Akademie věd České republiky
Ústav teorie informace a automatizace

Academy of Sciences of the Czech Republic
Institute of Information Theory and Automation

RESEARCH REPORT

I. Vajda and E.C. van der Meulen:

Divergences between Models and Data under Hypothetical and Empirical Quantizations

No. 2275 January 2010

ÚTIA AV ČR, P. O. Box 18, 182 08 Prague, Czech Republic
Telex: 122018 atom c, Fax: (+42) (2) 688 4903
E-mail: utia@utia.cas.cz

This report constitutes an unrefereed manuscript which is intended to be submitted for publication. Any opinions and conclusions expressed in this report are those of the author(s) and do not necessarily represent the views of the Institute.

January 17, 2010

On Divergences Between Models and Data Under Hypothetical and Empirical Quantizations

I. VAJDA[1] and E. C. VAN DER MEULEN[2]
[1] *Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod vodárenskou věží 4, 182 08 Prague, Czech Republic. E-mail: vajda@utia.cas.cz*
[2] *Department of Mathematics, Katholieke Universiteit Leuven, Celestijnenlaan 200B, B-3001 Heverlee, Belgium. E-mail: edward.vandermeulen@wis.kuleuven.ac.be*

This paper summarizes the research on goodness-of-fit disparity statistics obtained as appropriately scaled ϕ-divergences or ϕ-disparities of quantized hypothetical and empirical distributions. It is shown that the classical Pearson-type statistics are obtained if we quantize by means of hypothetical quantiles, and modified versions of the spacings-based disparity statistics known from the literature are obtained if we quantize by means of empirical quantiles. The main attention is paid to the asymptotic properties of the new modified disparity statistics and their comparisons with the classical spacings-based statistics known from the literature. First the asymptotic equivalence between both of them is proved. Then for the new statistics are proved the law of large numbers and the asymptotic distributions under the hypothesis and under local and fixed alternatives. Special attention is devoted to the limit laws for the power divergence statistics of orders $\alpha \in \mathbb{R}$. Parameters of these laws are evaluated for $\alpha \in (-1, \infty)$ in a closed form and their continuity in α on the subinterval $(-1/2, \infty)$ is proved. These closed form expressions are used to compare local asymptotic powers of the tests based on these statistics, which allows to extend previous asymptotic optimality results to the class of power divergence statistics. Tables of values of the asymptotic parameters are presented for selected representative orders of $\alpha \geq -1/2$.

Key words: asymptotic normality, asymptotic optimality, consistency, goodness-of-fit, power divergences, spacings, ϕ-disparities, ϕ-divergences

1. Introduction and basic concepts

We consider real-valued independent identically distributed observations $X_1, \ldots, X_n$ with a distribution function $F(x)$ and the problem of testing the hypothesis $\mathcal{H}_0$ that F is a given continuous increasing distribution function F_0. As is well known, we can then assume without loss of generality that the observation space is the interval $\mathcal{X} = (0,1)$ and $F_0(x) = x$ on $\mathcal{X}$. Further, we can restrict ourselves to test statistics T_n which are functions of sufficient statistics. Examples of sufficient statistics are the empirical distribution function

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{I}(x \geq X_i), \quad x \in \mathcal{X} \tag{1.1}$$

where $\boldsymbol{I}$ is the indicator function, and the order statistics

$$0 = Y_0 \leq Y_1 = X_{n:1} \leq \cdots \leq Y_n = X_{n:n} \leq Y_{n+1} = 1 \tag{1.2}$$

where Y_0 and Y_{n+1} are dummy variables and the inequalities are typically strict with probability one.

It is natural to consider test statistics of the form $T = T_n = c_n D(F_0, F_n)$ where c_n is an appropriate scaling constant and $D(F,G)$ is a nonnegative measure of disparity between two distribution functions F and G on $\mathcal{X} = (0,1)$. Let ϕ be a continuous function $\phi : (0,\infty) \mapsto \mathbb{R}$. We shall deal with ϕ-*disparities* of the form

$$D_\phi(F,G) = \int_0^1 g(x)\phi\left(\frac{f(x)}{g(x)}\right) \mathrm{d}x \tag{1.3}$$

when F and G are defined by densities f and g, denoted by $F \sim f, G \sim g$, or of the form

$$D_\phi(\boldsymbol{p},\boldsymbol{q}) = \sum_{j=1}^{k} q_j \phi\left(\frac{p_j}{q_j}\right) \tag{1.4}$$

when F, G are quantized into discrete distributions $\boldsymbol{p} = (p_1, p_2, ..., p_k)$, $\boldsymbol{q} = (q_1, q_2, ..., q_k)$ by an interval partition $\mathcal{P}$ of $\mathcal{X} = (0,1)$ using certain cutpoints

$$0 = a_0 < a_1 \cdots < a_{k-1} < a_k = 1 \quad \text{for} \quad k > 1. \tag{1.5}$$

If the function ϕ is convex then the ϕ-disparities are traditionally called ϕ-*divergences*. For convex ϕ definitions (1.3) and (1.4) represent special cases of general ϕ-divergences introduced for arbitrary probability measures by Csiszár (1963). In this paper we restrict ourselves to ϕ-divergences for ϕ from *the class* $\boldsymbol{\Phi}$ *of convex functions* $\phi : (0,\infty) \mapsto \mathbb{R}$ *which are twice continuously differentiable in a neighborhood of* 1 *with* $\phi''(1) > 0$ *and* $\phi(1) = 0$. Then the integrals $D_\phi(F,G)$ and sums $D_\phi(\boldsymbol{p},\boldsymbol{q})$ in (1.3) and (1.4) are well defined for all distribution functions $F \sim f$, $G \sim g$ and all discrete distributions $\boldsymbol{p}, \boldsymbol{q}$ by applying behind the integral and sum a lower semicontinuous and convex extension of the function $s\phi(t/s)$from the open domain $t, s > 0$ into the closure $t, s \geq 0$. Moreover, then $D_\phi(F,G)$ and $D_\phi(\boldsymbol{p},\boldsymbol{q})$ are nonnegative and are equal to zero if and only if $F = G$ or $\boldsymbol{p} = \boldsymbol{q}$ respectively (for details about the definition of ϕ-divergences and their properties see Liese and Vajda (2006)).

If the function ϕ considered in (1.3) and (1.4) is differentiable at $t = 1$ with $\phi(1) = 0$ and with the difference $\phi(t) - \phi'(1)\,(t-1)$ decreasing for $t \in (0,1)$ and increasing for

$t \in (1, \infty)$, then $D_\phi(F, G)$ and $D_\phi(\boldsymbol{p}, \boldsymbol{q})$ are *ϕ-disparities* in the sense considered by Lindsay (1994), Morales *et al.* (2003) and others cited there. For $\phi \in \boldsymbol{\Phi}$ the difference $\phi(t) - \phi'(1)\,(t-1)$ is automatically decreasing on $(0, 1)$ and increasing on $(1, \infty)$ so that the concept of ϕ-disparity generalizes the concept of ϕ-divergence.

An example of functions $\phi \in \boldsymbol{\Phi}$ is provided by the class of functions $\{\phi_\alpha : \alpha \in \mathbb{R}\}$ defined on $(0, \infty)$ by

$$\phi_\alpha(t) = \frac{t^\alpha - \alpha(t-1) - 1}{\alpha(\alpha-1)} \tag{1.6}$$

for $\alpha \neq 0$, $\alpha \neq 1$, and by the continuous extensions

$$\phi_1(t) = t \ln t - t + 1 \quad \text{and} \quad \phi_0(t) = -\ln t + t - 1 \tag{1.7}$$

otherwise. The corresponding divergences $D_{\phi_\alpha}(F, G)$ and $D_{\phi_\alpha}(\boldsymbol{p}, \boldsymbol{q})$ are denoted by $D_\alpha(F, G)$ and $D_\alpha(\boldsymbol{p}, \boldsymbol{q})$, respectively. The class of divergences $D_\alpha(\boldsymbol{p}, \boldsymbol{q})$ contains the following classical examples: the *quadratic divergence*

$$D_2(\boldsymbol{p}, \boldsymbol{q}) = \frac{1}{2}\chi^2(\boldsymbol{p}, \boldsymbol{q}) = \frac{1}{2}\sum_{j=1}^{k} \frac{(p_j - q_j)^2}{q_j} \tag{1.8}$$

where $\chi^2(\boldsymbol{p}, \boldsymbol{q})$ is also known as *χ^2-divergence*, the *harmonic divergence*

$$D_{-1}(\boldsymbol{p}, \boldsymbol{q}) = D_2(\boldsymbol{q}, \boldsymbol{p}) = \frac{1}{2}\sum_{j=1}^{k} \frac{(p_j - q_j)^2}{q_j}, \tag{1.9}$$

the *logarithmic divergences*

$$D_0(\boldsymbol{p}, \boldsymbol{q}) = D_1(\boldsymbol{q}, \boldsymbol{p}) \quad \text{and} \quad D_1(\boldsymbol{p}, \boldsymbol{q}) = I(\boldsymbol{p}, \boldsymbol{q}) = \sum_{j=1}^{k} p_j \ln \frac{p_j}{q_j} \tag{1.10}$$

where $I(\boldsymbol{p}, \boldsymbol{q})$ is known as the *information divergence* (often denoted also as $D(\boldsymbol{p} \parallel \boldsymbol{q})$), and the *square root divergence*

$$D_{1/2}(\boldsymbol{p}, \boldsymbol{q}) = 4H^2(\boldsymbol{p}, \boldsymbol{q}) = 4\sum_{j=1}^{k} \left(\sqrt{p_j} - \sqrt{q_j}\right)^2 \tag{1.11}$$

where $H(\boldsymbol{p}, \boldsymbol{q})$ is known as *Hellinger distance.*

We admit that the size $k = k_n$ of the interval partition $\mathcal{P} = \{(a_{j-1}, a_j] : 1 \leq j \leq k\}$ of $\mathcal{X}$ introduced in (1.5), and also the cutpoints $a_1, \ldots, a_{k-1}$ themselves, may in general depend on the sample size n, but this dependence is not always explicitly denoted in this paper. Quantizations of F_0 and F_n by means of such partition lead to discrete *hypothetical* and *empirical* distributions

$$\boldsymbol{p}_0 = (p_{0j} : 1 \leq j \leq k) \qquad \text{and} \qquad \boldsymbol{p}_n = (p_{nj} : 1 \leq j \leq k) \tag{1.12}$$

where

$$p_{0j} = F_0(a_j) - F_0(a_{j-1}) = a_j - a_{j-1} > 0 \tag{1.13}$$

and

$$p_{nj} = F_n(a_j) - F_n(a_{j-1}) > 0 \quad \text{a. s.} \tag{1.14}$$

These distributions can serve as arguments of the disparities D_ϕ in (1.4), yielding $D_\phi(\boldsymbol{p}_0, \boldsymbol{p}_n)$, and of the corresponding *ϕ-disparity statistics* T_ϕ defined by

$$T_\phi = T_{\phi,n} := n\, D_\phi(\boldsymbol{p}_0, \boldsymbol{p}_n) = n \sum_{j=1}^{k} p_{nj} \phi\left(\frac{p_{0j}}{p_{nj}}\right). \tag{1.15}$$

In this paper we restrict ourselves to the simplest ϕ-disparity statistics T_ϕ, which are obtained when one of the distributions $\boldsymbol{p}_0, \boldsymbol{p}_n$ in (1.15) is uniform, that is, equal to

$$\boldsymbol{u}_k = (u_{kj} = 1/k : 1 \le j \le k). \tag{1.16}$$

This takes place when the cutpoints a_j of (1.13) or (1.14) are the quantiles

$$a_j = G^{-1}(j/k) = \inf\{x \in (0,1] : G(x) \ge j/k\} \tag{1.17}$$

of the distribution functions $G = F_0$ or $G = F_n$, respectively. Proceeding this way we obtain two versions of $D_\phi(\boldsymbol{p}_0, \boldsymbol{p}_n)$ and T_ϕ.

(I) Applying the rule (1.17) to $G = F_0$ we get the **hypothetical quantiles**

$$a_j = F_0^{-1}(j/k) = j/k, \quad 1 \le j \le k-1 \tag{1.18}$$

leading, according to (1.13) and (1.15), to the **uniform hypothetical distribution** $\boldsymbol{p}_0 = \boldsymbol{u}_k$ and the **frequency-type disparity statistics**

$$T_\phi^{(f)} := n\, D_\phi(\boldsymbol{u}_k, \boldsymbol{p}_n) = n \sum_{j=1}^{k} p_{nj} \phi\left(\frac{1}{k\, p_{nj}}\right) \tag{1.19}$$

where the p_{nj} *'s* are given by (1.14) for the a_j of (1.18) and $a_0 = 0$, $a_k = 1$. We denote the corresponding partition by $\mathcal{P}_0$. In other words

$$p_{nj} = \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{I}_{(a_{j-1}, a_j]}(Y_i) \tag{1.20}$$

is the relative frequency of the observations $Y_1, \ldots, Y_n$ in the cell $(a_{j-1}, a_j] = ((j-1)/k, j/k)]$, $j = 1, \cdots, k$, of the partition $\mathcal{P}_0$.

Example 1.1. A well-known subclass of the frequency-type disparity statistics $T_\phi^{(f)}$ consists of the power divergence statistics

$$T_\alpha^{(f)} = n\, D_\alpha(\boldsymbol{u}_k, \boldsymbol{p}_n) = n \sum_{j=1}^{k} p_{nj}\, \phi_\alpha \left(\frac{1}{k p_{nj}} \right), \quad \alpha \in \mathbb{R} \tag{1.21}$$

systematically studied in Read and Cressie (1988).

Classical examples of such $T_\alpha^{(f)}$ are the Neyman statistic $T_2^{(f)} = n\, D_2(\boldsymbol{u}_k, \boldsymbol{p}_n)$, the Pearson statistic $T_{-1}^{(f)} = n\, D_{-1}(\boldsymbol{u}_k, \boldsymbol{p}_n)$, the log-likelihood ratio statistic $T_0^{(f)} = n\, D_0(\boldsymbol{u}_k, \boldsymbol{p}_n)$, the reversed log-likelihood ratio statistic $T_1^{(f)} = n\, D_1(\boldsymbol{u}_k, \boldsymbol{p}_n)$, and the Freeman–Tukey statistic $T_{1/2}^{(f)} = n\, D_{1/2}(\boldsymbol{u}_k, \boldsymbol{p}_n)$.

(II) The main focus of this paper is on disparity statistics T_ϕ obtained from (1.15) when rule (1.17) is applied to the empirical distribution $G = F_n$, leading to the **empirical quantiles** $a_j = F_n^{-1}(j/k)$. For simplicity we assume that n is divisible by k. Then, using the integers $m = n/k \geq 1$, we get the $k-1$ empirical quantiles

$$a_j = F_n^{-1}(j/k) = Y_{mj}, \quad 1 \leq j \leq k-1 \tag{1.22}$$

and the partition $\mathcal{P} = \mathcal{P}_n$ consisting of the k cells

$$(a_{j-1}, a_j] = (Y_{m(j-1)}, Y_{mj}], \quad 1 \leq j \leq k-1, \qquad (a_{k-1}, a_k] = (Y_{m(k-1)}, 1] \tag{1.23}$$

where $a_0 = Y_{m0} = Y_0 = 0$ (cf (1.2) and (1.5)), leading to the hypothetical distribution $\boldsymbol{p}_0 = (p_{0j} : 1 \leq j \leq k)$ with

$$p_{0j} = Y_{mj} - Y_{m(j-1)} \quad \text{for} \quad 1 \leq j \leq k-1, \quad \text{and} \quad p_{0k} = 1 - Y_{m(k-1)}. \tag{1.24}$$

(Note that here and in the sequel mj, $m(j-1)$ and so forth denote the products of integers and not the pairs of integers as in (1.12)–(1.15) and elsewhere. We believe that the correct meaning of mj can always be recognized. Also note that the order statistic Y_{mk} does not occur as an endpoint in the definition (1.23) of the cells $(a_{j-1}, a_j], 1 \leq j \leq k$.)

