Statistics & Probability Letters 13 (1992) 365–372
North-Holland

7 April 1992

Tests of fit using spacings statistics with estimated parameters

Martin T. Wells

Department of Economic and Social Statistics, Cornell University, Ithaca, NY 14851-0952, USA

Sreenivasa R. Jammalamadaka

Department of Statistics and Applied Probability, University of California, Santa Barbara, CA 93106, USA

Ram C. Tiwari

Department of Mathematics, University of North Carolina, Charlotte, NC 28223, USA

Received April 1990
Revised July 1991

Abstract: Let $X_1, \ldots, X_n$ be a sequence of independent and identically distributed random variables with an unknown underlying continuous cumulative distribution function F. Relative to this unknown distribution function suppose one would like to test a null hypothesis concerning the goodness of fit of F to some distribution function using symmetric functions of sample spacings. In some applications the null hypothesis is simple while in others it may be composite. In this article we present the large sample theory of tests based on symmetric functions of sample spacings under composite null hypotheses and contiguous alternatives. It is shown that these test statistics have the same asymptotic distribution in the case when parameters must be estimated from the sample as in the case when parameters are specified. Optimal goodness of fit tests are also constructed for these hypotheses.

Keywords: Spacings tests, goodness of fit tests, nuisance parameters, optimal tests, Pitman efficiency.

1. Introduction

Suppose we have a sample of independent observations $X_1, \ldots, X_{n-1}$ on $\mathbb{R}$ from the family of absolutely continuous distributions given by $\mathscr{F}(\beta_0, \theta) = \{F(x; \beta_0, \theta): \theta \in \Theta\}$ where β_0 is a p-dimensional column vector of *known or specified* parameters and θ is a q-dimensional column vector of *unknown* parameters, belonging to a given subset Θ of $\mathbb{R}^q$. One is often interested in testing whether the true distribution function F of the i.i.d. sequence $\{X_n\}$ belongs to the family $\mathscr{F}(\beta_0, \theta)$, that is, we wish to test the following composite null hypothesis: H_0: $F \in \mathscr{F}(\beta_0, \theta)$. Tests of H_0 based on the empirical distribution function type statistics have been discussed, for example, by Durbin (1973), for the χ^2-type statistics, for instance, by Moore and Spurill (1975). See D'Agostino and Stephens (1986) for a good review on the theory of goodness of fit test in the presence of nuisance parameters. Our aim here is to discuss the asymptotic distribution theory of test statistics based on symmetric functions of spacings, under H_0 as well as under a sequence of contiguous alternatives $\{A_n\}$. Spacings tests of H_0 have been considered recently by Cheng and Stephens (1989) using Moran's statistics. It was shown that the test statistics have the same asymptotic distribution when parameters must be estimated from the sample as they do when the parameters are known. In this note we will show this phenomena is true for a large class of statistics. Furthermore, we will derive the test which is locally most powerful in this class.

When $\theta = \theta_0$ is a specified value, we can define the one-step uniform spacings as $D_k = F(X_{(k)};\ \beta_0, \theta_0) - F(X_{(k-1)};\ \beta_0, \theta_0)$ $(k = 1,\dots,n)$ where $X_{(k)}$ is the kth order statistic from a sample of size $n-1$. Let $h(\cdot)$ be real valued functions satisfying some regularity condition. Consider the symmetric spacings statistic

$$T_n = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} h(nD_k). \tag{1.1}$$

Note that if $h(x) = \log x$, T_n is Moran's statistic. See D'Agostino and Stephens (1986) for a good review on spacings statistics. In the case where the function in (1.1) is also a function of the index of summation, say $\{h_k(\cdot),\ k = 1,\dots,n\}$, then the statistic at hand is called a 'nonsymmetric' spacings statistic.

Rao and Sethuraman (1975) study T_n through the weak convergence of the empirical spacings process. They show that the class of symmetric tests discriminate alternatives converging to H_0 at a rate faster than $n^{-1/4}$. Hence T_n has poor asymptotic performance as compared to, say, the Kolmogorov–Smirnov, Cramer–von Mises and χ^2 tests for goodness of fit. It is also shown that $h(x) = x^2$ is asymptotically most powerful in this class. Holst and Rao (1981) investigated nonsymmetric spacings statistic and have found that the class of nonsymmetric tests can discriminate alternatives converging to H_0 at a rate of $n^{-1/2}$, as in the Kolmogorov–Smirnov, Cramer–von Mises and χ^2 tests.

