i+1)}[(n-i)\log(n-i) \\ &\quad - (n-i-1)\log(n-i-1) - \log n - 1].\end{aligned}$$

Letting $j = i+1$ we have

$$\begin{aligned}\int_0^1 F_n^-(p) F_0^-(p)\,\mathrm{d}p = \frac{\beta}{n} \Bigg\{ &\sum_{j=1}^{n} [(n-j)\log(n-j) - (n-j+1)\log(n-j+1) \\ &+ \log(n)] y_{(j)} + \sum_{j=1}^{n} y_{(j)} \Bigg\}\end{aligned}$$

and the result follows. □

Proposition 2.1. *The* Q_n *statistic defined in* (2) *does not depend on the scale parameter, and can be written as*

$$Q_n = \frac{\sum_{l=1}^{n-1} a_{nl} z_l}{\sum_{l=1}^{n} z_l}, \tag{6}$$

where $a_{nl} = \log(n/(n-l))$, $l = 1,\ldots,n-1$, *and* $z_1,\ldots,z_n$ *are iid* $\sim \mathrm{Exp}(0,1)$ *random variables.*

Proof. Suppose $x_j = y_j/\beta$, $j = 1,\ldots,n$, so that

$$Q_n = \frac{\sum_{j=1}^{n} l_{nj} x_{(j)}}{\sum_{j=1}^{n} x_{(j)}}, \tag{7}$$

from David (1981, pp. 20–21), for $j=1,\ldots,n$, the transformation $z_j=(n-j+1)(x_{(j)}-x_{(j-1)})$, where, by convention, $x_{(0)}=0$, gives n iid $\sim \mathrm{Exp}(0,1)$ random variables, and the jth order statistic from the standard distribution can be written as

$$x_{(j)} = \sum_{k=1}^{j} (x_{(k)} - x_{(k-1)}) = \sum_{k=1}^{j} \frac{z_k}{n-k+1}. \tag{8}$$

Substituting (8) in (7),

$$Q_n = \frac{\sum_{j=1}^{n} \sum_{k=1}^{j} l_{nj}\, z_k/(n-k+1)}{\sum_{j=1}^{n} \sum_{k=1}^{j} z_k/(n-k+1)} = \frac{\sum_{j=1}^{n} \log(n/(n-j+1))\, z_j}{\sum_{j=1}^{n} z_j}.$$

Define $a_{nj}=\log(n/(n-j+1))$, $j=1,\ldots,n$; then $a_{n1}=0$. We obtain (6) making $l=j-1$ and rearranging terms. □

Let $N = \sum_{j=1}^{n-1} a_{nj} z_j$ and $D = \sum_{j=1}^{n} z_j$, so that $Q_n = N/D$.

Lemma 2.2. *The characteristic function of the vector* (N,D) *is*

$$\varphi_{(N,D)}(t_1,t_2) = \left[\prod_{j=1}^{n-1} (1 - \mathrm{i}(a_{nj}t_1 + t_2))^{-1}\right] (1 - \mathrm{i}t_2)^{-1}$$

with $(t_1,t_2)' \in \mathbb{R}^2$, $\mathrm{i} = \sqrt{-1}$.

Proof. In matrix notation, $N=\mathbf{a}'\mathbf{z}$, $D=\mathbf{1}'\mathbf{z}$, where $\mathbf{a}=(a_{n1},\ldots,a_{n,n-1},0)'$, $\mathbf{z}=(z_1,\ldots,z_n)'$ and $\mathbf{1}=(1,\ldots,1)'$. We also consider the matrix $\mathbf{A}=(\mathbf{a},\mathbf{1})'$. Then,

$$\begin{pmatrix} N \\ D \end{pmatrix} = \mathbf{A}\,\mathbf{z}.$$

From the transformation formula of characteristic functions by the action of an affine map

$$\varphi_{\mathbf{Az}}(\mathbf{t}) = \varphi_{\mathbf{z}}(\mathbf{A}'\mathbf{t}),$$