Since all cells $(a_{j-1}, a_j]$, $1 \leq j \leq k$, in (1.23) contain exactly m of the observations $Y_1, \ldots, Y_n$, formulas (1.14) and (1.15) lead to the **uniform empirical distribution** $\boldsymbol{p}_n = \boldsymbol{u}_k$ and to the **spacings-type disparity statistics**

$$T_\phi^{(m)} := n\, D_\phi(\boldsymbol{p}_0, \boldsymbol{u}_k) = m \sum_{j=1}^{k} \phi(k p_{0j}) \tag{1.25}$$

where the p_{0j} 's are given by (1.24) with $Y_0 = 0$. The use of the spacings terminology is justified by the fact that, since $k = n/m$, formula (1.25) can be given the form

$$T_\phi^{(m)} = m \sum_{j=1}^{k-1} \phi \left(\frac{n}{m} (Y_{mj} - Y_{m(j-1)}) \right) + m\, \phi \left(\frac{n}{m} (1 - Y_{m(k-1)}) \right) \tag{1.26}$$

where $Y_{mj} - Y_{m(j-1)}$ are m-spacings. For $m = 1$, (1.23) reduces to

$$(a_{j-1}, a_j] = (Y_{j-1}, Y_j], \quad 1 \le j \le n-1, \qquad (a_{n-1}, a_n] = (Y_{n-1}, 1], \tag{1.27}$$

where $a_0 = Y_0 = 0$ (giving rise to a partition $\mathcal{P}^1$), the distribution (1.24) reduces to

$$p_{0j} = Y_j - Y_{j-1} \quad \text{for} \quad 1 \le j \le n-1, \quad \text{and} \quad p_{0n} = 1 - Y_{n-1}, \tag{1.28}$$

and the statistic $T_\phi^{(m)}$ of (1.26) reduces to the simple-spacings-formula

$$T_\phi = \sum_{j=1}^{n-1} \phi\left(n(Y_j - Y_{j-1})\right) + \phi\left(n(1 - Y_{n-1})\right). \tag{1.29}$$

Remark 1.1. Formulas (1.26) and (1.29) employ the dummy observations $Y_0 = 0$ and $Y_{n+1} = 1$ introduced in (1.2). Unless otherwise explicitly stated, these dummy observations are also assumed in the formulas below, notably in (1.30) and (1.32).

It seems that the spacings-based goodness-of-fit test statistics given in the literature lacked sofar the motivation of taking into account the notion of disparity between hypothetical and empirical distributions $\boldsymbol{p}_0$ and $\boldsymbol{p}_n$. This contrasts with the goodness-of-fit statistics based on deterministic partitions derived from the a_j in (1.18) and the related frequency counts (1.20), where the typical statistics, including the most classical Pearson statistic T_1 and likelihood ratio statistic T_0, can easily be recognized as appropriately scaled power divergences between $\boldsymbol{p}_0$ and $\boldsymbol{p}_n$.

The classical spacings-based statistics, however, appear to have been motivated rather by other considerations such as the analytic simplicity of formulas and the possibility to achieve desired asymptotic properties. In fact, as pointed out by Pyke(1965) in his landmark paper, most of the classical spacings-based statistics were proposed within the context of testing the randomness of events in time, in which differences between successive order statistics (spacings) were considered to play an important role. Also, in the period 1946-1953, when most of the classical tests based on spacings were proposed, research focused mostly on studying the behavior of these tests under the null-hypothesis, rather than under an alternative, making it unnecessary to motivate the test statistic from the point of view of divergence or disparity. Although the concept of dispersion of spacings around the uniform distribution has been mentioned as a motivation for a test statistic by some authors, no known spacings-based statistic happens to be the disparity statistic $T_\phi^{(m)}$ of (1.26) or T_ϕ of (1.29) for some $\phi \in \boldsymbol{\Phi}$. This situation is illustrated in the next two examples for the simple-spacings statistics where $m = 1$. Then T_ϕ is of the form $R_\phi + W_\phi$ for

$$R_\phi = \sum_{j=1}^{n+1} \phi\left(n\left(Y_j - Y_{j-1}\right)\right) \tag{1.30}$$

and

$$W_\phi = \phi\left(n(1 - Y_{n-1})\right) - \phi\left(n(Y_n - Y_{n-1}) - \phi\left(n(1 - Y_n)\right)\right), \tag{1.31}$$

while the classical simple-spacings statistics are of the form

$$S_\phi = \sum_{j=1}^{n+1} \phi\left((n+1)\right)(Y_j - Y_{j-1}))\,. \tag{1.32}$$

With reference to the above discussion, we mention here that Pyke(1965) writes that it is more convenient to weight the spacings by $n+1$ instead of n if one is concerned entirely with uniform observations.

Let us now turn to comparing our m-spacings–based disparity statistics $T_\phi^{(m)}$ from (1.25) - (1.26) and the m-spacings–based statistics known from the literature for general $m \geq 1$. We shall start with Del Pino's (1979) class of statistics of the form

$$S_\phi^{(m)} = m \sum_{j=1}^{k} \phi\left(\frac{n+1}{m}(Y_{mj} - Y_{m(j-1)})\right) \tag{1.33}$$

where it is assumed that $n+1$ is divisible by k and that $m = (n+1)/k \geq 1$. Hence the notation in our paper is consistent in the sense that (1.33) reduces for $m = 1$ to the formula for S_ϕ in (1.32). Del Pino found $\phi(t) = t^2$ to be optimal among the functions ϕ considered by him. The class (1.33) was later investigated by Jammalamadaka *et al.* (1989), Jimenez and Shao (2009) and many others cited there. Jimenez and Shao (2009) studied the asymptotics of $S_\phi^{(m)}$ for fixed $m \geq 1$. Jammalamadaka *et al* (1989) studied the asymptotics of $S_\phi^{(m)}$ for m tending slowly to infinity as $n \to \infty$. In such case these asymptotics depend only on the local properties of $\phi(t)$ in the neighborhood of $t = 1$ and a wide class of functions ϕ can be admitted including those with $\phi''(1) = 0$. However, as we have seen in the examples above, even for ϕ from the above introduced ϕ-divergence class $\mathbf{\Phi}$, the statistics (1.33) differ from those in (1.25) or (1.26). Other examples of well-known spacings-based statistics which differ from our spacings-type ϕ-disparity statistics (1.25) and (1.26) will be given in the next section. Therefore it is important to look at the problem whether the classical spacings-based statistics and our spacings-type disparity statistics are asymptotically equivalent for $n \to \infty$, and, if yes, then in what precise sense.

The **first objective** of the present paper is to prove the mutual asymptotic equivalence of the statistics of the two mentioned origins. This equivalence helps to understand why many ad hoc defined spacings-based statistics exhibit desirable asymptotic properties.The **second objective** of this paper is to prove the consistency and asymptotic normality under fixed and local alternatives for a sufficiently wide variety of our spacings-type ϕ-disparity or ϕ-divergence statistics. These results may also be useful in the estimation of functionals of the type of ϕ-disparity or ϕ-divergence. The **last objective** is to apply this asymptotic theory to the spacings-based power divergence statistics and compare their asymptotic parameters and properties for various divergence orders $\alpha \in \mathbb{R}$. To achieve all these objectives on a reasonably limited space, we restrict ourselves in this paper to the simple spacings with $m = 1$.

Sofar we have defined for the case $m = 1$ three different spacings-based statistics, viz. T_ϕ of (1.29), R_ϕ of (1.30), and S_ϕ of (1.32). Whereas among those three statistics, only T_ϕ is of the form $nD_\phi(\boldsymbol{p}_0, \boldsymbol{u}_k)$, the formulas of T_ϕ, R_ϕ, and S_ϕ are quite similar, and we shall refer to all three of them as **spacings-type disparity statistics** in the sequel. In the rest of the paper we introduce somer new spacings-type disparity statistics and study the asymptotic properties of all of them together. Let us describe briefly how the paper is organized.

Section 2 compares the structure of the new spacings-type disparity statistic T_ϕ is compared with that of the spacings-based statistics known from the literature, and three more spacings-type disparity statistics are introduced.

Section 3 deals with the asymptotic equivalence of these six different disparity statistics.

Section 4 presents a general asymptotic theory of spacings-type disparity statistics.

Section 5 introduces spacings-type power divergence statistics and presents results about their consistency.

Section 6 presents theorems on the asymptotic normality of the spacings-type power divergence statistics under local and fixed alternatives.

Section 7 comments on comparable results in previous papers in the literature.

2. Spacings-based statistics

This section reviews various types of spacings-based goodness-of-fit statistics known from the literature. As before, $0 \le Y_1 \le \cdots \le Y_n \le 1$ are the ordered observations. Unless otherwise explicitly stated, we use also the dummy observations $Y_0 = 0$ and $Y_{n+1} = 1$.

Let us start with our spacings-type ϕ-disparity statistic $T_\phi^{(m)}$ introduced in (1.26). This statistic is not efficient if $m > 1$ because then it ignores the observations Y_{mj+r} for $1 \le j \le k-1$ and $1 \le r \le m-1$. Shifting the orders j/k of the quantiles in (1.22) by a quantity depending on r, we obtain the additional quantiles

$$a_j^{(r)} = F_n^{-1}\left(\frac{mj+r}{n}\right) = Y_{mj+r}, \quad 1 \le j \le k-1, \quad 1 \le r \le m-1 \tag{2.1}$$

and, instead of $p_{0j} = Y_{mj} - Y_{m(j-1)} = p_{0j}^{(0)}$, the shifted hypothetical probabilities $p_{0j}^{(r)} = Y_{mj+r} - Y_{m(j-1)+r}$, while still preserving the uniform shifted empirical probabilities $p_{nj}^{(r)} = 1/k = m/n$ on the cells $(a_{j-1}^{(r)}, a_j^{(r)}]$, $1 \le r \le m-1$. Replacing each term $\phi(\frac{n}{m}(Y_{mj} - Y_{m(j-1)}))$ in (1.26) by the average

$$\frac{1}{m}\sum_{r=0}^{m-1} \phi\left(\frac{n}{m}(Y_{mj+r} - Y_{m(j-1)+r})\right) \tag{2.2}$$

of all $\phi(\frac{n}{m}(p_{0j}^{(r)}))$ for $0 \leq r \leq m-1$, we get a more efficient version of $T_\phi^{(m)}$, namely

$$\hat{T}_\phi^{(m)} = \sum_{j=0}^{n-m-1} \phi\left(\frac{n}{m}(Y_{j+m} - Y_j)\right) + m\phi\left(\frac{n}{m}(1 - Y_{n-m})\right) \tag{2.3}$$

which for $m = 1$ reduces to (1.29), so that the notation of our paper is consistent.

A similar procedure can be carried out for $S_\phi^{(m)}$ of (1.33), which involves the observations $Y_{mj}, j = 1, \cdots, k$, but ignores the observations Y_{mj+r} for $0 \leq j \leq k-1$ and $1 \leq r \leq m-1$. Applying the averaging and substitution from the previous paragraph, with n replaced by $n+1$ in (1.33), and excluding the terms containing undefined expressions (that is, the terms $Y_{mk+r} - Y_{m(k-1)+r}, 1 \leq r \leq m-1$, where $mk + r > n+1$), we get a similar more efficient version

$$\hat{S}_\phi^{(m)} = \sum_{j=0}^{n-m+1} \phi\left(\frac{n+1}{m}(Y_{j+m} - Y_j)\right) \tag{2.4}$$

of Del Pino's statistic $S_\phi^{(m)}$ of (1.33). Notice that if $m = 1$, then $\hat{S}_\phi^{(m)}$ of (2.4) reduces to S_ϕ of (1.32) above, so that our notation is in this sense still consistent.

The statistics (2.4) are formally well defined for all $1 \leq m \leq n$, and not only for $m = (n+1)/k \geq 1$ corresponding to the integers $1 < k \leq n+1$. Cressie (1976, 1979), Hall (1986), and Ekström (1999) are among the authors dealing with the statistics (2.4) for fixed $m \geq 1$ and eventually also for m slowly tending to ∞ when $n \to \infty$.

If $m > 1$, and in particular if $m \to \infty$, then the statistics (2.4) assign more weight to central spacings than to those in the tails. To avoid this, Hall (1986) proposed to wrap the observations $Y_1, Y_2, ..., Y_n$ around the circle of unit circumference and to define the m-spacings $Y_{m+j} - Y_j$ for arbitrary $1 \leq m \leq n$ and j as the distance between observations Y_j and Y_{j+m} on this circle. This leads either to the extension of the ordered observations $Y_1, \ldots, Y_n$ by the formula

$$Y_{n+j} = 1 + Y_j \quad \text{for} \quad j = 1, 2, ..., n \tag{2.5}$$

where the previous dummy observation $Y_0 = 0$ is suppressed and the other dummy observation $Y_{n+1} = 1$ is redefined in accordance with (2.5) by $Y_{n+1} = 1 + Y_1$, leading to the m-spacing $Y_{j+m} - Y_j$ to be equal to $1 + Y_{m+j-n} - Y_j$ if $n+1-m \leq j \leq n$, or to the extension by the alternative formula

$$Y_{n+j} = 1 + Y_{j-1} \quad \text{for} \quad j = 0, 1, \cdots, n \tag{2.6}$$

where the dummy observations $Y_0 = 0$ and $Y_{n+1} = 1$ are placed on the circle as well, resulting in the m-spacing $Y_{j+m} - Y_j$ to be defined as $1 + Y_{m+j-n-1} - Y_j$ if $n+2-m \leq j \leq n$. These extensions of the ordered observations Y_j beyond $j > n$ allow to add in (2.4) the tail evidence missing there, namely by adding to the substituted averages (2.2) also the previously excluded terms. Depending on whether we use (2.5) or the alternative extension (2.6), we get in this manner two different extensions of (2.4), namely

$$\tilde{S}_\phi^{(m)} = \sum_{j=1}^{n} \phi\left((n+1)\,\frac{Y_{j+m}-Y_j}{m}\right) \quad \text{where} \quad Y_{j+m} = 1 + Y_{j+m-n} \tag{2.7}$$

if $j = n+1-m, \cdots, n$, or

$$\tilde{\tilde{S}}_\phi^{(m)} = \sum_{j=0}^{n} \phi\left((n+1)\,\frac{Y_{j+m}-Y_j}{m}\right) \quad \text{where} \quad Y_{j+m} = 1 + Y_{j+m-n-1} \tag{2.8}$$

if $j = n+2-m, \cdots, n$, and $Y_0 = 0$ (cf (2.6)). The statistics from the class (2.7) were studied for example by Hall (1986) and Morales *et al.* (2003), while those from the class (2.8) were investigated among others by Cressie (1978), Rao and Kuo (1984), Ekström (1999) and Misra and van der Meulen (2001).

As said in the Introduction, this paper deals only with the ordinary spacings where $m = 1$. This means that we use the statistics T_ϕ in the form presented in (1.29). If $m = 1$, then not only $S_\phi^{(m)}$ of (1.33) and $\hat{S}_\phi^{(m)}$ of (2.4) reduce for all ϕ to the statistic

$$S_\phi = \sum_{j=0}^{n} \phi\left((n+1)\left(Y_{j+1}-Y_j\right)\right), \quad \text{where } Y_{n+1} = 1 \text{ and } Y_0 = 0, \tag{2.9}$$

introduced in (1.32), but also $\tilde{\tilde{S}}_\phi^{(m)}$ of (2.8) does so for all ϕ. However, $\tilde{S}_\phi^{(m)}$ of (2.7) does not reduce to S_ϕ, unless ϕ is linear. Indeed, if $m = 1$, $\tilde{S}_\phi^{(m)}$ reduces to

$$\tilde{S}_\phi = \sum_{j=1}^{n-1} \phi\left((n+1)\left(Y_{j+1}-Y_j\right)\right) + \phi\left((n+1)\left(Y_1+1-Y_n\right)\right) \tag{2.10}$$

which coincides with

$$S_\phi = \sum_{j=1}^{n-1} \phi\left((n+1)\left(Y_{j+1}-Y_j\right)\right) + \phi\left((n+1)\,Y_1\right) + \phi\left((n+1)\left(1-Y_n\right)\right) \tag{2.11}$$

only if

$$\phi\left((n+1)\,Y_1\right) + \phi\left((n+1)\left(1-Y_n\right)\right) = \phi\left((n+1)\left(Y_1+1-Y_n\right)\right)$$

which takes place with a positive probability only for linear ϕ.