One can also define higher-order uniform spacings or m-step uniform spacings, namely $D_k^{(m)} = F(X_{(k+m-1)};\ \beta_0, \theta_0) - F(X_{(k-1)};\ \beta_0, \theta_0))$ and discuss the associated asymptotic distribution theory (see e.g. Cressie, 1976; Kuo and Rao, 1981; Hall, 1986). For the sake of simplicity, we restrict our discussion to one step spacings. The results of this paper could directly be generalized to statistics based on higher order spacings. It will not be done here.

In the next section, we will discuss the limit theory for the estimated version of T_n. We will also derive the uniformly most powerful test. The proofs of the results are deferred to the appendix.

2. Limit theory for symmetric functions of spacings

Consider the parametric family of distribution functions given by $\mathscr{F}(\beta_0, \theta)$. It is desired to test the null hypothesis H_0: $f \in \mathscr{F}(\beta_0, \theta)$ against the sequence of alternatives, A_n: $F \in \mathscr{F}(\beta_n, \theta)$. These alternatives will be discussed further in Assumption A3. We will discuss the distribution theory of $\hat{T}_n$ under the sequence of alternatives $\{A_n\}$. Since θ is unknown the usual probability integral transform cannot be applied. Instead, define $\hat{U}_{(k)} = F(X_{(k)};\ \beta_n, \hat{\theta}_n)$, where $X_{(k)}$ denotes the kth order statistic from the sample of the X's and $\hat{\theta}_n$ is an estimate of θ (see Assumption A4). Define $\hat{D}_k = \hat{U}_{(k)} - \hat{U}_{(k-1)}$ $(k = 1,\dots,n)$ as the one-step spacing with the estimated parameters. Define the symmetric spacings statistic with estimated parameters as

$$\hat{T}_n = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} h\left(n\hat{D}_k\right). \tag{2.1}$$

We will need the following assumptions to prove the desired results. See Kuo and Rao (1981) for further details on this class of functions $\{h(\cdot)\}$. Assumptions A3–A5 are the usual type of assumptions needed to study goodness of fit problems in the presence of nuissance parameters. See for instance, Durbin (1973) and Moore and Spurill (1975) for more discussion on these regularity conditions.

Assumption A1. Assume the function $h(\cdot)$ is twice differentiable.

Assumption A2. (without loss of generality) $E(h(W)) = 0$, where W is an exponential random variable with mean equal to one.

Assumption A3. Let $n^{1/4}(\beta_n - \beta_0) \to \gamma$ as $n \to \infty$ for some p-dimensional vector γ.

A typical example of the setup we are using is, when one is testing the null hypothesis that a sample is normally distributed with unknown mean θ, but with a variance specified to equal β_0. An alternative which one may be interested in is the shift on the variance parameter, $\beta_n + \beta_0 + \gamma/n^{1/4}$. Having the asymptotic distribution theory under $\{A_n\}$ will allows us to discuss asymptotic power and efficiency of the testing procedure.

Let Λ denote the closure of a given neighborhood of θ_0, the true unknown value of θ, and of β_0, the value of β specified under H_0. Concerning the sequence of estimators of θ assume:

Assumption A4. Under the sequence of alternatives $\{A_n\}$ the estimator of the nuissance parameter is such that $\sqrt{n}(\hat{\theta}_n - \theta_0) = O_p(1)$. Denote $\xi_n = \sqrt{n}(\hat{\theta}_n - \theta_0)$. Assume $E|\xi_n|^3 < \infty$ for all n.

Assumption A5. (i) Let the vector valued functions $g_\beta(u; \beta, \theta) = (\partial/\partial\beta)F(x; \beta, \theta)$ and $g_\theta(u; \beta, \theta) = (\partial/\partial\theta)F(x; \beta, \theta)$ be uniformly continuous in $u \in (0, 1)$ for all $(\beta, \theta) \in \Lambda$, here the right hand side of each of these functions are expressed as a function of u by mean of the transformation $u = F(x; \beta, \theta)$.