for $\mathbf{t}=(t_1,t_2)' \in \mathbb{R}^2$. Since $z_1,\ldots,z_n$ are random variables iid $\sim \mathrm{Exp}(0,1)$,

$$\varphi_{\mathbf{z}}(\mathbf{A}'\mathbf{t}) = \left[\prod_{j=1}^{n-1} \varphi_{z_j}((a_{nj}t_1,t_2))\right] \varphi_{z_n}((0,t_2)),$$

where

$$\varphi_{z_j}((w_1,w_2)) = (1-\mathrm{i}\,(w_1+w_2))^{-1}, \quad j=1,\dots,n. \qquad \square$$

Since Q_n is a quotient of linear combinations of iid $\sim$ Exp(0, 1) random variables, to obtain its exact probability density function we can adapt the technique used in Dwass (1961) and Matsunawa (1985).

Proposition 2.2. *The exact probability density function of* Q_n *is given by*

$$f(t) = (n-1)\sum_{j=1}^{n-1} a_{nj}^{n-3} \prod_{k=1,k\neq j}^{n-1} (a_{nj}-a_{nk})^{-1}\,(1-t/a_{nj})^{n-2}\,\mathbf{1}_{(0,a_{nj})}(t), \tag{9}$$

where $a_{nj} = \log(n/(n-j))$, $j=1,\dots,n-1$, *and* $\mathbf{1}$ *denotes an indicator function.*

Proof. Using Lemma 2.2 from Matsunawa (1985), we obtain the characteristic function of (N,D):

$$\varphi_{(N,D)}(t_1,t_2) = \prod_{j=1}^{n-1} a_{nj}^{-1} \sum_{l=1}^{n-1} C_l\, a_{nl}[(1-\mathrm{i}t_2) - ia_{nl}t_1]^{-1}(1-\mathrm{i}t_2)^{1-n},$$

where

$$C_l = \prod_{k=1,k\neq j}^{n-1} \left(\frac{1}{a_{nk}} - \frac{1}{a_{nj}}\right)^{-1}.$$

Inverting this characteristic function, following (2.16) in Matsunawa (1985),

$$f_{(N,D)}(w,s) = \sum_{l=1}^{n-1} C_l \prod_{j=1}^{n-1} a_{nj}^{-1} \mathbf{1}_{(0,+\infty)}(w)\,\mathbf{1}_{(0,s\,a_{nl})}(w)\,\mathrm{e}^{-s}\left(s-\frac{w}{a_{nl}}\right)^{n-2} \Big/ \Gamma(n-1),$$

for $(w,s)\in\mathbb{R}_+^2$. Applying a change of variables and noticing that

$$C_l \prod_{j=1}^{n-1} a_{nj}^{-1} = a_{nj}^{n-3} \prod_{k=1,k\neq j}^{n-1} (a_{nj}-a_{nk})^{-1},$$

we get the probability density function of Q_n,

$$f(t) = \int_{-\infty}^{+\infty} |s|\, f_{(N,D)}(ts,s)\,\mathrm{d}s. \qquad \square$$

We have developed a *Mathematica* program implementing these algorithms. Fig. 1 shows two examples of probability densities of Q_n, and Table 1 in the appendix contains exact critical values of Q_n for a bilateral test at the 5% significance level.

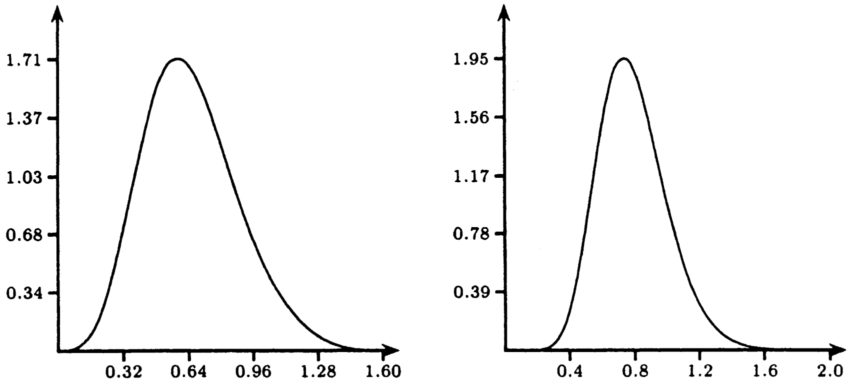

Fig. 1. Probability density functions of Q_n for $n = 5$, on the left, and for $n = 10$, on the right.