In what follows we use the functions

$$\phi^{(n)}(t) = \phi\left(\frac{n+1}{n}\,t\right) \tag{2.12}$$

and, in addition to T_ϕ, S_ϕ, $\tilde{S}_\phi$, also the statistics R_ϕ introduced earlier in (1.30). Moreover, we study another new type of spacings-type disparity statistic denoted by $\tilde{T}_\phi$. To obtain it, we redefine the former partition $\mathcal{P}^1 = \{(0, Y_1], \dots, (Y_{n-2}, Y_{n-1}], (Y_{n-1}, 1)\}$ of $(0,1)$ defined in (1.27), which led to the hypothetical distribution $\boldsymbol{p}_0$ of (1.28) and the empirical distribution $\boldsymbol{p}_n = \boldsymbol{u}_n$ on $\mathcal{P}^1$ which both served as arguments of our general ϕ-disparity statistic T_ϕ of (1.29) derived from (1.25). The new partition of $(0,1)$ is obtained by rearranging the n intervals of $\mathcal{P}^1$ into n new intervals by the rule

$$(0, Y_1] \mapsto (0, Y_1] \cup (Y_n, 1) \quad \text{and} \quad (Y_{n-1}, 1) \mapsto (Y_{n-1}, Y_n], \tag{2.13}$$

with the intervals $(Y_{j-1}, Y_j]$, $\quad 2 \le j \le n-1$, remaining the same. This new partition, denoted by $\tilde{\mathcal{P}}^1$, leads to the *modified* hypothetical distribution

$$\tilde{\boldsymbol{p}}_0 = (\tilde{p}_{01} = Y_1 + 1 - Y_n,\, \tilde{p}_{02} = Y_2 - Y_1, \dots, \tilde{p}_{0n} = Y_n - Y_{n-1})$$

but preserves the original uniform empirical distribution $\boldsymbol{p}_n = \boldsymbol{u}_n$ on the cells of $\tilde{\mathcal{P}}^1$, as each of the new n intervals still contains exactly one of the observations $Y_1, \dots, Y_n$. Therefore the new partition $\tilde{\mathcal{P}}^1$ leads to the new spacings-type disparity statistic

$$\begin{aligned} \tilde{T}_\phi &= n\, D_\phi\left(\tilde{\boldsymbol{p}}_0, \boldsymbol{u}_n\right) = \sum_{j=1}^{n} \phi(n\tilde{p}_{0j}) \quad \text{(cf (1.25))} \\ &= \sum_{j=2}^{n} \phi\left(n(Y_j - Y_{j-1})\right) + \phi\left(n(Y_1 + 1 - Y_n)\right), \end{aligned} \tag{2.14}$$

which differs from T_ϕ of (1.29). Applying (2.12), we obtain the useful relations

$$\tilde{S}_\phi = \tilde{T}_{\phi^{(n)}} \quad \text{and} \quad S_\phi = R_{\phi^{(n)}}. \tag{2.15}$$

In addition to the statistics $R_\phi, S_\phi, \tilde{S}_\phi$, T_ϕ, $\tilde{T}_\phi$, defined above in (1.30), (1.32) , (2.10), (1.29), and (2.14), respectively, we use in this paper also the auxiliary spacings-based statistics

$$\tilde{R}_\phi = \sum_{j=1}^{n-1} \phi\left(n(Y_{j+1} - Y_j)\right) = R_\phi - \phi(nY_1) - \phi\left(n(1 - Y_n)\right), \tag{2.16}$$

investigated previously by authors neglecting the tail probabilities such as for example Hall (1984).

3. Asymptotic equivalence

The spacings-type ϕ-disparity statistics $R_\phi, \tilde{R}_\phi, S_\phi, \tilde{S}_\phi, T_\phi$, and $\tilde{T}_\phi$ are with probability one formally well defined by (1.30), (2.16), (1.32), (2.10), (1.29), and (2.14) for all functions $\phi : (0, \infty) \mapsto \mathbb{R}$. However, our original functionals (1.3), (1.4) are justified as measures of disparity only for some of these functions. They are best justified for ϕ from the class $\boldsymbol{\Phi}$ defined in Section 1 in the paragraph after (1.5) when they represent ϕ-divergences. In what follows we relax the restrictions imposed on $\boldsymbol{\Phi}$ and we consider the *larger class* $\boldsymbol{\Phi}_0$

of all continuous functions $\phi : (0, \infty) \mapsto \mathbb{R}$ *which are twice continuously differentiable in a neighborhood of* 1 *with* $\phi''(1) > 0$ *and* $\phi(1) = 0$. We see that this larger class does not only contain the *convex* functions which define ϕ-divergences, but also those *continuous* functions which define ϕ-disparities in the sense made precise in Section 1 in the paragraph preceding (1.6).

In fact, in this section and in the rest of this paper we study the subset

$$\{R_\phi, \tilde{R}_\phi, S_\phi, \tilde{S}_\phi, T_\phi, \tilde{T}_\phi\} \tag{3.1}$$

of the mentioned spacings-type ϕ-disparity statistics for ϕ from one of the subsets $\mathbf{\Phi}_2 \subset \mathbf{\Phi}_1 \subset \mathbf{\Phi}_0$ defined by the condition that there exist functions ξ, $\eta, \zeta : (0, \infty) \mapsto \mathbb{R}$ such that every $\phi \in \mathbf{\Phi}_1$ satisfies for all s, $t \in (0, \infty)$ the functional equation

$$\phi(st) = \xi(s)\, \phi(t) + \zeta(t)\, \phi(s) + \eta(s)\, (t-1) \tag{3.2}$$

and every $\phi \in \mathbf{\Phi}_2$ satisfies the stronger functional equation

$$\phi(st) = \xi(s)\, \phi(t) + \phi(s) + \eta(s)\, (t-1). \tag{3.3}$$

Lemma 3.1. The functions ξ, ζ and η are continuous on $(0, \infty)$ and satisfy the relations

$$\xi(1) = \zeta(1) = 1 \quad \text{and} \quad \eta(1) = 0. \tag{3.4}$$

Proof. The continuity of ξ and η can be obtained by putting $t = 2$ and $t = 3$, and that of ζ by putting $s = 2$ in (3.2). If we put $s = 1$ in (3.2) or (3.3) and use the assumption $\phi(1) = 0$, then we obtain that for all $t \in (0, \infty)$

$$(\xi(1) - 1)\, \phi(t) + \eta(1)\, (t-1) = 0.$$

This contradicts the assumption $\phi''(1) > 0$, unless $\xi(1) = 1$ which implies also $\eta(1) = 0$. By putting $t = 1$ in (3.2) we find that $\zeta(1) = 1$. □

Lemma 3.2. Every $\phi \in \mathbf{\Phi}_1$ is differentiable on $(0, \infty)$, the corresponding functions ξ and η are differentiable at 1, and for every $t > 0$

$$\phi'(t) = \xi'(1)\, \frac{\phi(t)}{t} + \phi'(1)\, \frac{\zeta(t)}{t} + \eta'(1)\, \frac{t-1}{t}. \tag{3.5}$$

Proof. Putting $s = 1 + \varepsilon$ and

$$\xi^*(\varepsilon) = \frac{\xi(1+\varepsilon) - \xi(1)}{\varepsilon}, \quad \eta^*(\varepsilon) = \frac{\eta(1+\varepsilon) - \eta(1)}{\varepsilon}$$

we obtain from (3.2) for every $t > 0$ and ε close to 0

$$t\, \frac{\phi(t+\varepsilon t) - \phi(t)}{\varepsilon t} = \xi^*(\varepsilon)\, \phi(t) + \frac{\phi(1+\varepsilon) - \phi(1)}{\varepsilon}\, \zeta(t) + \eta^*(\varepsilon)\, (t-1). \tag{3.6}$$

Since ϕ is differentiable in a neighborhood of 1, we have for t close to 1

$$\xi^*(\varepsilon)\,\phi(t) + \eta^*(\varepsilon)\,(t-1) = t\,\phi'(t) - \phi'(1)\,\zeta(t) + o(\varepsilon) \quad \text{as } \varepsilon \to 0.$$

By assumptions concerning $\mathbf{\Phi}$, $\phi(t)$ is not linear in a neighborhood of $t = 1$. Therefore the last relation implies that the limits of $\xi^*(\varepsilon)$ and $\eta^*(\varepsilon)$ for $\varepsilon \to 0$ exist, that is,

$$\xi^*(\varepsilon) = \xi'(1) + o(\varepsilon) \quad \text{and} \quad \eta^*(\varepsilon) = \eta'(1) + o(\varepsilon) \quad \text{as } \varepsilon \to 0.$$

Now (3.5) follows from (3.6) for all $t > 0$. □

Example 3.1. The function $\phi(t) = (1-t)/t$, $t > 0$, belongs to $\mathbf{\Phi}$ and satisfies (3.3) for $\xi(t) = 1/t$ and $\eta(t) \equiv 0$. Therefore it belongs to $\mathbf{\Phi}_2 \subset \mathbf{\Phi}$. The function $\phi(t) = (1-t)^2/t$, $t > 0$, belongs to $\mathbf{\Phi}$ too and satisfies (3.3) for the same $\xi(t)$ as above and $\eta(t) = t - 1/t$. Therefore it belongs to $\mathbf{\Phi}_2$. The functions defined on $(0,\infty)$ by

$$\phi_\alpha(t) = \frac{t^\alpha \ln t}{(2\alpha - 1)}, \quad \alpha \in \mathbb{R} - \{\tfrac{1}{2}\}$$

belong to $\mathbf{\Phi}$ and satisfy (3.2) for $\xi(t) = \zeta(t) = t^\alpha$ and $\eta(t) \equiv 0$. Therefore

$$\{\phi_\alpha : \alpha \in \mathbb{R} - \{\tfrac{1}{2}\}\} \subset \mathbf{\Phi}_1$$

and $\phi_0 \in \mathbf{\Phi}_2$. But ϕ_1 satisfies also (3.3) for $\xi(t) = t$ and $\eta(t) = t \ln t$ and therefore ϕ_1 belongs to $\mathbf{\Phi}_2$.

In the remainder of this paper the observations are assumed to be distributed on $(0,1]$ in two possible ways:

(i) *under a fixed alternative*,

(ii) *under local alternatives.*

Case (i) means that the observations are distributed by a fixed distribution function $F \sim f$ with f positive and continuous on $[0,1]$. Case (ii) means that the observations from samples of sizes $n = 1, 2, \ldots$ are distributed by distribution functions

$$F^{(n)}(x) = F_0(x) + \frac{L_n(x)}{\sqrt[4]{n}} = x + \frac{L_n(x)}{\sqrt[4]{n}} \tag{3.7}$$

on $[0,1]$, where the functions $L_n : \mathbb{R} \mapsto \mathbb{R}$ are continuously differentiable, with $L_n(0) = L_n(1) = 0$, and with derivatives $\ell_n(x) = L_n'(x)$ tending on $[0,1]$ to a continuously differentiable function $\ell : \mathbb{R} \mapsto \mathbb{R}$ uniformly in the sense that

$$\sup_{0 \le x \le 1} |\ell_n(x) - \ell(x)| = o(1) \quad \text{as } n \to \infty. \tag{3.8}$$

The two possibilities (i) and (ii) are not mutually exclusive: their conjunction is "under the hypothesis $\mathcal{H}_0$ " where $F(x) = F_0(x)$, $f(x) = f_0(x) = \boldsymbol{I}_{[0,1]}(x)$ and $L_n(x) \equiv 0$ on $\mathbb{R}$ for all n. This means that the asymptotic results obtained under local alternatives for $\ell(x)$ of (3.8) being identically equal to 0 must coincide with the results obtained under the fixed alternative for $F(x) = F_0(x)$.

The theorems below demonstrate that if $\phi \in \mathbf{\Phi}_2$ defines a ϕ-divergence or ϕ-disparity, then the statistics S_ϕ, $\tilde{S}_\phi$, R_ϕ and $\tilde{R}_\phi$, which are formally not scaled ϕ-divergences or ϕ-disparities of the hypothetical and empirical distributions F_0 and F_n, share the most important statistical properties with the statistics T_ϕ and $\tilde{T}_\phi$, which are scaled ϕ-divergences or ϕ-disparities of this type . Therefore they provide a key argument for the thesis of the present paper formulated in Section 2, that the spacings-based goodness-of-fit statistics considered in the previous literature actually measure a disparity between the hypothetical and empirical distributions F_0 and F_n, although this was possibly not so intended by the various authors. But the main purpose of the following theorems is to present a systematic asymptotic theory for the whole set of statistics (3.1) and to demonstrate that the small modifications distinguishing these statistics from one another are asymptotically negligible. The restriction to the functions from $\mathbf{\Phi}_2$ or even $\mathbf{\Phi}_1$ is not essential – it only simplifies the proof of the next theorem.

Theorem 3.1. Consider the observations under fixed or local alternatives, and the set of statistics $\{R_\phi, \tilde{R}_\phi, S_\phi, \tilde{S}_\phi, T_\phi, \tilde{T}_\phi\}$ defined in (1.30), (2.16), (1.32), (2.10), (1.29), and (2.14). If $\phi \in \mathbf{\Phi}_1$ then for any statistic $U_\phi \in \{R_\phi, S_\phi, \tilde{S}_\phi, T_\phi\}$

$$U_\phi - \tilde{R}_\phi = O_p(1) \quad \text{as } n \to \infty \tag{3.9}$$

and if $\phi \in \mathbf{\Phi}_2$ then

$$S_\phi - R_\phi = \varepsilon_n R_\phi + \delta_n \quad \text{and} \quad \tilde{S}_\phi - \tilde{T}_\phi = \varepsilon_n \tilde{T}_\phi + \delta_n \tag{3.10}$$

where $\varepsilon_n = o(1)$ and $\delta_n = \phi'(1) + o(1)$ as $n \to \infty$.