(ii) The functions g_β and g_θ are uniformly bounded in u for $(\beta, \theta) \in \Lambda$. Also $(\partial/\partial\alpha)g_\alpha$, $\alpha = \beta, \theta$ are uniformly bounded in u for $(\beta, \theta) \in \Lambda$.

(iii) The functions $\dot{g}_\alpha(u; \beta, \theta) = (d/du)g_\alpha(u; \beta, \theta)$, for $\alpha = \beta$, θ, are uniformly continuous in u for $(\beta, \theta) \in \Lambda$.

Assumption A6. Assume that first and second partial differentiation of $\int f(x; \beta, \theta)\,dx$, with respect to the components of $\beta$and θ, may be passed under the integral sign.

Define the function $\phi: \mathbb{R} \times \mathbb{R}^p \times \Theta \to \mathbb{R}$ implicitly by $F(x; \beta, \theta) = F(\phi(x, \beta, \theta))$. In the case of location and scale parameters, $\phi(x, \beta, \theta) = (x - \beta)/\theta$. From Assumption A5, we know that

$$(\partial^{i+j}/\partial\beta^i\,\partial\theta^j)F(x; \beta, \theta) = f(\phi(x, \beta, \theta))(\partial^{i+j}/\partial\beta^i\,\partial\theta^j)\phi(x, \beta, \theta)$$
$$(i = 1, 2,\ j = 1, 2,\ i + j \leqslant 2)$$

exists and are finite. We will use the notation $L(x)|_a^b = L(b) - L(a)$.

Recall T_n is the symmetric spacings statistic where the distribution function $F(\cdot; \beta, \theta)$ is completely specified, in our case, say $F(\cdot; \beta_0, \theta_0)$. Let $D_k = F(X_{(k)}; \beta_0, \theta_0) - F(X_{(k-1)}; \beta_0, \theta_0)$ and T_n be as in (1.1). The asymptotic behavior of the spacings statistic for the simple hypothesis is given by the following result of Kuo and Rao (1981).

Theorem 2.1. *Let $h(\cdot)$ satisfy Assumptions* A1 *and* A2. *Then $T_n \xrightarrow{D} T$ where $T \sim N(0, \sigma^2)$, $\sigma^2 = \text{Var}(h(W)) - \text{Cov}^2(h(W), W))$ and W is an exponential random variable with mean equal to one.* □

Upon expanding $\hat{T}_n$ in a Taylor series about θ_0 and β_0 we have

$$\hat{T}_n = T_n + \sqrt{n}\left(\hat{\theta}_n - \theta_0\right)^t \Psi_{1n} + \sqrt{n}\,(\beta_n - \beta_0)^t \Psi_{2n} + \tfrac{1}{2}\sqrt{n}\,(\beta_n - \beta_0)^t \Psi_{22n}(\beta_n - \beta_0) + o_p(1), \tag{2.2}$$

where

$$\begin{aligned}\Psi_{1n} &= \frac{1}{n}\sum_{k=1}^{n} \frac{\partial}{\partial\theta} h(nD_k)\\ &= \frac{1}{n}\sum_{k=1}^{n} nh'(nD_k)\cdot\frac{\partial}{\partial\theta}\left[F(X_{(k)}; \beta_0, \theta_0) - F(X_{(k-1)}; \beta_0, \theta_0)\right]\\ &= \frac{1}{n}\sum_{k=1}^{n} (nD_k)h'(nD_k)\dot{g}_\theta\big(\tilde{U}_{1k}\big) \quad \text{for } U_{(k-1)} \leqslant \tilde{U}_{1k} \leqslant U_{(k)},\end{aligned} \tag{2.3}$$

$$\Psi_{2n}=\frac{1}{n}\sum_{k=1}^{n}(nD_k)h'(nD_k)\dot{g}_\beta(\tilde{U}_{2k})\quad\text{for } U_{(k-1)}\leqslant\tilde{U}_{2k}\leqslant U_{(k)},\tag{2.4}$$