3. Asymptotic distribution of L_n^a under H_0

Proposition 3.1. (i) *The function* L_n^a *defined in* (4) *is a linear combination of the order statistics* $x_{(1)}, \ldots, x_{(n)}$ *from the* Exp(0, 1) *distribution*

$$L_n^a = \frac{\beta}{n} \sum_{j=1}^{n} c_{nj}^a x_{(j)}$$

with coefficients

$$c_{nj}^a = l_{nj} - a, \quad 1 \leqslant j \leqslant n$$

and l_{nj} *were defined in* (5).

(ii) *Letting* $\mu_n^a = E(L_n^a)$ *and* $\sigma_{n,a}^2 = n\,\mathrm{var}(L_n^a)$, *we have*

$$\mu_n^a = \beta \left[\log \left(\frac{n^n}{n!} \right)^{1/n} - a \right],$$

$$\sigma_{n,a}^2 = \beta^2 [A_0(n) + A_1(n) a + a^2], \tag{10}$$

where $A_0(n)$ *and* $A_1(n)$ *depend only on* n.

The sequence $\{\sigma_{n,a}^2\}$ *converges to* $\sigma_a^2 = \beta^2(a^2 - 2\,a + 2)$, *when* $n \to \infty$.

(iii) *We have the following convergences in law*:

$$\sqrt{n}[L_n^a - \mu_n^a] \underset{n \to \infty}{\overset{\mathscr{L}}{\to}} \mathrm{N}(0, \sigma_a^2), \tag{11}$$

$$\sqrt{n}\, \frac{[L_n^a - \mu_n^a]}{\sigma_{n,a}} \underset{n \to \infty}{\overset{\mathscr{L}}{\to}} \mathrm{N}(0, 1). \tag{12}$$

Proof. (i) From Lemma 2.1

$$L_n^a = \frac{1}{n}\sum_{j=1}^{n} l_{nj}\, y_{(j)} - a\bar{y}_n = \frac{1}{n}\sum_{j=1}^{n} c_{nj}^a y_{(j)} = \frac{\beta}{n}\sum_{j=1}^{n} c_{nj}^a x_{(j)}$$

with

$$c_{nj}^a = l_{nj} - a, \quad 1 \leqslant j \leqslant n.$$

(ii) From the general asymptotic theory of L-statistics (Shorack and Wellner, 1986), L_n^a can be written as

$$L_n^a = \int_0^1 J_n^a(t)\, F_n^-(t)\, \mathrm{d}t,$$

where

$$J_n^a(t) = n(1-t)\log\left(\frac{1-t}{1-t+1/n}\right) + \log\left(\frac{1}{1-t+1/n}\right) - a$$

and F_n^- is the pseudoinverse of the empirical distribution function. The expectation of L_n^a under H_0 can be computed as

$$\mu_n^a = \int_0^1 J_n^a(t)\, F_0^-(t)\, \mathrm{d}t,$$

where $F_0^-(t) = -\beta\log(1-t)$, $0 \leqslant t \leqslant 1$. Alternatively, it can be computed directly from the expression in (i)

$$\mu_n^a = \frac{\beta}{n}\sum_{j=1}^{n}(l_{nj} - a)m_j$$

where $m_j = \sum_{k=1}^{j}(n-k+1)^{-1}$ is the expected value of the jth order statistic from the standard exponential distribution. See Fortiana and Grané (1999) for more details.