Proof. We shall consider the fixed alternative $F(x)$ with a continuous density $f(x) > 0$ for $0 \le x \le 1$. For the local alternatives the argument is similar. By inspecting the definitions of T_ϕ, $\tilde{T}_\phi$ and R_ϕ we see that for (3.9) it suffices to prove that as $n \to \infty$

$$\phi(np_{01}) = O_p(1) \quad \text{and} \quad \phi(n(p_{01} + p_{02})) = O_p(1). \tag{3.11}$$

It is known (see for example page 208 in Hall (1986)) that $p_{01} = F^{-1}(Z_1/W_{n+1})$ and $p_{01} + p_{02} = F^{-1}((Z_1 + Z_2)/W_{n+1})$, where $Z_1, \dots, Z_{n+1}$ are independent standard exponential variables and $W_{n+1} = Z_1 + \cdots + Z_{n+1}$, so that, for $n \to \infty$,

$$\frac{W_{n+1}}{n} \xrightarrow{p} 1 \quad \text{and} \quad V_n = \frac{Z_1}{W_{n+1}} \xrightarrow{p} 0.$$

Setting

$$\Upsilon_n = \frac{F^{-1}(V_n)}{V_n} = \frac{F^{-1}(V_n) - F^{-1}(0)}{V_n}$$

and using the mean value theorem and the assumed continuity of f in the neighborhood of 0, we find that

$$\Upsilon_n \xrightarrow{p} \frac{1}{f(0)} \quad \text{as } n \to \infty$$

where, by assumptions about f, $0 < f(0) < \infty$. Thus

$$np_{01} = \frac{n}{W_{n+1}}\, Z_1\, \Upsilon_n$$

and, by applying (3.2),

$$\phi(np_{01}) = \xi\left(\frac{n}{W_{n+1}}\right)\, \phi(Z_1\, \Upsilon_n) + \zeta(Z_1\, \Upsilon_n)\, \phi\left(\frac{n}{W_{n+1}}\right) + \eta\left(\frac{n}{W_{n+1}}\right)\, (Z_1\, \Upsilon_n - 1).$$

Since $Z_1\, \Upsilon_n = O_p(1)$ as $n \to \infty$, we obtain from Lemma 3.1

$$\begin{aligned} \phi(np_{01}) &= \left[\xi\left(\frac{n}{W_{n+1}}\right) + \phi\left(\frac{n}{W_{n+1}}\right) + \eta\left(\frac{n}{W_{n+1}}\right)\right] O_p(1) \\ &= [\xi(1) + \phi(1) + \eta(1) + o_p(1)]\, O_p(1) \\ &= O_p(1) \quad \text{(cf (3.4))}, \end{aligned}$$

thus proving the first relation of (3.11). Replacing $V_n = Z_1/W_{n+1}$ by $V_n = (Z_1+Z_2)/W_{n+1}$, and using the fact that now

$$(Z_1 + Z_2)\, \Upsilon_n = (Z_1 + Z_2)\, \frac{F^{-1}(V_n) - F^{-1}(0)}{V_n} = O_p(1)$$

we obtain the second relation of (3.11). Next we prove (3.10). From (3.3) we get for any $p > 0$

$$\phi((n+1)\, p) = \xi\left(\frac{n+1}{n}\right)\, \phi(np) + \phi\left(\frac{n+1}{n}\right) + \eta\left(\frac{n+1}{n}\right)(np - 1)$$

so that

$$\phi((n+1)\, p) - \phi(np) = \varepsilon_n \phi(np) + \phi\left(\frac{n+1}{n}\right) + \eta\left(\frac{n+1}{n}\right)(np - 1) \tag{3.12}$$

where $\varepsilon_n = \xi((n+1)/n) - 1 = o(1)$ as $n \to \infty$ by Lemma 3.1. Replacing p by the probabilities $p_{0j} = Y_j - Y_{j-1}$ figuring in the definitions of S_ϕ and R_ϕ (cf (1.32) and (1.30)), and summing over $1 \leq j \leq n+1$, we get the equality

$$S_\phi - R_\phi = \varepsilon_n\, R_\phi + \delta_n$$

for

$$\begin{aligned}\delta_n &= (n+1)\,\phi\left(\frac{n+1}{n}\right) - \eta\left(\frac{n+1}{n}\right)\\ &= \frac{n+1}{n}\,\frac{\phi\left(1+\frac{1}{n}\right)-\phi(1)}{\frac{1}{n}} - \eta\left(\frac{n+1}{n}\right).\end{aligned}$$

By Lemma 3.1,

$$\delta_n = \phi'(1) + o(1) \quad \text{as } n \to \infty.$$

This completes the proof of the first relation in (3.10). The proof of the second relation is the same: we just replace p in (3.12) by the probabilities $\tilde{p}_{0j}$ figuring in the definition (2.14) of $\tilde{T}_\phi$.

□

4. General asymptotic theory

In this section we study the same spacings-type ϕ-disparity statistics $R_\phi, \tilde{R}_\phi, S_\phi, \tilde{S}_\phi, T_\phi$ and $\tilde{T}_\phi$, defined by (1.30), (2.16), (1.32), (2.10), (1.29), and (2.14), for ϕ from $\mathbf{\Phi}_2$ or $\mathbf{\Phi}_1$ as in the previous section. Unless otherwise explicitly stated, these statistics are assumed to be distributed under the fixed or local alternatives introduced as case (i) and case (ii) in Section 3.

For every continuous function $\psi : (0,\infty) \mapsto \mathbb{R}$ we define the condition

$$\lim_{t\to\infty} t^{-\alpha}|\psi(t)| = \lim_{t\downarrow 0} t^{\beta}|\psi(t)| = 0 \quad \text{for some } \alpha \geq 0 \text{ and } \beta < 1 \tag{4.1}$$

and the integral

$$\langle\psi\rangle = \langle\psi(t)\rangle = \int_0^\infty \psi(t)\, e^{-t}\,\mathrm{d}t. \tag{4.2}$$

Obviously, if (4.1) holds then $\langle\psi\rangle$ exists and is finite.

Let $\phi \in \mathbf{\Phi}_1$ satisfy (4.1) and let

$$\xi = \xi_\phi, \quad \zeta = \zeta_\phi \quad \text{and} \quad \eta = \eta_\phi \tag{4.3}$$

be the corresponding functions satisfying the functional equation (3.2). Then all functions

$$\psi(t) = \phi(ts) - \phi(t)\,\zeta(s), \quad s > 0,$$

satisfy (4.1) too, and by (3.2) the linear combinations

$$\psi(t) = \xi(t)\,\phi(s) + \eta(t)\,(s-1), \quad s > 0,$$

of functions $\xi(t)$ and $\eta(t)$ also satisfy (4.1). Since $\phi(s)$ is not linear in the neighborhood of $s = 1$, it follows from here that $\xi(t)$ and $\eta(t)$ themselves satisfy (4.1). Therefore the integrals $\langle\xi\rangle$ and $\langle\eta\rangle$ exist and are finite.

For the fixed alternatives $F \sim f$ we shall consider the linear combinations

$$\mu_\phi(f) = \langle\xi\rangle\, D_\phi(F_0, F) + \langle\phi\rangle\, D_\zeta(F_0, F) \tag{4.4}$$

where

$$D_\phi(F_0, F) = \int_0^1 f(x)\phi\left(\frac{1}{f(x)}\right) \mathrm{d}x \tag{4.5}$$

and

$$D_\zeta(F_0, F) = \int_0^1 f(x)\zeta\left(\frac{1}{f(x)}\right) \mathrm{d}x \tag{4.6}$$

are disparities of the distributions F_0 and F, well defined by (1.3) under the present assumptions about the densities f_0 and f, and are finite. If $\phi(t)$ is convex on $(0, \infty)$, or $\phi(t) - \phi'(1)\,(t-1)$ is monotone on $(0, 1)$ and $(1, \infty)$, then $D_\phi(F_0, F)$ is a nonnegative ϕ-divergence or ϕ-disparity of F_0 and F. Similarly, if $\zeta(t)$ is convex on $(0, \infty)$, or $\zeta(t) - \zeta(1) - \zeta'(1)\,(t-1)$ is monotone on $(0, 1)$ and $(1, \infty)$, then the ϕ^*-divergence or ϕ^*-disparity of F_0 and F for

$$\phi^*(t) = \zeta(t) - \zeta(1) = \zeta(t) - 1 \quad \text{(cf (3.4))}$$

satisfies the relation $D_{\phi^*}(F_0, F) = D_\zeta(F_0, F) - 1$. Hence the formula for $\mu_\phi(f)$ can be written for every $\phi \in \mathbf{\Phi}_1$ in the more intuitive form

$$\mu_\phi(f) = \langle\xi\rangle\, D_\phi(F_0, F) + \langle\phi\rangle\, [D_{\phi^*}(F_0, F) + 1] \tag{4.7}$$

where ξ and ϕ^* depend on ϕ as specified above, and $D_\phi(F_0, F)$, $D_{\phi^*}(F_0, F)$ are divergences or disparities between the hypothesis F_0 and the alternative F for typical $\phi \in \mathbf{\Phi}_1$. For $\phi \in \mathbf{\Phi}_2 \subset \mathbf{\Phi}_1$ it holds that $\zeta \equiv 1$, so that (4.7) then simplifies to

$$\mu_\phi(f) = \langle\xi\rangle\, D_\phi(F_0, F) + \langle\phi\rangle. \tag{4.8}$$

In particular for $\phi \in \mathbf{\Phi}_2$

$$\mu_\phi(f_0) = \langle\phi\rangle. \tag{4.9}$$

Theorem 4.1. Consider the observations under a fixed alternative $F \sim f$ with f positive and continuous on $[0, 1]$, and denote by U_ϕ any statistic from the class $\{R_\phi, \tilde{R}_\phi, T_\phi, \tilde{T}_\phi\}$. If $\phi \in \mathbf{\Phi}_1$ satisfies (4.1), then

$$\frac{U_\phi}{n} \xrightarrow{p} \mu_\phi(f) \quad \text{for } n \to \infty \tag{4.10}$$

where $\mu_\phi(f)$ is given by (4.7). If $\phi \in \mathbf{\Phi}_2$ satisfies (4.1), then the asymptotic relation (4.10) remains valid also for $U_\phi = \tilde{S}_\phi$ and $U_\phi = S_\phi$, and $\mu_\phi(f)$ is given by the simpler formula (4.8).

Proof. By Theorem 1 of Hall (1984), the statistic $\tilde{R}_\phi$ defined by (2.16) satisfies under a fixed alternative $F \sim f$ the relation

$$\frac{\tilde{R}_\phi}{n} \xrightarrow{p} \tilde{\mu}_\phi(f) = \int_0^1 f^2(x) \left(\int_0^\infty \phi(t)\, e^{-tf(x)} \mathrm{d}t \right) \mathrm{d}x \quad \text{as } n \to \infty$$

provided $\phi : (0, \infty) \mapsto \mathbb{R}$ is continuous and exponentially bounded in the sense that $|\phi(t)| \leq K(t^\alpha + t^{-\beta})$ for some $K > 0$, $\alpha \geq 0$, $\beta < 1$, and f is bounded, piecewise continuous and bounded away from 0 (see also part (i) of Theorem 3.1 in Misra and van der Meulen (2001)). Thus (4.10) is proved for $U_\phi = \tilde{R}_\phi$ as soon as it is shown that for $\phi \in \mathbf{\Phi}_1$ the limit $\tilde{\mu}_\phi(f)$ coincides with $\mu_\phi(f)$. By substituting s for $tf(x)$ in the last integral, and using the assumption $0 < f(x) < \infty$ and the functional equation (3.2),

$$\begin{aligned} \tilde{\mu}_\phi(f) &= \int_0^1 f(x) \left(\int_0^\infty \phi\left(\frac{s}{f(x)}\right) e^{-s} \mathrm{d}s \right) \mathrm{d}x \qquad (4.11) \\ &= \int_0^1 f(x) \left(\int_0^\infty \left[\xi(s)\, \phi\left(\frac{1}{f(x)}\right) + \zeta\left(\frac{1}{f(x)}\right) \phi(s) + \eta(s) \left(\frac{1}{f(x)} - 1\right) \right] e^{-s} \mathrm{d}s \right) \mathrm{d}x \\ &= \mu_\phi(f) + \int_0^\infty \eta(s)\, e^{-s}\, \mathrm{d}s \int_0^1 (1 - f(x))\, \mathrm{d}x = \mu_\phi(f). \end{aligned}$$

The extension of (4.10) to $U_\phi \in \{T_\phi, \tilde{T}_\phi, R_\phi\}$ follows from Theorem 3.1. For $\phi \in \mathbf{\Phi}_2$ the extension of (4.10) to $U_\phi \in \{S_\phi, \tilde{S}_\phi\}$ follows from Theorem 3.1 too. □

In the sequel we use the L_2-norm

$$\|\ell\| = \left(\int_0^1 \ell^2(x)\, \mathrm{d}x \right)^{1/2}$$

and we denote the integral (4.2) usually by $\langle \psi(t) \rangle$ rather than $\langle \psi \rangle$.

Theorem 4.2. Consider the observations under the local alternatives (3.7) with a limit function $\ell(x)$ introduced in (3.8), and denote by U_ϕ any statistic from the set $\{R_\phi, \tilde{R}_\phi, S_\phi, \tilde{S}_\phi, T_\phi, \tilde{T}_\phi\}$. If $\phi \in \mathbf{\Phi}_2$ satisfies the stronger version of (4.1) with $\beta < 1/2$ then

$$\frac{1}{\sqrt{n}}(U_\phi - n\mu_\phi) \xrightarrow{\mathcal{D}} N(m_\phi(\ell), \sigma_\phi^2) \quad \text{as } n \to \infty \qquad (4.12)$$

where

$$\mu_\phi = \langle \phi(t) \rangle, \quad \sigma_\phi^2 = \langle \phi^2(t) \rangle - \langle \phi(t) \rangle^2 - \left(\langle t\phi(t) \rangle - \langle \phi(t) \rangle \right)^2 \qquad (4.13)$$

and

$$m_\phi(\ell) = \frac{\|\ell\|^2}{2} \left(\langle t^2 \phi(t) \rangle - 4\langle t\phi(t) \rangle + 2\langle \phi(t) \rangle \right). \qquad (4.14)$$

Proof. For $U_\phi = S_\phi$ the relations (4.12) – (4.14) follow from the result of Kuo and Rao (1981), cf also Del Pino (1979) and Theorem 3.2 in Misra and van der Meulen (2001). The extension to the remaining statistics U_ϕ follows from Theorem 3.1. $\square$

Let us now consider the fixed alternative $F \sim f$ defined in Section 3 under (i), and $\phi \in \mathbf{\Phi}_2$ with $\xi = \xi_\phi$, $\eta = \eta_\phi$, satisfying the functional equation (3.3), and denote by ϕ', ξ', η' the derivatives of ϕ, ξ, η as in Lemma 3.2. To express the asymptotic normality under this alternative, we need auxiliary functions $\Psi_i = \Psi_{i,\phi}$ of the variable $x \in (0,1)$:

$$\begin{aligned}\Psi_1(x) &= \xi'(1)\,\langle\phi(t)\rangle\, f(x)\,\xi\left(\frac{1}{f(x)}\right) + \xi'(1)\, f(x)\,\phi\left(\frac{1}{f(x)}\right) \\ &\quad + [\phi'(1) - \eta'(1)]\, f(x) + \eta'(1) \end{aligned} \tag{4.15}$$

$$\begin{aligned}\Psi_2(x) &= \left(\langle\phi^2(t)\rangle - \langle\phi(t)\rangle^2\right) f(x)\,\xi^2\left(\frac{1}{f(x)}\right) + f(x)\,\eta^2\left(\frac{1}{f(x)}\right) \\ &\quad + 2(\langle t\phi(t)\rangle - \langle\phi(t)\rangle) f(x)\,\xi\left(\frac{1}{f(x)}\right)\eta\left(\frac{1}{f(x)}\right),\end{aligned} \tag{4.16}$$

$$\Psi_3(x) = (\langle t\phi(t)\rangle - \langle\phi(t)\rangle)\sqrt{f(x)}\,\xi\left(\frac{1}{f(x)}\right) + \sqrt{f(x)}\,\eta\left(\frac{1}{f(x)}\right), \tag{4.17}$$

and also

$$\Psi_4(x) = \frac{\sqrt{f(x)}}{F(x)}\int_0^x \left(1 - \frac{F(y)\, f'(y)}{f^2(y)}\right)\Psi_1(y)\,\mathrm{d}y \tag{4.18}$$

when the alternative density has a continuous derivative $f'(x)$ on $(0,1)$.

Theorem 4.3. Consider the observations under the fixed alternative $F \sim f$ with f positive and continuous on $[0,1]$ and continuously differentiable on $(0,1)$ with the derivative f' bounded. If U_ϕ is a statistic from the set $\{R_\phi, \tilde{R}_\phi, S_\phi, \tilde{S}_\phi, T_\phi, \tilde{T}_\phi\}$, and $\phi \in \mathbf{\Phi}_2$ satisfies the stronger version of (4.1) with $\beta < 1/2$, then

$$\frac{1}{\sqrt{n}}(U_\phi - n\mu_\phi(f)) \xrightarrow{\mathcal{D}} N(0, \sigma_\phi^2(f)) \quad \text{as} \quad n \to \infty \tag{4.19}$$

where $\mu_\phi(f)$ is given by (4.8) and

$$\sigma_\phi^2(f) = \int_0^1 \Psi_2(x)\,\mathrm{d}x - 2\int_0^1 \Psi_3(x)\,\Psi_4(x)\,\mathrm{d}x + \int_0^1 \Psi_4^2(x)\,\mathrm{d}x \tag{4.20}$$

for $\Psi_2(x)$, $\Psi_3(x)$ and $\Psi_4(x)$ defined by (4.16) – (4.18).