$$\Psi_{22n}=\frac{1}{n}\sum_{k=1}^{n}(nD_k)h''(nD_k)\dot{g}_\beta^2(\tilde{U}_k)+(nD_k)h(nD_k)\dot{g}_{\beta\beta}(\tilde{U}_k),\tag{2.5}$$

$$\dot{g}_{\beta\beta}(\tilde{U}_k)=\frac{\partial^2}{\partial\beta^2}g_\beta(u)\,|_{u=\tilde{U}_k}\quad\text{for } U_{(k-1)}\leqslant\tilde{U}_k\leqslant U_{(k)}.\tag{2.6}$$

Using the notation $\gamma_n=n^{1/4}(\beta_n-\beta_0)$ and $\xi_n=\sqrt{n}(\hat{\theta}_n-\theta_0)$, (2.2) may be written as

$$\hat{T}_n=T_n+\xi_n^t\Psi_{1n}+n^{1/4}\gamma_n^t\Psi_{2n}+\gamma_n^t\Psi_{22n}\gamma_n+o_p(1).\tag{2.7}$$

To simplify (2.7) we need the identities stated in Lemma A.1 of the appendix. The next lemma examines the second, third and fourth terms on the right hand side of (2.7).

Lemma 2.2. *Under Assumptions* A1–A6 *and the sequence of alternatives* $\{A_n\}$:

(i) $n^{1/4}\gamma_n^t\Psi_{2n}\xrightarrow{P}0$ *as* $n\to\infty$;

(ii) $\xi_n^t\Psi_{1n}\xrightarrow{P}0$ *as* $n\to\infty$;

(iii) $\gamma_n^t\Psi_{22n}\gamma_n\xrightarrow{P}E(W^2h''(W))\int_0^1(\gamma^t\dot{g}_\beta(u))^2\,du$ *as* $n\to\infty$.

Now we have our main result by applying Lemma 2.2 and Theorem 2.1 to (2.7). The theorem states that the test statistics with and without estimated nuisance parameters have the same asymptotic behavior. Therefore, the estimation of the nuisance parameters has no effect on the asymptotic distribution theory.

Theorem 2.3. *Under Assumptions* A1–A6 *and under the sequence of alternatives* $\{A_n\}$, *it follows that* $\hat{T}_n\xrightarrow{D}\hat{T}\sim N(\mu,\sigma^2)$ *as* $n\to\infty$, *where* $\mu=E(W^2h''(W))\int_0^1(\gamma^t\dot{g}_\beta(u))^2\,du$ *and* $\sigma^2=\mathrm{Var}(h(W))-\mathrm{Cov}^2(h(W),W)$.

Proof. By the representation in (2.7) and an application of Lemma 2.2 it can be seen that $\hat{T}_n=T_n+\mu+o_p(1)$. Thus $\hat{T}_n$ is a translation of T_n in Theorem 2.1, hence the result follows. □

Although it seems unlikely that the test statistics with and without estimated nuisance parameters have the same asymptotic behavior, there is a simple intuitive explanation for this phenomenon. Recall that the type of alternatives under consideration are at a distance of $n^{-1/4}$ away from the null hypothesis. Also recall that the estimates used are $\sqrt{n}$-consistent estimates, hence at a distance proportional to $n^{-1/2}$ away from the true value. Therefore, the test statistic can not distinguish between the estimate and the true value of the nuisance parameter. If the test statistics under consideration could discriminate alternatives at a distance of $n^{-1/2}$, which the test statistics discussed here can not, then one would find that the parameter estimation truly matters. For instance, the Kolmogorov–Smirnov, Cramer–von Mises and χ^2 tests can discriminate alternatives that are at a distance proportional to $n^{-1/2}$ away. For these statistics it is well known that the asymptotic distribution theory for the composite hypothesis and the simple hypothesis are quite different. See for instance, Durbin (1973), Moore and Spurill (1975) and D'Agostino and Stephens (1986) for more discussion.