The variance of L_n^a is

$$\mathrm{var}(L_n^a) = \sigma_{n,a}^2/n,$$

where

$$\sigma_{n,a}^2 = \int_0^1\int_0^1 J_n^a(s)J_n^a(t)[\min(s,t) - st]\,\mathrm{d}F_0^-(s)\,\mathrm{d}F_0^-(t). \tag{13}$$

Since the integrand is symmetrical,

$$\sigma_{n,a}^2 = 2\,\beta^2 \int_0^1 \left(J_n^a(s)\int_0^s J_n^a(t)\,\frac{t}{1-t}\,\mathrm{d}t\right)\,\mathrm{d}s$$

and, partitioning the region $\{(s,t)\in\mathbb{R}^2 : 0 \leqslant s \leqslant 1, 0 \leqslant t \leqslant s\}$ in three parts,

$$\mathrm{A} = \{(s,t)\in\mathbb{R}^2\colon 0 \leqslant s \leqslant 1-1/n,\ 0 \leqslant t \leqslant s\},$$

$$\mathrm{B} = \{(s,t)\in\mathbb{R}^2\colon 1-1/n \leqslant s \leqslant 1,\ 0 \leqslant t \leqslant 1-1/n\},$$

$$\mathrm{C} = \{(s,t)\in\mathbb{R}^2\colon 1-1/n \leqslant s \leqslant 1,\ 1-1/n \leqslant t \leqslant s\},$$

we consider the following expansions:

$$\log(1-s+1/n) \approx \log(1-s) + \sum_{k=1}^{m} \frac{(-1)^{k+1}}{k}(n(1-s))^{-k}, \tag{14}$$

when $s < 1-1/n$, and

$$\log(1-s+1/n) \approx \log(1/n) + \sum_{k=1}^{m} \frac{(-1)^{k+1}}{k}(n(1-s))^{k}, \tag{15}$$

when $s \geqslant 1-1/n$. We have used *Mathematica* to compute the integrals on the three regions, obtaining $A_0(n)$ and $A_1(n)$ in (10). To attain the intended precision, (14) was expanded up to $m=2$ and (15) up to $m=5$.

Since J_n^a is continuous and bounded a.e. (F_0^-), we can permute the limit with the integral. Thus, to compute $\sigma_a^2 = \lim_{n\to\infty}\sigma_{n,a}^2$, we just need to substitute

$$J^a(t) = \lim_{n\to\infty} J_n^a(t) = -[1+\log(1-t)+a]$$

for J_n^a in integral (13). Analogously, we have

$$\begin{aligned}\sigma_a^2 &= 2\,\beta^2 \int_0^1 \left(J^a(s) \int_0^s J^a(t)\,\frac{t}{1-t}\,\mathrm{d}t \right) \mathrm{d}s \\ &= 2\,\beta^2 \int_0^1 \left([1+\log(1-s)+a] \int_0^s [1+\log(1-t)+a]\,\frac{t}{1-t}\,\mathrm{d}t \right) \mathrm{d}s \\ &= 2\,\beta^2(1-a+a^2/2).\end{aligned}$$

(iii) Convergence (11) is obtained from Theorem 1 of Shorack and Wellner (1986, pp. 664–665). Convergence (12) is immediate from (11) and from the fact that $\sigma_a^2 = \lim_{n\to\infty}\sigma_{n,a}^2$. □

To compute critical values we have used the normal approximation based on (12) which, having a higher convergence rate than (11), gives a more powerful test. This normal approximation is rather good (for example for $n=20$ we obtain a relative error of 0.6% on the right tail and a relative error of 1.8% on the left tail). On the other hand, while it is actually possible to compute the exact critical values from (9), the computational cost required for solving equations involving piecewise polynomial functions of degree n increases steeply with the sample size.