Proof. Consider $U_\phi = \tilde{R}_\phi$ for $\phi \in \mathbf{\Phi}_2$. By Lemma 3.2, $\phi(t)$ has a continuous derivative $\phi'(t)$ on $(0,\infty)$. By (3.5), for every $c \in \mathbb{R}$

$$t^c|\phi'(t)| \le |\xi'(1)|\, t^{c-1}|\phi(t)| + |\phi'(1)|\, t^c + |\eta'(1)|\, t^{c-1}|t-1|.$$

Thus if ϕ satisfies (4.1) with $\beta < 1/2$ then there exists $\alpha \geq 0$ such that

$$\lim_{t\to\infty} t^{-\alpha}|\phi'(t)| = \lim_{t\downarrow 0} t^{1+\beta}|\phi'(t)| = 0.$$

This means that under the assumptions of the theorem there exist $a > 0$, $K > 0$ and $b < 1/2$ such that for every $t \in (0, \infty)$

$$|\phi(t)| \leq K(t^a + t^{-b}) \quad \text{and} \quad |\phi'(t)| \leq K(t^a + t^{-b-1}).$$

For continuously differentiable functions ϕ satisfying these assumptions, and fixed alternatives with densities f continuously differentiable on $(0, 1)$, it follows from Theorem 2 in Hall (1984) (cf also part (ii) of Theorem 3.1 in Misra and van der Meulen (2001)) that $U_\phi = \tilde{R}_\phi$ satisfies the relation

$$\frac{1}{\sqrt{n}}(U_\phi - n\tilde{\mu}_\phi(f)) \xrightarrow{\mathcal{D}} N(0, \tilde{\sigma}^2_\phi(f)) \quad \text{for } n \to \infty$$

where: (1) the asymptotic mean $\tilde{\mu}_\phi(f)$ was presented and proved to be equal to $\mu_\phi(f)$ in the proof of Theorem 4.1 under assumptions weaker than here and, (2) the asymptotic variance $\tilde{\sigma}^2_\phi(f)$ can be specified by means of the standard exponential variable Z and the auxiliary function

$$G(x) = \int_0^x \left(1 - \frac{F(y)\, f'(y)}{f^2(y)}\right) E\left[Z\, \phi'\left(\frac{Z}{f(y)}\right)\right] \mathrm{d}y, \quad 0 < x < 1, \tag{4.21}$$

as the sum of

$$s_1^2(f) \;=\; \int_0^1 \left(E\phi^2\left(\frac{Z}{f(x)}\right) - \left[E\phi\left(\frac{Z}{f(x)}\right)\right]^2\right) f(x)\,\mathrm{d}x \tag{4.22}$$

$$s_2^2(f) \;=\; -2\int_0^1 E\left[(Z-1)\,\phi\left(\frac{Z}{f(x)}\right)\right] \frac{G(x)}{F(x)}\, f(x)\,\mathrm{d}x \tag{4.23}$$

and

$$s_3^2(f) = \int_0^1 \left(\frac{G(x)}{F(x)}\right)^2 f(x)\,\mathrm{d}x. \tag{4.24}$$

It remains to be proved that for every $x \in (0, 1)$

$$\left(E\,\phi^2\left(\frac{Z}{f(x)}\right) - \left[E\,\phi\left(\frac{Z}{f(x)}\right)\right]^2\right) f(x) = \Psi_2(x), \tag{4.25}$$

$$E\left[(Z-1)\,\phi\left(\frac{Z}{f(x)}\right)\right]\sqrt{f(x)} = \Psi_3(x) \tag{4.26}$$

and

$$\frac{G(x)\sqrt{f(x)}}{F(x)} = \Psi_4(x). \tag{4.27}$$

Indeed, then $\tilde{\sigma}_\phi^2(t) = \sigma_\phi^2(f)$ so that (4.19) is proved for $U_\phi = R_\phi$, and the extension of (4.19) to the remaining statistics $U_\phi \in \{\tilde{R}_\phi, S_\phi, \tilde{S}_\phi, T_\phi, \tilde{T}_\phi\}$ follows from Theorem 3.1. We shall prove (4.25) – (4.27) in the reversed order. By substituting $t = Z/f(y)$ in (3.5) and taking into account that $\zeta(t) \equiv 1$ we obtain

$$\begin{aligned} E\left[Z\phi'\left(\frac{Z}{f(y)}\right)\right] &= f(y)\,E\left[\xi'(1)\,\phi\left(\frac{Z}{f(y)}\right)+\phi'(1)+\eta'(1)\left(\frac{Z}{f(y)}-1\right)\right] \\ &= f(y)\left[\xi'(1)\,E\,\phi\left(\frac{Z}{f(y)}\right)+\phi'(1)+\eta'(1)\left(\frac{1}{f(y)}-1\right)\right] \end{aligned}$$

and, by putting $s = 1/f(x)$ and $t = Z$ in (3.3), we get

$$\phi\left(\frac{Z}{f(x)}\right) = \phi(Z)\,\xi\left(\frac{1}{f(x)}\right)+\phi\left(\frac{1}{f(x)}\right)+\eta\left(\frac{1}{f(x)}\right)(Z-1). \tag{4.28}$$

Therefore

$$E\,\phi\left(\frac{Z}{f(x)}\right) = \langle\phi\rangle\xi\left(\frac{1}{f(x)}\right)+\phi\left(\frac{1}{f(x)}\right) \tag{4.29}$$

and, consequently,

$$E\left[Z\,\phi'\left(\frac{Z}{f(y)}\right)\right] = \Psi_1(y). \tag{4.30}$$

This, together with the definitions of $\Psi_4(x)$ and $G(x)$ in (4.18) and (4.21), implies (4.27). Further, from (4.28) and the definition of $\Psi_3(x)$ in (4.17) we get (4.26). Finally, from (4.28), (4.29) and the definition of $\Psi_2(x)$ in (4.16) we obtain (4.25) which completes the proof.

□

Remark 4.1. Under the hypothesis $F_0 \sim f_0 \equiv 1$ both Theorems 4.2 and 4.3 deal with the same statistical model. Therefore the asymptotic parameters $(\mu_\phi, \sigma_\phi^2)$ from (4.13) and $(\mu_\phi(f_0), \sigma_\phi^2(f_0))$ from (4.8) and (4.20) must be the same, that is, the equalities

$$\mu_\phi(f_0) = \langle\phi\rangle \quad \text{and} \quad \sigma_\phi^2(f_0) = \langle\phi^2\rangle - \langle\phi\rangle^2 - (\langle t\phi(t)\rangle - \langle\phi\rangle)^2$$

must hold. The first equality is clear from (4.9). For $f = f_0$ we get from (4.30) by partial integration

$$\Psi_1(y) = \langle t\phi'(t)\rangle = \langle t\phi(t)\rangle - \langle\phi\rangle \quad \text{for all } y \in (0,1).$$

Thus, by (4.18), $\Psi_4(x)$ is under the hypothesis constant, equal to $\langle t\phi(t)\rangle - \langle\phi\rangle$. Similarly, by (4.16), (4.17) and Lemma 3.1, $\Psi_2(x) = \langle\phi^2\rangle - \langle\phi\rangle^2$ and $\Psi_3(x) = \Psi_4(x)$. Hence (4.20) implies the desired result

$$\sigma_\phi^2(f_0) = \Psi_2(x) - 2\Psi_4^2(x) + \Psi_4^2(x) = \sigma_\phi^2.$$

Remark 4.2. The expressions μ_ϕ, σ^2_ϕ are well defined by (4.13) for every continuous function $\phi : (0, \infty) \mapsto \mathbb{R}$ satisfying the condition (4.1) with $\beta < 1/2$. If this condition holds for some function $\psi : (0, \infty) \mapsto \mathbb{R}$, then it holds also for all linear transformations $\phi(t) = a\psi(t) + b(t-1) + c$ and

$$\mu_\phi = a\mu_\psi + c, \quad \sigma^2_\phi = a^2\sigma^2_\psi. \tag{4.31}$$

Let us now consider a fixed alternative $F \sim f$ with the density continuously differentiable on $(0,1)$. Then, using expression (4.11) for $\mu_\phi(f)$, and (4.22)–(4.24) for $s_i^2(f)$, the formulas

$$\mu_\phi(f) = \int_0^1 f(x) \left\langle \phi\left(\frac{t}{f(x)}\right)\right\rangle \mathrm{d}x \quad \text{and} \quad \sigma^2_\phi(f) = s_1^2(f) + s_2^2(f) + s_3^2(f) \tag{4.32}$$

define $\mu_\phi(f)$ and $\sigma^2_\phi(f)$ for all continuously differentiable functions $\phi : (0, \infty) \mapsto \mathbb{R}$ such that both $\phi(t)$ and $\tilde{\phi}(t) = t\phi'(t)$ satisfy (4.1) with $\beta < 1/2$. If ψ is one of the functions satisfying all these conditions then all linear transformations $\phi(t) = a\psi(t) + b(t-1) + c$ satisfy these conditions too and

$$\mu_\phi(f) = a\mu_\psi(f) + c, \quad \sigma^2_\phi(f) = a^2\sigma^2_\psi(f). \tag{4.33}$$

Formulas (4.31) and (4.33) are verifiable from the definitions mentioned in this remark and are useful for the evaluation of asymptotic means and variances.

Remark 4.3 We observe that the asymptotic results of Theorems 4.1, 4.2 and 4.3 are in each case for a fixed ϕ the same for *any* statistic U_ϕ from the class of statistics considered, thus demonstrating the asymptotic equivalence of these statistics announced and alluded to in Sections 1 and 3.

5. Power divergence statistics

The remaining part of this paper pays special attention to the subclass of spacings-based ϕ-disparity statistics studied in the previous section which are defined by the class of convex functions $\phi = \phi_\alpha : (0, \infty) \mapsto \mathbb{R}$ parametrized by $\alpha \in \mathbb{R}$ and defined by (1.6), (1.7). All these functions belong to the subset $\mathbf{\Phi}_2 \subset \mathbf{\Phi}$, that is, they satisfy the functional equation (3.3) with

$$\xi(t) = \xi_\alpha(t) = t^\alpha \quad \text{and} \quad \eta(t) = \eta_\alpha(t) = \begin{cases} \frac{t^\alpha - t}{\alpha - 1} & \text{if} \quad \alpha \neq 1 \\ \lim\limits_{\alpha \to 1} \frac{t^\alpha - t}{\alpha - 1} = t\ln t & \text{if} \quad \alpha = 1 \end{cases}. \tag{5.1}$$

In other words, if $\alpha \in \mathbb{R}$ then

$$\phi_\alpha(st) = s^\alpha\phi_\alpha(t) + \phi_\alpha(s) + (t-1) \cdot \begin{cases} \frac{s^\alpha - s}{\alpha - 1} & \text{if} \quad \alpha \neq 1 \\ s\ln s & \text{if} \quad \alpha = 1 \end{cases} \tag{5.2}$$

for all $s,\, t > 0$. We use the simplified notation

$$D_\alpha(\boldsymbol{p},\boldsymbol{q}) = D_{\phi_\alpha}(\boldsymbol{p},\boldsymbol{q}) \quad \text{and} \quad D_\alpha(F,G) = D_{\phi_\alpha}(F,G) \qquad \text{(cf (1.3)–(1.4))}$$

for the ϕ_α-divergences, and also the easily verifiable facts that

$$\tilde{\phi}_\alpha(t) := \phi_\alpha(t) + \frac{t-1}{\alpha-1} = \frac{t^\alpha - 1}{\alpha(\alpha-1)}, \quad \alpha \in \mathbb{R} - \{0,1\}$$

and

$$\tilde{\phi}_0(t) := \phi_0(t) - t + 1 = -\ln t, \quad \tilde{\phi}_1(t) := \phi_1(t) + t - 1 = t \ln t$$

are convex functions belonging to $\mathbf{\Phi}_2$ too, that the $\tilde{\phi}_\alpha$-divergences coincide with the ϕ_α-divergences, and that $\langle \tilde{\phi}_\alpha(t) \rangle = \langle \phi_\alpha(t) \rangle$. We also use freely the symbols for the concrete power divergences introduced in (1.8) – (1.11), therein replacing $\boldsymbol{p}$ and $\boldsymbol{q}$ by F and G.

In this and the following section we study the sets

$$\mathcal{U}_\alpha = \{R_{\phi_\alpha}, \tilde{R}_{\phi_\alpha}, S_{\phi_\alpha}, \tilde{S}_{\phi_\alpha}, T_{\phi_\alpha}, \tilde{T}_{\phi_\alpha}\} \tag{5.3}$$

of spacings-type ϕ_α-divergence statistics for $\alpha \in \mathbb{R}$. Similarly as Section 4, we restrict ourselves to the simple spacings (1-spacings), so that these statistics are well defined almost surely by (1.29), (1.30), (2.9), (2.10), (2.14) and (2.16) for functions ϕ replaced by ϕ_α. Similarly as the corresponding ϕ_α-divergences themselves, the ϕ_α-divergence statistics T_{ϕ_α}, $\tilde{T}_{\phi_\alpha}$ and S_{ϕ_α} are not altered if the nonnegative convex functions $\phi_\alpha \in \mathbf{\Phi}_2$ are replaced by the simpler convex functions $\tilde{\phi}_\alpha \in \mathbf{\Phi}_2$. Note that throughout this paper the spacings-type ϕ_α-divergence statistics T_{ϕ_α} are distinguished in notation from the frequency-type ϕ_α-divergence statistics $T_\alpha^{(f)}$ discussed in Section 1, Example 1.1, by having no superscript and using the subscript ϕ_α instead of just α.

The statistics T_{ϕ_α} and $\tilde{T}_{\phi_\alpha}$ are scaled ϕ_α-divergences of hypothetical and empirical distributions F_0 and F_n quantized by the empirical quantile partitions of the observation space $(0,1)$ discussed in part (**II**) Section 1 and in Section 2. For the other statistics from $\mathcal{U}_\alpha$ one cannot find partitions of $(0,1)$ enabling such a direct ϕ_α-divergence interpretation, but these statistics still reflect a proximity of F_0 and F reduced by some partitions of $(0,1)$, and depending on the functions ϕ_α or $\tilde{\phi}_\alpha$. Some of the statistics from $\mathcal{U}_\alpha$ are closely related to the spacings-based statistics studied in the previous literature, as it is mentioned in the following remark.

Remark 5.1. The statistic

$$\mathcal{G} = \sum_{j=1}^{n+1} (Y_j - Y_{j-1})^2 = \frac{1}{n+1}\left(1 + \frac{2S_{\phi_2}}{n+1}\right) = \frac{1}{n+1}\left(1 + \frac{2S_{\tilde{\phi}_2}}{n+1}\right) \tag{5.4}$$

with $Y_0 = 0$, $Y_{n+1} = 1$ was introduced by Greenwood (1946) the formula for $S_\psi = 2S_{\tilde{\phi}_2} + n + 1$ is presented based on $\psi(t) = t^2 = 2\tilde{\phi}_2(t) + 1)$ and studied later by Moran (1951) and many others. The statistic

$$\mathcal{M} = S_{\phi_0} = S_{\tilde{\phi}_0} \tag{5.5}$$

was introduced by Moran (1951) and studied later by Cressie (1976), van Es (1992), Ekström (1999) and many others cited by them. A class of statistics containing $\{\tilde{R}_{\phi_\alpha} : \alpha > -1/2\}$ was studied by Hall (1984), and classes containing $\{\tilde{S}_{\phi_\alpha} : \alpha \in \mathbb{R}\}$ and $\{R_{\phi_\alpha} : \alpha \in \mathbb{R}\}$ were studied by Hall (1986) and Jammalamadaka *et al.* (1986, 1989), respectively. Recently Misra and van der Meulen (2001) investigated the statistic $S_{\phi_1} = S_{\tilde{\phi}_1}$ (including its generalization to the m-spacings for fixed $m > 1$). The only papers dealing sofar with the spacings-based statistics with a direct ϕ_α-divergence interpretation appear to be those of Morales *et al.* (2003), Vajda and van der Meulen (2006), Vajda (2007) and Jimenez and Shao (2009). Morales *et al.* (2003) studied a class of statistics containing $\{\tilde{T}_{\phi_\alpha} : \alpha \in \mathbb{R}\}$, but the asymptotic theory was restricted there to the m-spacings with $m = m_n$ increasing to infinity for $n \to \infty$, similarly as in Hall (1986) or Jammalamadaka *et al.* (1986, 1989).