Similar results may be found for statistics that are functions of the multinomial cell frequencies. If one uses a statistic based on 'symmetric' functions of cell frequencies, the test procedure can only distinguish alternatives at a distance of $n^{-1/4}$ away from the null hypothesis. However, if one uses 'nonsymmetric' functions of cell frequencies, then the test procedure can distinguish alternatives at a distance of $n^{-1/2}$ away from the null hypothesis, see Holst (1972).

In summary, to test H_0 versus A_n one has to estimate the unknown parameter by a $\sqrt{n}$-consistent estimate and use it in a probability integral transform to transform the data to values in [0,1]. Then proceed as if one is testing a simple hypothesis. One may tabulate asymptotic critical values using Theorem 2.3. Once again, as in the case of the simple null hypothesis, the proposed tests will not be as powerful as the ones based on the empirical distribution function.

We now turn our attention to the question of finding the asymptotically locally most powerful test, i.e., the test with the maximum Pitman efficacy, against a specific sequence of alternatives. For the definition of Pitman efficacy see Serfling (1981). For the symmetric statistic based on the estimated spacings we have from Theorem 2.3 the efficacy equals

$$e(h) = E(W^2 h''(W)) \int_0^1 \left(\gamma' \dot{g}_\beta(u)\right)^2 \mathrm{d}u / (\mathrm{Var}(h(W)) - \mathrm{Cov}^2(h(W), W)). \tag{2.8}$$

Our goal is to find a function h such that (2.8) is maximized. That is, we wish to find the test with the maximum efficacy. The following result solves this problem.

Theorem 2.4. *The value of $e(h)$ is maximized by taking $h(x) = x^2$.*

It is important to note that the optimal $h(\cdot)$ does not depend on a specific sequence of alternatives as in the non-symmetric case. It is shown in Holst and Rao (1981) that the asymptotically locally most powerful tests of the simple goodness of fit hypothesis also uses $h(x) = x^2$, leading to the so-called Greenwood statistic. This is no surprise in view of Theorem 3.2. Note that the Moran statistic is not as powerful as the Greenwood statistic in both the simple and composite hypothesis testing problem.

Appendix: Auxiliary results and proofs

The following result may be established by a careful use of the differentiability under the integral sign of $f(x; \beta, \theta)$ and is stated without proof.

Lemma A1. *Suppose Assumption* A6 *holds, then*:

(i) $$f(x; \beta, \theta) \frac{\partial}{\partial \theta} \phi(x, \beta, \theta) \Big|_{x=-\infty}^{x=+\infty} = 0;$$

(ii) $$f(x; \beta, \theta) \frac{\partial}{\partial \beta} \phi(x, \beta, \theta) \Big|_{x=-\infty}^{x=+\infty} = 0;$$

(iii) $$f(x; \beta, \theta) \left(\frac{\partial}{\partial \beta} \phi(x, \beta, \theta) \right)^2 + f(x; \beta, \theta) \frac{\partial^2}{\partial \beta^2} \phi(x, \beta, \theta) \Big|_{x=-\infty}^{x=+\infty} = 0. \qquad \square$$

Proof of Lemma 2.2. (i) Since $\gamma_n \to \gamma < \infty$ we need only examine $n^{1/4} \Psi_{2n}$. From the existence of the limiting distribution of the Kolmogrov–Smirnov statistic

$$\sup_{1 \leqslant k \leqslant n} \sqrt{n} \left| U_{(k)} - k/n \right| = O_p(1).$$

Since $\dot{g}_\beta$ assumed to be continuous, for all $\delta > 0$, we have

$$\sup_{1 \leqslant k \leqslant n} n^{1/2} - \delta \left| \dot{g}_\beta(\tilde{U}_k) - \dot{g}_\beta(k/n) \right| = o_p(1).$$

By the Lebesgue dominated convergence theorem and the fundamental theorem of calculus

$$\frac{1}{n} \sum_{k=1}^{n} \dot{g}_\beta(k/n) \to \int_0^1 \dot{g}_\beta(u) \, du = g_\beta(u) \,|_{u=0}^{u=1} \quad \text{as } n \to \infty.$$