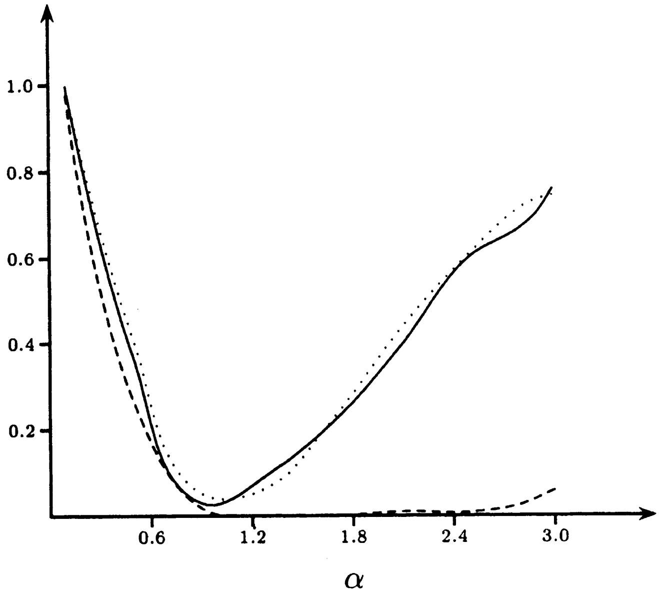

Fig. 2. Power curves of Q_n for A2 family of alternatives, with $\beta = 1$, (see Section 4.1 for notation) from three different critical regions (c.r.).

Fig. 2 shows the power of the test based on Q_n for the A2 family of alternatives, see Section 4.1 below for notation, computed from three different critical regions: the one computed from (9), the one computed from (12) and the one computed from (11).

The critical value a is obtained by solving the equation:

$$-\sqrt{n}\,\frac{\mu_n^a}{\sigma_{n,a}} = c_\varepsilon,$$

where c_ε is the $(1-\varepsilon)$-quantile of the standard normal distribution. We have developed a set of programs in *Mathematica* to compute these critical values. We reproduce some of them in Table 2 in the appendix.

4. Power of Q_n

4.1. Families of alternatives

We consider alternatives of the form $F(x;\theta_1,\theta_2)$. In each case, one of θ_1,θ_2 is the scale parameter and can be fixed to a given arbitrary value, without loss of generality.

(A1) Generalized Pareto

$$F(x;a,k) = 1 - \left(1 - \frac{k}{a}x\right)^{1/k}, \quad a > 0,$$

with $0 \leqslant x < \infty$, if $k \leqslant 0$, and $0 \leqslant x \leqslant a/k$, if $k > 0$. Observe that $\lim_{k\to 0} F(x;a,k) = \mathrm{Exp}(0,a)$. See Choulakian and Stephens (1997) for notation and relations.

(A2) Gamma with shape parameter α and scale parameter β, with density function

$$f(x;\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}}\, x^{\alpha-1}\,\mathrm{e}^{-x/\beta}, \quad \alpha,\beta > 0, x \geqslant 0.$$

In this case, $F(x;1,\beta) = \mathrm{Exp}(0,\beta)$.

(A3) Weibull with shape parameter α and scale parameter β, with density function

$$f(x;\alpha,\beta) = \frac{\alpha}{\beta^{\alpha}}\, x^{\alpha-1}\exp\{-(x/\beta)^{\alpha}\}, \quad \alpha,\beta > 0, x \geqslant 0.$$

In this case, $F(x;1,\beta) = \mathrm{Exp}(0,\beta)$.

(A4) Generalized exponential with location parameter α and scale parameter β.

$$F(x;\alpha,\beta) = 1 - \mathrm{e}^{-(x-\alpha)/\beta}, \quad x \geqslant \alpha,$$

with $\beta > 0$, $-\infty < \alpha < +\infty$. Note that $F(x;0,\beta) = \mathrm{Exp}(0,\beta)$.

4.2. Power comparisons

Power against an alternative distribution $F(x;\theta_1,\theta_2)$ has been estimated by the relative frequency of values of the statistic in the critical region for $N = 500$ simulated n-samples of $F(x;\theta_1,\theta_2)$. For $(x;\theta_1,\theta_2)$ we have taken distributions in each of the A1–A4 families with one fixed parameter and let the other vary within its range. For each family we have taken 30 different values of the parameter.