Since the general asymptotic theory of the statistics $U_\alpha \in \mathcal{U}_\alpha$ specified by (5.3) is covered by Theorem 3.1 and Theorems 4.1–4.3, the theorems that follow in the current and following sections are their corollaries. However, they bring explicit formulas and concrete results, the proofs of which are not trivial. These proofs are partly based on a continuity theory for the asymptotic parameters

$$\mu_\alpha(f) = \mu_{\phi_\alpha}(f),\ \sigma^2_\alpha(f) = \sigma^2_{\phi_\alpha}(f),\ \mu_\alpha = \mu_{\phi_\alpha},\ \sigma^2_\alpha = \sigma^2_{\phi_\alpha} \text{ and } m_\alpha(\ell) = m_{\phi_\alpha}(\ell), \tag{5.6}$$

defined by (4.32), (4.13) and (4.14), as functions of the structural parameter $\alpha \in \mathbb{R}$. Such a theory enables us to avoid a direct calculation of the asymptotic parameters at some $\alpha_0 \in \mathbb{R}$, if these calculations are tedious and the asymptotic parameters are known at the neighboring parameters α. This theory is summarized in Theorem 5.1 below using the next lemma. In Theorem 5.1 we take the representations (4.8) and (4.20) for $\mu_{\phi_\alpha}(f)$ and $\sigma^2_{\phi_\alpha}(f)$ rather than (4.32).

Lemma 5.1. Let $g(y)$ be a continuous positive function on a compact interval $[a, b] \subset \mathbb{R}$ and $\Phi(u, v)$ a continuous function of variables $u, v \in \mathbb{R}$. Furthermore let, for all α from an interval $(c, d) \subset \mathbb{R}$, $\psi_\alpha : (0, \infty) \mapsto \mathbb{R}$ be convex or concave functions differentiable at some point $t_* \in (0, \infty)$. If the values $\psi_\alpha(t)$, $t \in (0, \infty)$, and the derivatives $\psi'_\alpha(t_*)$ depend continuously on $\alpha \in (c, d)$, then for every $\alpha_0 \in (c, d)$

$$\lim_{\alpha \to \alpha_0} \int_a^b \Phi(g, \psi_\alpha(g))\, \mathrm{d}y = \int_a^b \Phi(g, \psi_{\alpha_0}(g))\, \mathrm{d}y. \tag{5.7}$$

Proof. By the assumptions about g,

$$t_0 = \min_{y \in [a,b]} g(y) > 0 \quad \text{and} \quad t_1 = \max_{y \in [a,b]} g(y) < \infty.$$

If $\psi_\alpha(t)$ is convex, then for every $t \in [t_0, t_1]$ and $\alpha \in (c, d)$

$$\psi'_\alpha(t_*)\,(t - t_*) \le \psi_\alpha(t) \le \psi_\alpha(t_0) + \psi_\alpha(t_1).$$

If $\psi_\alpha(t)$ is concave, then, similarly,

$$\psi_\alpha(t_0) + \psi_\alpha(t_1) \le \psi_\alpha(t) \le \psi'_\alpha(t_*)\,(t - t_*).$$

Therefore in both cases

$$\max_{t_0 \le t \le t_1} |\psi_\alpha(t)| \le \max\left\{|\psi_\alpha(t_0) + \psi_\alpha(t_1)|,\ |\psi'_\alpha(t_*)| \cdot |t_1 - t_0|\right\}.$$

The assumed continuity of $\psi'_\alpha(t_*)$ and $\psi_\alpha(t_0) + \psi_\alpha(t_1)$ in the variable $\alpha \in (c, d)$ implies that for all compact neighborhoods $N \subset (c, d)$ of α_0 the constant

$$k = \sup_{\alpha \in N} \max_{t_0 \le t \le t_1} |\psi_\alpha(t)| = \sup_{\alpha \in N} \max_{y \in [a,b]} |\psi_\alpha(g(y))|$$

is finite. Put

$$K = \max_{[t_0,t_1] \times [-k,k]} \Phi(u, v).$$

The function $|\Phi(g, \psi_\alpha(g))|$ of variables $(y, \alpha) \in [a, b] \times (c, d)$ is bounded on $[a, b] \times N$ by $K < \infty$. Since for every $y \in [a, b]$

$$\lim_{\alpha \to \alpha_0} \Phi(g, \psi_\alpha(g)) = \Phi(g, \psi_{\alpha_0}(g)),$$

the Lebesgue dominated convergence theorem for integrals implies (5.7). □

Theorem 5.1. The asymptotic parameters μ_α, σ^2_α and $m_\alpha(\ell)$ specified by (5.6) and (4.13), (4.14) are continuous in the variable $\alpha \in (-1/2, \infty)$. If the density f satisfies the assumptions of Theorem 4.1, then the asymptotic mean $\mu_\alpha(f)$ specified by (5.6) and (4.8) is continuous in the variable $\alpha \in (-1, \infty)$. If f satisfies the stronger assumptions of Theorem 4.3, then the asymptotic variance $\sigma^2_\alpha(f)$ specified by (5.6) and (4.20) is continuous in the variable $\alpha \in (-1/2, \infty)$.

Proof. Since $\mu_\alpha = \mu_\alpha(f_0)$ and $\sigma^2_\alpha = \sigma^2_\alpha(f_0)$, where the hypothetical density f_0 satisfies the assumptions of Theorems 4.1 and 4.3, the continuity of μ_α and σ^2_α follows from the continuity of $\mu_\alpha(f)$ and $\sigma^2_\alpha(f)$ proved below. By (5.6) and (4.14),

$$m_\alpha(\ell) = \frac{\|\ell\|^2}{2}\left(\langle t^2\phi_\alpha(t)\rangle - 4\langle t\phi_\alpha(t)\rangle + 2\langle \phi_\alpha(t)\rangle\right)$$

where ϕ_α is given by (1.6), (1.7), and, by (4.2),

$$\langle t^j \phi_\alpha(t)\rangle = \int_0^\infty t^j \phi_\alpha(t)\,\mathrm{d}H(t), \quad j \in \{0, 1, 2\} \tag{5.8}$$

for $H(t) = 1 - e^{-t}$. All integrals (5.8) are finite if and only if $\alpha \in (-1, \infty)$. Further, for every fixed $t > 0$

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\,\alpha\phi_\alpha(t) \ge 0 \quad \text{at any } \alpha \in \mathbb{R}. \tag{5.9}$$

Hence the continuity of the products $\alpha\langle t^j \phi_\alpha(t)\rangle$ in the variable $\alpha \in \mathbb{R}$ follows from the monotone convergence theorem for integrals, and this implies also the desired continuity of the integrals (5.8) at any $\alpha \in (-1,\infty) - \{0\}$. Further, for every fixed $t > 0$

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\,(\alpha-1)\,\phi_\alpha(t) \geq 0 \quad \text{for any } \alpha \in \mathbb{R}. \tag{5.10}$$

Hence the continuity of the products $(\alpha-1)\,\langle t^j \phi_\alpha(t)\rangle$ in the variable $\alpha \in \mathbb{R}$ follows as well from the monotone convergence theorem for integrals. Similarly as above, this implies the continuity of the integrals (5.8) at the remaining point $\alpha = 0$. Further, by (5.6) and (4.8),

$$\mu_\alpha(f) = \langle \xi_\alpha \rangle\, D_\alpha(F_0, F) + \langle \phi_\alpha \rangle$$

where, by (4.2) and (5.1)

$$\langle \xi_\alpha \rangle = \int_0^\infty t^\alpha \mathrm{d}H(t) \quad \text{and} \quad \langle \phi_\alpha \rangle = \int_0^\infty \phi_\alpha(t)\, \mathrm{d}H(t).$$

These integrals are finite if and only if $\alpha \in (-1,\infty)$. The continuity of $\langle \phi_\alpha \rangle$ at $\alpha \in (-1,\infty)$ was proved above, the continuity of $D_\alpha(F_0, F)$ at $\alpha \in \mathbb{R}$ follows from the assumptions about the densities f_0 and f and from Proposition 2.14 in Liese and Vajda (1987). The continuity of $\langle \xi_\alpha \rangle$ at $\alpha \in (-1,\infty)$ follows from the monotone convergence theorem for integrals applied separately to the integration domains $(0,1)$ and $(1,\infty)$. Finally, let us consider $\sigma_\alpha^2(f)$ defined by (4.15)–(4.20) for $\phi = \phi_\alpha$, $\xi = \xi_\alpha$, and $\eta = \eta_\alpha$ given by (1.6), (1.7) and (5.1). The integrals $\langle t\phi_\alpha(t)\rangle$, $\langle \phi_\alpha(t)\rangle$ and $\langle \phi_\alpha^2(t)\rangle$ are finite if and only if $\alpha \in (-1/2,\infty)$, and their continuity at $\alpha \in (-1/2,\infty)$ was either proved above or can be proved similarly as above. The continuity of the integral

$$\int_0^1 \left[f\xi_\alpha^2\left(\frac{1}{f}\right) + f\eta_\alpha^2\left(\frac{1}{f}\right)\right] \mathrm{d}x$$

at $\alpha \in (-1/2,\infty)$ follows from Lemma 4.1, which establishes the continuity of the component $\int \Psi_2(x)\,\mathrm{d}x$ of $\sigma_\alpha^2(f)$ in (4.20). For the continuity of the remaining two components, we take into account that $F(x) > c_1 x$ for some $c_1 > 0$ on $[0,1]$, because f is bounded away from zero on $[0,1]$. Furthermore, both $f(x)$ and $f'(x)$ are bounded on $[0,1]$, so that there exists a constant c_2 such that in (4.18)

$$\frac{\sqrt{f(x)}}{F(x)} \int_0^x \left| 1 - \frac{F(y)\, f'(y)}{f^2(y)} \right| \mathrm{d}y < c_2 \quad \text{for all } x \in [0,1]. \tag{5.11}$$

Using the function $\varphi_\alpha(t) = \alpha\phi_\alpha(t)$, which is for every $t > 0$ continuous and monotone in $\alpha \in \mathbb{R}$ (cf (5.9)), we obtain from (4.15)

$$\Psi_1(x) = \alpha\langle \phi_\alpha \rangle\, f(x)^{1-\alpha} + f(x)\, \varphi_\alpha\left(\frac{1}{f(x)}\right) + 1 - f(x)$$

where the right-hand side is bounded on $[0,1]$, locally uniformly in α, and continuous at any $\alpha \in \mathbb{R}$. By (4.18) and (5.11), this implies that also $\Psi_4(x)$ is bounded on $[0,1]$, locally uniformly in α, and continuous at any $\alpha \in \mathbb{R}$. Since the integrands in

$$\int_0^1 \left[\sqrt{f}\xi_\alpha\left(\frac{1}{f}\right) + \sqrt{f}\eta_\alpha\left(\frac{1}{f}\right)\right]\Psi_4 \,\mathrm{d}x \quad \text{and} \quad \int_0^1 \Psi_4^2 \,\mathrm{d}x$$

are continuous on $[0,1]$ and locally bounded in the variable $\alpha \in \mathbb{R}$, the continuity of both these integrals in the variable $\alpha \in \mathbb{R}$ follows from the Lebesgue dominated convergence theorem for integrals. This clarifies the continuity of the second and third component of $\sigma_\alpha^2(f)$ in (4.20) and thus completes the proof. □

5.1. Consistency of power divergence statistics

In the theorems below we use the gamma function of the variable $\alpha \in \mathbb{R}$ and the Euler constant,

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}\, e^{-t}\,\mathrm{d}t \quad \text{and} \quad \gamma = 0.577\ldots. \tag{5.12}$$

Theorem 5.2. Consider the observations under the fixed alternative $F \sim f$ assumed in Theorem 4.1 and denote by U_α any statistic from the class $\mathcal{U}_\alpha$ of (5.3). If $\alpha > -1$, then

$$\frac{U_\alpha}{n} \xrightarrow{p} \mu_\alpha(f) \quad \text{as } n \to \infty \tag{5.13}$$

for

$$\mu_\alpha(f) = D_\alpha(F_0, F)\,\Gamma(\alpha+1) + \mu_\alpha, \tag{5.14}$$

where

$$\mu_0 = \gamma, \quad \mu_1 = 1-\gamma, \quad \text{and} \quad \mu_\alpha = \frac{\Gamma(\alpha+1) - \Gamma(1)}{\alpha(\alpha-1)} \quad \text{for } \alpha \notin \{0,1\} \tag{5.15}$$

and $D_\alpha(F_0, F)$ are the ϕ_α-divergences

$$D_0(F_0,F) = \int_0^1 f \ln \frac{f}{f_0}\,\mathrm{d}x = \int_0^1 f(x)\ln f(x)\,\mathrm{d}x, \tag{5.16}$$

$$D_1(F_0,F) = \int_0^1 f_0 \ln \frac{f_0}{f}\,\mathrm{d}x = -\int_0^1 \ln f(x)\,\mathrm{d}x, \tag{5.17}$$

$$\begin{aligned} D_\alpha(F_0,F) &= \frac{1}{\alpha(\alpha-1)}\left(\int_0^1 f\left(\frac{f_0}{f}\right)^\alpha \mathrm{d}x - 1\right) \\ &= \frac{1}{\alpha(\alpha-1)}\left(\int_0^1 f(x)^{1-\alpha}\mathrm{d}x - 1\right) \qquad \text{for } \alpha \notin \{0,1\}. \end{aligned} \tag{5.18}$$

The ϕ_α-divergences are zero if and only if $F = F_0$, so that under the hypothesis $F = F_0$

$$\mu_\alpha(f_0) = \mu_\alpha, \quad \alpha \in \mathbb{R}. \tag{5.19}$$

Both parameters μ_α and $\mu_\alpha(f)$ are continuous in the variable $\alpha \in (-1,\infty)$ and satisfy the inequality $\mu_\alpha(f) \geq \mu_\alpha$, which is strict unless $F = F_0$.