Since

$$g_\beta(u) = \frac{\partial}{\partial \beta} F(x, \beta, \theta) = f(x, \beta, \theta) \frac{\partial}{\partial \beta} \phi(x, \beta, \theta)$$

where $u = F(x, \beta, \theta)$, we have by Lemma A.1(ii),

$$g_\beta(u) \,|_{u=0}^{u=1} = f(x, \beta, \theta) \frac{\partial}{\partial \beta} \phi(x, \beta, \theta) \,|_{x=-\infty}^{x=+\infty} = 0.$$

Hence, by applying Lebesgue dominated convergence theorem and the fundamental theorem of calculus we have

$$S_n = \frac{1}{n} \sum_{k=1}^{n} \dot{g}_\beta(U_{(k)}) \xrightarrow{\text{p}} 0. \tag{A.1}$$

Also since S_n is a function of order statistics it can be shown that

$$\sqrt{n}\, S_n \to \text{N}(0, \tau^2) \quad \text{as } n \to \infty, \tag{A.2}$$

where

$$\tau^2 = \int_0^1 \int_0^1 (\min(s, t) - st) \, d\dot{g}_\beta(s) \, d\dot{g}_\beta(t)$$

(see Shorack, 1972). We have by (A.1) and (A.2),

$$\frac{1}{n} \sum_{k=1}^{n} \dot{g}_\beta(\tilde{U}_k) = O_p(n^{-1/2}).$$

Hence by the Lebesgue dominated convergence theorem it follows

$$n^{-1/4} \Psi_{2n} n^{-3/4} \sum_{k=1}^{n} (nD_k) h(nD_k) \dot{g}_\beta(\tilde{U}_k) \xrightarrow{\text{p}} 0 \quad \text{as } n \to \infty.$$

(ii) Recall from Assumption A4 that $\xi_n = O_p(1)$. The weak law of large numbers and the Lebesgue dominated convergence theorem imply

$$\Psi_{1n} \xrightarrow{\text{p}} \Psi_1 = E(Wh(W)) \int_0^1 \dot{g}_\theta(u) \, du = E(Wh(W)) \, g_\theta(u) \,|_{u=0}^{u=1}.$$

Since

$$g_\theta(u) = \frac{\partial}{\partial \theta} F(x, \beta, \theta) = f(x, \beta, \theta) \frac{\partial}{\partial \theta} \phi(x, \beta, \theta)$$

where $u = F(x, \beta, \theta)$, we have by Lemma A.1(i),

$$g_\theta(u)\,|_{u=0}^{u=1} = f(x, \beta, \theta)\frac{\partial}{\partial\theta}\phi(x, \beta, \theta)\,|_{x=-\infty}^{x=+\infty} = 0.$$

Therefore $\xi_n^{\mathrm{t}}\Psi_{1n} \xrightarrow{\mathrm{P}} 0$.

(iii) By applying Lemma A.1 and a similar proof to the one above one shows

$$\gamma_n^{\mathrm{t}}\Psi_{22n}\gamma_n \xrightarrow{\mathrm{P}} E(W^2h''(W))\int_0^1\left(\gamma^{\mathrm{t}}\dot{g}_\beta(u)\right)^2 \mathrm{d}u + E(Wh(W))\int_0^1\gamma^{\mathrm{t}}\dot{g}_{\beta\beta}(u)\gamma\,\mathrm{d}u. \tag{A.3}$$

The second term on the right hand side of (A.3) equals $E(Wh(W))\gamma^{\mathrm{t}}g_{\beta\beta}(u)\gamma\,|_{u=0}^{u=1}$. Since

$$\begin{aligned} g_{\beta\beta}(u) &= (\partial^2/\partial\beta^2)F(x, \beta, \theta) \\ &= (\partial/\partial\beta)[f(x, \beta, \theta)(\partial/\partial\beta)\phi(x, \beta, \theta)] \\ &= (\partial/\partial\beta)f(x, \beta, \theta)(\partial/\partial\beta)\phi(x, \beta, \theta) + f(x, \beta, \theta)(\partial^2/\partial\beta)\phi(x, \beta, \theta) \\ &= f(x, \beta, \theta)[(\partial/\partial\beta)\phi(x, \beta, \theta)]^2 + f(x, \beta, \theta)(\partial^2/\partial\beta)\phi(x, \beta, \theta) \end{aligned}$$