We have compared Q_n with the Shapiro–Wilk statistic for exponentiality

$$W = \frac{n(\bar{x} - x_{(1)})^2}{(n-1)\sum_{i=1}^{n}(x_i - \bar{x})^2}, \tag{16}$$

the Stephens' modification of the Shapiro–Wilk statistic

$$W_{\mathrm{S}} = \frac{(\sum_{i=1}^{n} x_i)^2}{n(n+1)\sum_{i=1}^{n} x_i^2 - n(\sum_{i=1}^{n} x_i)^2}, \tag{17}$$

the Gini statistic

$$G = \frac{\sum_{i=1}^{n}(2i-n-1)x_{(i)}}{n(n-1)\bar{x}} \tag{18}$$

and

$$Q_n^* = \frac{\sum_{i=1}^{n} l_i x_{(i)}}{\sum_{i=1}^{n} b_i x_{(i)}}, \tag{19}$$

where, for $i = 1,\ldots,n$, $l_i = (n-i)\log(n-i) - (n-i+1)\log(n-i+1) + \log(n)$, with $0\log 0 = 0$, and $b_i = i/n - (n+1)/(2n)$. This statistic, also derived from Hoeffding's maximum correlation, was designed to test for the two-parameter exponential family (see Fortiana and Grané, 1999). Its distribution is invariant under both translations and dilations. For (16) we have used the critical values computed in Shapiro and Wilk

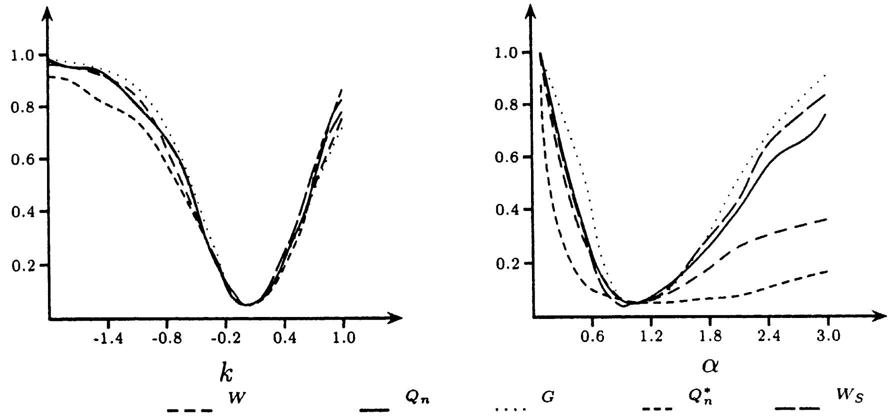

Fig. 3. Power curves for the A1 family ($a = 1$), on the left, and for the A2 family ($\beta = 1$), on the right.

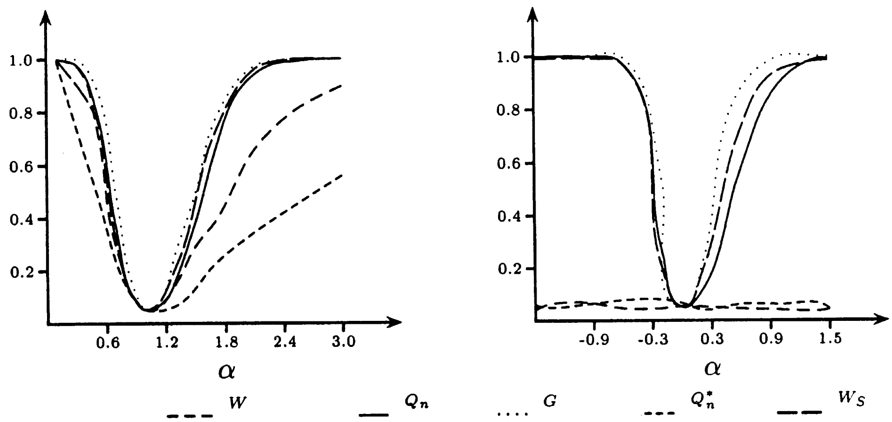

Fig. 4. Power curves for the A3 family ($\beta = 1$), on the left, and for the A4 family ($\beta = 1$), on the right.