Proof. The functions from the class $\{\phi_\alpha : \alpha \in (-1,\infty)\} \subset \mathbf{\Phi}_2$ satisfy all assumptions of Theorem 4.1. Hence (5.13) holds for all $\alpha > -1$ and the limit $\mu_\alpha(f)$ is given in accordance with (4.8) and (5.1) by the formula

$$\mu_\alpha(f) = \langle \xi_\alpha(t)\rangle\, D_\alpha(F_0,F) + \langle \phi_\alpha(t)\rangle = \langle t^\alpha\rangle\, D_\alpha(F_0,F) + \langle \tilde{\phi}_\alpha(t)\rangle$$

where $\langle t^\alpha\rangle = \Gamma(\alpha+1)$ for all $\alpha \in \mathbb{R}$. If $\alpha \notin \{0,1\}$ then

$$\langle \tilde{\phi}_\alpha(t)\rangle = \frac{1}{\alpha(\alpha-1)}\langle t^\alpha - 1\rangle = \frac{\Gamma(\alpha+1)-\Gamma(1)}{\alpha(\alpha-1)}$$

but the expressions

$$\langle \tilde{\phi}_0(t)\rangle = \langle -\ln t\rangle \quad \text{and} \quad \langle \tilde{\phi}_1(t)\rangle = \langle t\ln t\rangle$$

lead to the evaluation of unpleasant integrals. This evaluation can be avoided by employing Theorem 5.1. From the continuity of $\mu_\alpha = \langle \tilde{\phi}_\alpha(t)\rangle$, it follows that

$$\mu_j = \langle \tilde{\phi}_j(t)\rangle = \lim_{\alpha\to j} \frac{\Gamma(\alpha+1)-\Gamma(1)}{\alpha(\alpha-1)} \quad \text{for } j \in \{0,1\},$$

where the limit on the right can be easily evaluated by using L'Hospital's rule and the known formulas $\Gamma'(1) = -\gamma$, $\Gamma'(2) = 1-\gamma$, thus leading to the values μ_j, $j \in \{0,1\}$, given in (5.15). The continuity and the inequality $\mu_\alpha(f) \geq \mu_\alpha$ for $\alpha \in (-1,\infty)$ follow from (5.14) and (5.15) because $D_\alpha(F_0,F)$ is nonnegative and continuous in $\alpha \in \mathbb{R}$ and $\Gamma(\alpha+1)$ is positive and continuous in $\alpha \in (-1,\infty)$. The condition for equality follows from the fact that $D_\alpha(F_0,F)$ is positive unless $F = F_0$. □

Since $\Gamma(\alpha+1) = \alpha(\alpha-1)\,\Gamma(\alpha-1)$, (5.15) and (5.14) can be replaced for $\alpha \notin \{0,1\}$ by

$$\mu_\alpha = \Gamma(\alpha-1) - \frac{1}{\alpha(\alpha-1)} \quad \text{and} \quad \mu_\alpha(f) = \Gamma(\alpha-1)\int_0^1 f^{1-\alpha}\,\mathrm{d}x - \frac{1}{\alpha(\alpha-1)}. \tag{5.20}$$

Theorem 5.2 can be illustrated by Table 5.1, in which actual values of the parameters μ_α and $\mu_\alpha(f)$ are presented for selected parameters α. In this table, f denotes any density considered in Theorem 5.2, and the expressions for $D_\alpha(F_0,F)$, $I(F_0,F)$, $H(F_0,F)$, and $\chi^2(F_0,F)$ can be easily discerned from (1.8)–(1.11) replacing $\boldsymbol{p}$ and $\boldsymbol{q}$ by F_0 and F and sums by integrals.

Table 5.1: Values of μ_α and $\mu_\alpha(f)$ for selected $\alpha > -1$.

α	μ_α	$\mu_\alpha(f)$
$-\frac{1}{2}$	$\frac{4}{3}(\sqrt{\pi}-1) \doteq 1.030$	$\sqrt{\pi}\, D_{-1/2}(F_0, F) + \mu_{-1/2} = \frac{4\sqrt{\pi}}{3} \int_0^1 f^{3/2}\,\mathrm{d}x - \frac{4}{3}$
0	$\gamma \doteq 0.577$	$I(F, F_0) + \mu_0 = \int_0^1 f \ln f\,\mathrm{d}x + \gamma$
$\frac{1}{2}$	$4 - 2\sqrt{\pi} \doteq 0.455$	$2\sqrt{\pi}\, H(F_0, F) + \mu_{1/2} = 4 - 2\sqrt{\pi} \int_0^1 \sqrt{f}\mathrm{d}x$
1	$1 - \gamma \doteq 0.423$	$I(F_0, F) + \mu_1 = 1 - \gamma - \int_0^1 \ln f\,\mathrm{d}x$
2	$\frac{1}{2} = 0.500$	$\chi^2(F_0, F) + \mu_2 = \int_0^1 \frac{\mathrm{d}x}{f} - \frac{1}{2}$
3	$\frac{5}{6} \doteq 0.833$	$6D_3(F_0, F) + \mu_3 = \int_0^1 \frac{\mathrm{d}x}{f^2} - \frac{1}{6}$

6. Asymptotic laws for power divergence statistics

In this section we continue with the spacings-type power divergence statistics from the sets $\mathcal{U}_\alpha = \{R_{\phi_\alpha}, \tilde{R}_{\phi_\alpha}, S_{\phi_\alpha}, \tilde{S}_{\phi_\alpha}, T_{\phi_\alpha}, \tilde{T}_{\phi_\alpha}\}$ introduced in Section 5 (cf (5.3)) for $\alpha > -1/2$. We study the asymptotic distributions of these statistics both under the local alternatives assumed in Theorem 4.2 and under the fixed alternatives assumed in Theorem 4.3.

6.1. Asymptotic laws under local alternatives

Theorem 6.1. Consider the observations under the local alternatives (3.7) with the limit function $\ell(x)$ introduced in (3.8), and denote by U_α any statistic from the class $\mathcal{U}_\alpha$ of (5.3). If $\alpha > -1/2$, then

$$\frac{1}{\sqrt{n}}(U_\alpha - n\mu_\alpha) \xrightarrow{\mathcal{D}} N(m_\alpha(\ell), \sigma_\alpha^2) \qquad \text{as } n \to \infty \tag{6.1}$$

where the parameters μ_α, $m_\alpha(\ell)$, and σ_α^2 are continuous in the variable $\alpha \in (-1/2, \infty)$, and are given by (5.15) and the formulas

$$m_\alpha(\ell) = \frac{\|\ell\|^2}{2}\,\Gamma(\alpha+1) \tag{6.2}$$

$$\sigma_\alpha^2 = \frac{\Gamma(2\alpha+1) - (\alpha^2+1)\,\Gamma^2(\alpha+1)}{\alpha^2(\alpha-1)^2} \qquad \text{for } \alpha \notin \{0,1\} \tag{6.3}$$

and

$$\sigma_0^2 = \frac{\pi^2}{6} - 1, \qquad \sigma_1^2 = \frac{\pi^3}{3} - 3. \tag{6.4}$$

Proof. Similarly as we applied Theorem 4.1 in the proof of Theorem 5.2, (6.1) follows for all $\alpha > -1/2$ from Theorem 4.2. If $\alpha \notin \{0,1\}$, then the expressions for $m_\alpha(\ell)$ and σ_α^2 given in (6.2) and (6.3) follow easily from the formulas given for $m_{\phi_\alpha}(\ell)$ and $\sigma_{\phi_\alpha}^2$ in Theorem 4.2, but the direct evaluation of $m_j(\ell)$ and σ_j^2 from these formulas for $j \in \{0,1\}$ is a somewhat tedious task. However, by using the continuity of $m_\alpha(\ell)$ and σ_α^2 established in Theorem 5.1, we obtain $m_j(\ell)$ and σ_j^2 given in (6.2) and (6.4) as the limits

$$m_j(\ell) = \lim_{\alpha \to j} m_\alpha(\ell) \quad \text{and} \quad \sigma_j^2 = \lim_{\alpha \to j} \sigma_\alpha^2 \quad \text{for } j \in \{0,1\},$$

which expressions can be easily evaluated by using the continuity of the right-hand side of (6.2) and L'Hospital's rule, thereby employing the formulas

$$\begin{aligned} \Gamma(\alpha+k+1) &= (\alpha+k)(\alpha+k-1)\cdots(\alpha+1)\,\Gamma(\alpha+1), \\ \Gamma''(\alpha+1) &= 2\Gamma'(\alpha) + \alpha\Gamma''(\alpha) \end{aligned}$$

and

$$\Gamma''(1) = \frac{\pi^2}{6} + \gamma^2, \;\; \Gamma''(2) = \frac{\pi^2}{6} - 2\gamma + \gamma^2, \;\; \Gamma''(3) = \frac{\pi^2}{3} + 2 - 6\gamma + 2\gamma^2$$

in addition to the previously used $\Gamma'(1) = -\gamma$ and $\Gamma'(2) = 1 - \gamma$. $\square$

Theorem 6.1 provides the possibility to compute and compare asymptotic relative efficiencies of tests of the hypothesis $\mathcal{H}_0 : F_0 \sim f_0$ based on the statistics $U_\alpha \in \mathcal{U}_\alpha$, $\alpha > -1/2$, for various values of α. The Pitman asymptotic relative efficiency (ARE) of one test relative to another is defined as the limit of the inverse ratio of sample sizes required to obtain the same limiting power at the sequence of alternatives converging to the null hypothesis. If we define the "efficacies" of the statistics $U_\alpha \in \mathcal{U}_\alpha$ of Theorem 6.1 by

$$\text{eff}(U_\alpha) = \frac{\Gamma^2(\alpha+1)}{\sigma_\alpha^2} = \frac{(m_\alpha(\ell))^2}{\sigma_\alpha^2}\left(\frac{2}{\|\ell\|^2}\right)^2 \quad \text{for } \|\ell\|^2 \neq 0$$

then under the assumptions of Theorem 6.1 we get in accordance with Section 4 in Del Pino (1979)

$$\text{ARE}(U_{\alpha_1}, U_{\alpha_2}) = \frac{\text{eff}(U_{\alpha_1})}{\text{eff}(U_{\alpha_2})}$$

where U_{α_1} and U_{α_2} are arbitrary statistics from $\mathcal{U}_{\alpha_1}$ and $\mathcal{U}_{\alpha_2}$. Notice that arbitrary statistics U_α from the set $\mathcal{U}_\alpha$, α fixed, all have the same efficacy (cf also Remark 4.3). In Table 6.1 we present the parameters $m_\alpha(\ell)$, σ_α^2 and $\Gamma^2(\alpha+1)/\sigma_\alpha^2$ for selected values of $\alpha > -1/2$. This table indicates that the statistics $U_2 \in \{R_{\phi_2}, \tilde{R}_{\phi_2}, S_{\phi_2}, \tilde{S}_{\phi_2}, T_{\phi_2}, \tilde{T}_{\phi_2}\}$ are most asymptotically efficient in the Pitman sense among all statistics U_α, $\alpha > -1/2$. This extends the result on p. 1457 in Rao and Kuo (1984) about the asymptotic efficiency of the Greenwood statistic $\mathcal{G} = (2S_{\phi_2} + n + 1)/(n+1)^2$ (cf. (5.4) in Remark 5.1 and formula (7.2) below).

Table 6.1: The asymptotic parameters $m_\alpha(\ell)$, σ^2_α and eff(U_α) for selected statistics U_α of Theorem 6.1.

α	$m_\alpha(\ell)$	σ^2_α	eff(U_α)
0	$\frac{\|\ell\|^2}{2}$	$\frac{\pi^2}{6} - 1 \doteq 0.645$	1.550
$\frac{1}{2}$	$\|\ell\|^2 \frac{\sqrt{\pi}}{4} \doteq \frac{\|\ell\|^2}{2} \times 0.886$	$16 - 5\pi \doteq 0.292$	2.690
1	$\frac{\|\ell\|^2}{2}$	$\frac{\pi^2}{3} - 3 \doteq 0.290$	3.448
2	$\|\ell\|^2 = \frac{\|\ell\|^2}{2} \times 2$	1	4.000
3	$\|\ell\|^2\, 3 = \frac{\|\ell\|^2}{2} \times 6$	10	3.600

The general form of the asymptotic normality (6.1), as well as the continuity of the parameters μ_α, $m_\alpha(\ell)$ and σ^2_α in $\alpha \in (-1/2, \infty)$ established in Theorem 6.1 appear to be new results. The special result for $\alpha = 0$ also seems to be new. The particular result for $\alpha \in (-1/2, \infty) - \{0, 1\}$ and $U_\alpha = S_{\phi_\alpha}$ follows from the asymptotic normality obtained for the statistics

$$\sum_{j=1}^{n+1} ((n+1)\,(Y_j - Y_{j-1}))^\alpha = \alpha(\alpha - 1)\, S_{\phi_\alpha} + n + 1 \tag{6.5}$$

(cf (7.3) below) by Del Pino, see p. 1062 in Del Pino (1979). The particular result for $\alpha = 1$ and the statistics $U_1 = S_{\phi_1}$ with μ_1 and σ^2_1 given in the Tables 5.1 and 6.1 was obtained by Misra and van der Meulen (2001), who however considered m-spacings for arbitrary $m \geq 1$. They compared also the efficiency of the test statistics for $\alpha = 0$, $\alpha = 1$, and $\alpha = 2$ with a similar conclusion as in the Table 6.1.

6.2. Asymptotic laws under fixed alternatives

In the remaining part of this section we study the asymptotic distributions of the spacings-type power divergence statistics U_α from the sets $\mathcal{U}_\alpha = \{R_{\phi_\alpha}, \tilde{R}_{\phi_\alpha}, S_{\phi_\alpha}, \tilde{S}_{\phi_\alpha}, T_{\phi_\alpha}, \tilde{T}_{\phi_\alpha}\}$ for $\alpha > -1/2$ under the assumption that the observations are distributed by a fixed alternative $F \sim f$ satisfying the assumptions of Theorem 4.3. If $\alpha > -1/2$ then ϕ_α satisfies the assumption of Theorem 4.3 too. Therefore this theorem implies that

$$\frac{1}{\sqrt{n}}(U_\alpha - n\mu_\alpha(f)) \xrightarrow{\mathcal{D}} N(0, \sigma^2_\alpha(f)) \quad \text{for } n \to \infty \tag{6.6}$$

where the asymptotic parameters $\mu_\alpha(f)$, $\sigma^2_\alpha(f)$ are given by (5.6). Similarly as in the previous section, we are interested in explicit formulas for these parameters. By Theorem 4.3, the asymptotic mean is for all $\alpha \in \mathbb{R}$ given by the explicit formula (5.14) presented in Theorem 5.2. The only problem which remains is the formula for $\sigma_\alpha^2(f)$, $\alpha \in \mathbb{R}$.

The functions $\psi_\alpha(t) = t^\alpha$ with $\alpha > -1/2$ satisfy all assumptions of Remark 4.2 so that we can consider the quantities

$$\tau_\alpha^2(f) \equiv \sigma_{\psi_\alpha}^2(f), \quad \alpha \in (-1/2, \infty)$$

defined there. By (4.33),

$$\sigma_\alpha^2(f) = \frac{\tau_\alpha^2(f)}{\alpha^2(\alpha-1)^2} \quad \text{for } \alpha \in (-1/2, \infty) - \{0, 1\}. \tag{6.7}$$

One can find on p. 521 of Hall (1984) an expression for $\tau_\alpha^2(f)$ for all $\alpha \in (-1/2, \infty) - \{0, 1\}$, which for the case $m = 1$ can be given the form

$$\tau_\alpha^2(f) = \alpha^2(\alpha-1)^2 \left(\sigma_\alpha^2 \int_0^1 f^{1-2\alpha} \mathrm{d}x + \Gamma^2(\alpha+1)\, \Delta_\alpha(F_0, F) \right) \tag{6.8}$$

where σ_α^2 is defined by formula (6.3) and

$$\Delta_\alpha(F_0, F) = \frac{1}{\alpha^2} \int_0^1 \left(\frac{1}{(f(x))^\alpha} - \frac{1}{F(x)} \int_0^x (f(y))^{1-\alpha} \mathrm{d}y \right)^2 f(x)\, \mathrm{d}x \quad \text{for } \alpha \in \mathbb{R} - \{0\}. \tag{6.9}$$

Since Hall (1984) gave no hint about the derivation of his formula, let us mention that (6.8) is obtained if one substitutes ψ_α for ϕ in the expressions (4.22)– (4.24) for $s_j^2(f)$, $j \in \{1, 2, 3\}$, given in the proof of Theorem 4.3, thereby employing the expression

$$\begin{aligned} G(x) &= \alpha E(Z^\alpha) \int_0^x \left(1 - \frac{F f'}{f^2} \right) \frac{1}{f^{\alpha-1}} \, \mathrm{d}y \\ &= \Gamma(\alpha+1) \left((\alpha-1) \int_0^x (f(y))^{1-\alpha} \mathrm{d}y + (f(x))^{-\alpha} F(x) \right) \end{aligned}$$

for $G(x)$ of (4.21) when ϕ is replaced by ψ_α, and then forms the sum $s_1^2(f) + s_2^2(f) + s_3^2(f)$. By (6.7) and (6.8),

$$\sigma_\alpha^2(f) = \sigma_\alpha^2 \int_0^1 f^{1-2\alpha} \mathrm{d}x + \Gamma^2(\alpha+1)\, \Delta_\alpha(F_0, F), \quad \alpha \in (-1/2, \infty) - \{0, 1\}. \tag{6.10}$$

The final, intuitively appealing, form of the asymptotic variance

$$\sigma_\alpha^2(f) = (1 + 2\alpha(2\alpha-1)\, D_{2\alpha}(F_0, F))\, \sigma_\alpha^2 + \Gamma^2(\alpha+1)\, \Delta_\alpha(F_0, F) \tag{6.11}$$

(with $\sigma_\alpha^2(f_0) = \sigma_\alpha^2$ given in (6.3)), follows for $\alpha \in (-1/2, \infty) - \{0, 1\}$ by taking into account the formula for $D_{2\alpha}(F_0, F)$ obtained from (5.18). The peculiar expression $\Delta_\alpha(F_0, F)$ figuring in (6.9) and (6.11) can be better understood if we take into account the following Lemma, after which we extend (6.11) to include also the values $\alpha \in \{0, 1\}$.