where $u = F(x, \beta, \theta)$, hence Lemma A.1(iii) implies $g_{\beta\beta}(u)\,|_{u=0}^{u=1} = 0$. Therefore

$$\gamma_n^{\mathrm{t}}\Psi_{22n}\gamma_n \xrightarrow{\mathrm{P}} E(W^2h''(W))\int_0^1\left(\gamma^{\mathrm{t}}\dot{g}_\beta(u)\right)^2 \mathrm{d}u \quad \text{as } n\to\infty. \qquad \square$$

Proof of Theorem 2.4. It is easily established, by integration by parts, that $EW^2h''(W) = \mathrm{Cov}(h(W), (W-2)^2)$. This identity will simplify our task. Consider the nondegenerate statistic $\hat{T}_n(h) = (1/\sqrt{n})\Sigma_{k=1}^n h(n\hat{D}_k)$ with $\mathrm{Var}(\hat{T}_n(h)) = \sigma_h^2$. Note that the efficacy is uneffected by a linear transformation of the statistic. Therefore to find $h(\cdot)$ which maximizes $e(h)$ among all functions satisfying Assumptions A1–A6, one may consider without loss of generality, the class of $\{h\}$ with $\sigma_h^2 = 1$. Thus by the Cauchy–Schwartz inequality,

$$\begin{aligned} e(h) &= \int_0^1\left(\gamma^{\mathrm{t}}\dot{g}_\beta(u)\right)^2 \mathrm{d}u\ \mathrm{Cov}\left(h(W), (W-2)^2\right)/2 \\ &\leqslant \int_0^1\left(\gamma^{\mathrm{t}}\dot{g}_\beta(u)\right)^2 \mathrm{d}u\ [\mathrm{Var}(W)]^{1/2}\left[\mathrm{Var}(W-2)^2\right]^{1/2}/2 \end{aligned}$$

with equality if and only if $h(x) = a(x-2)^2 + b$ for some $a \neq 0$ and b. But since $\mathrm{Cov}(W, (W-2)^2) = 0$, the maximum of $e(h)$ is attained by taking $h(x) = x^2$. $\square$

References

Cheng, R.C.H. and M.A. Stephens (1989), A goodness of fit test using Moran's statistic with estimated parameters, *Biometrika* **76**, 385–392.

Cressie, N. (1976), On the logarithms of high-order spacings, *Biometrika* **63**, 343–355.

D'Agostino. R.B. and M.A. Stephens (1986), *Goodness-of-fit Techniques* (Dekker, New York).

Durbin, J. (1973), Weak convergence of the sample distribution function when parameters are estimated, *Ann. Statist.* **1**, 279–290.

Hall, P. (1986), On powerful distributional tests based on sample spacings, *J. Multivariate Anal.* **19**, 201–224.

Holst, L. (1982), Asymptotic normality and efficiency of certain goodness of fit tests, *Biometrika* **59**, 137–145.

Holst, L. and J.S. Rao (1981), Asymptotic spacings theory with applications to the two sample problem, *Canad. J. Statist.* **9**, 79–89.

Kuo, M. and J.S. Rao (1981), Limit theory and efficiencies for tests based on higher order spacings, in: *Statistics: Applications and New Directions. Proc. of the Golden Jubilee Conf.*

of the Indian Statist. Inst. (Statist. Publ. Soc., Calcutta) pp. 333–352.

Moore, D.S. and M.C. Spurill (1975), Unified large sample theory of general chi-squared statistics for tests of fit, *Ann. Statist.* **3**, 599–616.

Rao, J.S. and J. Sethuraman (1975), Weak convergence of empirical distribution functions of random variables subject to perturbations and scale factors, *Ann. Statist.* **3**, 299–313.

Serfling, R. (1980), *Approximation Theorems for Mathematical Statistics* (Wiley, New York).

Shorack, G.R. (1972), Functions of order statistics, *Ann. Math. Statist.* **52**, 591–611.