(1972), for (17) the points computed in Stephens (1986), for (18) the points computed in Gail and Gastwirth (1978) and for (19) the points computed in Fortiana and Grané (1999).

4.3. Power curves

Figs. 3 and 4 show the power functions of the bilateral tests at the 5%-significance level, for $n = 20$. Note that for the A4 alternative W and Q_n^* are not suitable, because both statistics are location and scale-free. Q_n is more powerful than W and Q_n^* for all the families studied. When the parameter is greater than that of the null hypothesis Q_n is less powerful than G and W_S, while for smaller values of the parameter the power of Q_n is between G and W_S.

Appendix A. Critical values of Q_n are given in Tables 1 and 2

Table 1
Exact critical values of Q_n, for $\varepsilon = 0.05$

n	Lower	Upper	n	Lower	Upper
5	0.254651	1.155460	11	0.454907	1.255941
6	0.304333	1.194334	12	0.474452	1.258626
7	0.344955	1.218537	13	0.491963	1.260081
8	0.378879	1.234278	14	0.507775	1.260478
9	0.407814	1.244687	15	0.522151	1.271918
10	0.432889	1.251537	20	0.578551	1.254820

Table 2
Critical values of Q_n, for $\varepsilon = 0.05$, computed with the asymptotic approximation

n	Lower-tail	Upper-tail
15	0.510208	1.307932
20	0.568069	1.262297
25	0.609831	1.237242
30	0.640928	1.220525
35	0.665574	1.208114
40	0.685736	1.198272
45	0.702634	1.190129
50	0.717065	1.183190
55	0.729578	1.177151
60	0.740566	1.171810
65	0.750315	1.167029
70	0.759043	1.162704
75	0.766918	1.158761
80	0.774069	1.155141
85	0.780602	1.151799
90	0.786601	1.148698
95	0.792135	1.145807
100	0.797262	1.143103
125	0.818219	1.131754
150	0.833809	1.122967
175	0.845993	1.115870
200	0.855854	1.109965

References

Cambanis, S., Simons, G., Stout, W., 1976. Inequalities for $\mathscr{E}k(x, y)$ when the marginals are fixed. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 36, 285–294.

Choulakian, V., Stephens, M., 1997. Goodness-of-fit tests for the generalized Pareto distribution, preprint.

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

David, H.A., 1981. Order Statistics, 2nd Edition. Wiley, New York.
Dwass, M., 1961. The distribution of linear combinations of random divisions of an interval. Trabajos de Estadística e Investigación Operativa 12, 11–17.
Fortiana, J., Grané, A., 1999. A location and scale-free goodness-of-fit statistic for the exponential distribution based on maximum correlations, preprint.
Gail, M., Gastwirth, J., 1978. A scale-free goodness-of-fit test for the exponential distribution based on the Gini statistic. J. Roy. Statist. Soc. B 40, 350–357.
Matsunawa, T., 1985. The exact and approximate distributions of linear combinations of selected order statistics from a uniform distribution. Ann. Inst. Statist. Math. 37, 1–16.
Shapiro, S., Wilk, M., 1972. An analysis of variance test for the exponential distribution complete samples. Technometrics 14, 355–370.
Shorack, G.R., Wellner, J.A., 1986. Empirical Processes with Applications to Statistics. Wiley, New York.
Stephens, M.A., 1986. Tests based on regression and correlation. In: Stephens, M.A., D'Agostino, R.B. (Eds.), Goodness-of-fit Techniques. Marcel Dekker, New York, pp. 195–233.