Lemma 6.2. If the fixed alternative $F \sim f$ satisfies the assumptions of Theorem 4.3 then the class $\{\Delta_\alpha(F_0, F) : \alpha \in \mathbb{R} - \{0\}\}$ consists of the variances

$$\begin{aligned} \Delta_\alpha(F_0, F) &= \int_0^1 \left(\frac{f^{-\alpha}}{\alpha} - \int_0^1 \frac{f^{-\alpha}}{\alpha} f \, \mathrm{d}y \right)^2 f \, \mathrm{d}x \\ &= \int_0^1 \left(\frac{f^{-\alpha}}{\alpha} \right)^2 f \, \mathrm{d}x - \left(\int_0^1 \frac{f^{-\alpha}}{\alpha} f \, \mathrm{d}x \right)^2 \end{aligned} \tag{6.12}$$

of the functions $f^{-\alpha}(X)/\alpha$ of the random argument X distributed by F. This class is continuously extended to all $\alpha \in \mathbb{R}$ by introducing the variance

$$\begin{aligned} \Delta_0(F_0, F) &= \int_0^1 \left(\ln f - \int_0^1 (\ln f) \, f \, \mathrm{d}y \right)^2 f \, \mathrm{d}x \\ &= \int_0^1 f \ln^2 f \, \mathrm{d}x - \left(\int_0^1 f \ln f \, \mathrm{d}x \right)^2 \end{aligned} \tag{6.13}$$

of the function $\ln f(X)$ of the random argument X introduced above. All $\Delta_\alpha(F_0, F)$, $\alpha \in \mathbb{R}$, are nonnegative measures of divergence of F_0 and F, reflexive in the sense that $\Delta_\alpha(F_0, F) = 0$ if and only if $F = F_0$.

Proof. If $\psi : [0, 1] \mapsto \mathbb{R}$ is continuous then by the assumptions about f

$$\inf_{x \in [0,1]} f(x) > 0 \quad \text{and} \quad \sup_{x \in [0,1]} |\psi(x) \, f(x)| < \infty$$

and, consequently, the function

$$\Psi(x) = \int_0^x \psi(y) \, f(y) \, \mathrm{d}y, \quad x \in (0, 1)$$

is well defined. Since

$$\frac{\mathrm{d}}{\mathrm{d}x} \frac{\Psi^2}{F} = - \left(\frac{\Psi}{F} \right)^2 f + \frac{2 \Psi \psi f}{F}$$

and

$$|\Psi(y)| \le y \sup_{x \in [0,1]} |\psi(x) \, f(x)| \quad \text{as well as} \quad F(y) \ge y \inf_{x \in [0,1]} f(x),$$

the function Ψ satisfies the relation

$$\int_0^1 (\psi - \Psi/F)^2 f \, \mathrm{d}x = \int_0^1 \psi^2 f \, \mathrm{d}x - \left(\int_0^1 \psi \, f \, \mathrm{d}x \right)^2. \tag{6.14}$$

To this end take into account the relations

$$\begin{aligned} \int_0^1 (\psi - \Psi/F)^2 f \, \mathrm{d}x &= \int_0^1 \psi^2 f \, \mathrm{d}x - \int_0^1 \frac{2 \Psi \psi f}{F} \, \mathrm{d}x + \int_0^1 \left(\frac{\Psi}{F} \right)^2 f \, \mathrm{d}x \\ &= \int_0^1 \psi^2 f \, \mathrm{d}x - \left(\frac{\Psi^2(1)}{F(1)} - \lim_{y \downarrow 0} \frac{\Psi^2(y)}{F(y)} \right) \\ &= \int_0^1 \psi^2 f \, \mathrm{d}x - \frac{\Psi^2(1)}{F(1)}. \end{aligned}$$

Now, using (6.14) we obtain (6.12) from the definition (6.9). Since f is assumed to be bounded and bounded away from 0,

$$\begin{aligned}\lim_{\alpha\to 0}\Delta_\alpha(F_0,F) &= \int_0^1\left(\lim_{\alpha\to 0}\frac{f^{-\alpha}-1}{\alpha}-\int_0^1\lim_{\alpha\to 0}\frac{f^{-\alpha}-1}{\alpha}f\,\mathrm{d}y\right)^2 f\,\mathrm{d}x\\ &= \int_0^1\left(\ln f-\int_0^1(\ln f)\,f\,\mathrm{d}y\right)^2 f\,\mathrm{d}x\\ &= \Delta_0(F_0,F)\end{aligned}$$

which proves the continuity at $\alpha = 0$. The reflexivity is clear from (6.12) and (6.13). □

We are now in a position to formulate the general results obtained in this paper regarding the asymptotic normality of spacings-type power divergence statistics U_α from the sets $\mathcal{U}_\alpha = \{R_{\phi_\alpha}, \tilde{R}_{\phi_\alpha}, S_{\phi_\alpha}, \tilde{S}_{\phi_\alpha}, T_{\phi_\alpha}, \tilde{T}_{\phi_\alpha}\}$ for $\alpha > -1/2$ under the assumption of the fixed alternative, thereby specifying the parameters $\mu_\alpha(f)$ and $\sigma_\alpha^2(f)$ in (6.6) for all $\alpha > -1/2$. Inspecting once more formula (6.11), we observe that if $\alpha > -1/2$ differs from 0 and 1, then the asymptotic variance $\sigma_\alpha^2(f)$ under the alternative f exceeds the asymptotic variance $\sigma_\alpha^2 = \sigma_\alpha^2(f_0)$ achieved under the hypothesis $F_0 \sim f_0$ by a linear function of σ_α^2 given by

$$2\alpha(2\alpha-1)\,D_{2\alpha}(F_0,F)\sigma_\alpha^2+\Gamma^2(\alpha+1)\,\Delta_\alpha(F_0,F) \tag{6.15}$$

with the coefficients $D_{2\alpha}(F_0,F)$ and $\Delta_\alpha(F_0,F)$ positive unless $F = F_0$. By using Theorem 5.1, we can now find the formulas for $\sigma_0^2(f)$ and $\sigma_1^2(f)$ which are missing in (6.10) by taking limits in (6.11) for $\alpha \to 0$ and $\alpha \to 1$. Since the limits σ_0^2 and σ_1^2 were already calculated in Theorem 6.1, and the limit $\Delta_0(F_0,F)$ is clear from Lemma 6.2, we obtain

$$\sigma_0^2(f)=\lim_{\alpha\to 0}\sigma_\alpha^2(f)=\sigma_0^2+\Delta_0(F_0,F) \tag{6.16}$$

and

$$\sigma_1^2(f)=\lim_{\alpha\to 1}\sigma_\alpha^2(f)=(1+2D_2(F_0,F))\,\sigma_1^2+\Delta_1(F_0,F) \tag{6.17}$$

where (cf (6.12))

$$\Delta_1(F_0,F)=\int_0^1\frac{1}{f}\mathrm{d}x-1. \tag{6.18}$$

Together with (6.10), (6.16) and (6.17) provide formulas for $\sigma_\alpha^2(f)$ for all $\alpha > -1/2$. It is clear that $\sigma_0^2(f)$ and $\sigma_1^2(f)$ are of the form (6.11), so that the representation (6.11) holds for all $\alpha > -1/2$. We summarize our results as follows.

Theorem 6.2. If the alternative $F \sim f$ satisfies the assumptions of Theorem 4.3, then the asymptotic formula of (6.6) is valid for all $\alpha > -1/2$. The asymptotic means $\mu_\alpha(f)$ are given by the explicit formulas (5.14) – (5.18). The asymptotic variances $\sigma_\alpha^2(f)$ are given by

(6.11), where the explicit formulas for $D_{2\alpha}(F_0, F)$ can be found in (5.16)−(5.18), those for σ_α^2 in (6.3) and (6.4), and the formulas for $\Delta_\alpha(F_0, F)$ in (6.12) and (6.13). The asymptotic means and variances are continuous in the variable $\alpha \in (-1/2, \infty)$. The asymptotic means satisfy the inequality $\mu_\alpha(f) \geq \mu_\alpha$ mentioned in Theorem 5.2. The asymptotic variances satisfy the inequality $\sigma_\alpha^2(f) \geq \sigma_\alpha^2$. Both inequalities become equalities if and only if $F = F_0$.

Proof. The proof should be clear from what was said above. The inequality $\sigma_\alpha^2(f) \geq \sigma_\alpha^2$ and the condition for equality follow from (6.11), because $D_{2\alpha}(F_0, F)$ and $\Delta_\alpha(F_0, F)$ are nonnegative measures of divergence of F_0 and F, which are equal to zero if and only if $F = F_0$, in which case the excess function (6.15) is 0. □

Concrete forms of $\mu_\alpha(f)$ and $\sigma_\alpha^2(f_0) = \sigma_\alpha^2$ were illustrated in Tables 5.1 and 6.1. The next table illustrates $\sigma_\alpha^2(f)$ given by (6.11) for arbitrary f satisfying the assumptions of Theorem 4.3 and selected values of α. In each line of Table 6.2 two expressions for $\sigma_\alpha^2(f)$ are given: the first one is obtained by substituting α in (6.11), the second one by actually calculating $D_{2\alpha}(F_0, F)$ and $\Delta_\alpha(F_0, F)$ in each case and putting the results in a closed form. As presumed, for $f = 1$ the illustrated values reduce to σ_α^2 from Table 6.1.

Table 6.2: Asymptotic variances $\sigma_\alpha^2(f)$ for selected $\alpha > -1/2$.

α		$\sigma_\alpha^2(f)$	
0	$\sigma_0^2 + \Delta_0(F_0, F)$	$=$	$\frac{\pi^2}{6} - 1 + \int_0^1 f \ln^2 f \mathrm{d}x - \left(\int_0^1 f \ln f \, \mathrm{d}x\right)^2$
1	$[1 + \chi^2(F_0, F)]\, \sigma_1^2 + \Delta_1(F_0, F)$	$=$	$\int_0^1 \frac{\mathrm{d}x}{f} \left(\frac{\pi^2}{3} - 2\right) - 1$
2	$[1 + 12 D_4(F_0, F)\, \sigma_2^2 + 4\Delta_2(F_0, F)]$	$=$	$2 \int_0^1 \frac{\mathrm{d}x}{f^3} - \left(\int_0^1 \frac{\mathrm{d}x}{f}\right)^2$
3	$[1 + 30 D_6(F_0, F)]\, \sigma_3^2 + 36\Delta_3(F_0, F)$	$=$	$14 \int_0^1 \frac{\mathrm{d}x}{f^5} - 4 \left(\int_0^1 \frac{\mathrm{d}x}{f^2}\right)^2$

7. Discussion

The general form of the asymptotic normality (6.6) established by Theorem 6.2, as well as the continuity of the asymptotic means and variances $\mu_\alpha(f)$ and $\sigma_\alpha^2(f)$ in the parameter $\alpha > -1/2$ proved in Theorem 5.1, and the explicit formulas (5.14) and (6.11) for these parameters for general α seem to be new results. However, in the references cited in Sections 1 and 2 one can find particular versions of these results for some of the statistics

U_α from the set $\{R_{\phi_\alpha}, \tilde{R}_{\phi_\alpha}, S_{\phi_\alpha}, \tilde{S}_{\phi_\alpha}, T_{\phi_\alpha}, \tilde{T}_{\phi_\alpha}\}$ or their linear functions, and for some $\alpha > -1/2$ and some distributions $F \sim f$.

Let us start with the statistic $\mathcal{M} = S_{\phi_0}$ proposed by Moran (1951) and given in (5.5). The asymptotic normality (6.6) for $\alpha = 0$, $U_0 = S_{\phi_0}$ and $f = f_0 \equiv 1$, with the parameters $\mu_0(f_0) = \mu_0$ and $\sigma_0^2(f_0) = \sigma_0^2$ given in Tables 5.1 and 6.1, was proved by Darling (1953), yielding specifically that under $\mathcal{H}_0$

$$\frac{1}{\sqrt{n}}\,(\mathcal{M} - n\gamma) \xrightarrow{\mathcal{D}} N\left(0, \frac{\pi^2}{6} - 1\right) \quad \text{as } n \to \infty. \tag{7.1}$$

The result of Darling was extended to all positively valued step functions f on $[0,1]$ by Cressie (1976), who also obtained $\mu_0(f)$ and $\sigma_0^2(f)$ given in Tables 5.1 and 6.2. The result of Cressie was extended by van Es (1992) to the alternatives f considered in the present paper which satisfy a Lipschitz condition on $[0,1]$, and to all f considered in this paper by Shao and Hahn (1995). Cressie(1976) and van Es(1992) studied S_{ϕ_0} as the special case obtained for $m = 1$ from a more general statistic based on m-spacings with $m \geq 1$. Van Es extended ideas and methods developed for $m > 1$ by Vasicek (1976) and Dudewicz and van der Meulen (1981) for proving the consistency and asymptotic normality of a spacings-based estimator of entropy. The latter authors considered only $\phi(t) = \phi_0(t) = -\ln t$.

Greenwood (1946) introduced the statistic

$$\mathcal{G} = \sum_{j=1}^{n+1} (Y_j - Y_{j-1})^2 = \frac{2S_{\phi_2} + n + 1}{(n+1)^2}, \tag{7.2}$$

given above in (5.4). Kimball (1950) proposed the generalization

$$\sum_{j=1}^{n+1} (Y_j - Y_{j-1})^\alpha = \frac{\alpha(\alpha-1)S_{\phi_\alpha} + n + 1}{(n+1)^\alpha}, \qquad \alpha > 0, \tag{7.3}$$

and Darling (1953) proved an asymptotic normality theorem equivalent to (6.6) for $U_\alpha = S_{\phi_\alpha}$, $\alpha \in (0,\infty) - \{1\}$, and $f = f_0 \equiv 1$. Weiss (1957) extended this result of Darling to positive piecewise constant densities f. Hall (1984) obtained the asymptotic normality

$$\frac{1}{\sqrt{n}}\left(\tilde{U}_\alpha - n\alpha(\alpha-1)\,\mu_\alpha(f) - n\right) \xrightarrow{\mathcal{D}} N(0, \alpha^2(\alpha-1)^2\sigma_\alpha^2(f)) \quad \text{as } n \to \infty \tag{7.4}$$

for all statistics

$$\begin{aligned} \tilde{U}_\alpha &= \sum_{j=2}^{n} (n(Y_j - Y_{j-1}))^\alpha \\ &= \alpha(\alpha-1)\,\tilde{R}_{\phi_\alpha} - \alpha n(1 - Y_n + Y_1) + n + \alpha - 1 = \alpha(\alpha-1)\,\tilde{R}_{\phi_\alpha} + n + O_p(1) \end{aligned}$$

with $\alpha \in (-1/2, \infty) - \{0, 1\}$ for any f considered in Theorem 6.2. Here $\mu_\alpha(f)$ and $\sigma_\alpha^2(f)$ are the same as in Theorem 6.2, with $\mu_\alpha(f)$ given by the right-hand side of (5.20) and $\sigma_\alpha^2(f)$ by (6.11), $\tilde{R}_{\phi_\alpha}$ is defined as in (2.16) with $\phi = \phi_\alpha$, and the $O_p(1)$ statement follows from the proof of Theorem 3.1. In fact, this result of Hall (1984) was one of the arguments used in the proof of Theorem 6.2.

The statistic S_{ϕ_1} was proposed recently by Misra and van der Meulen (2001), who proved the asymptotic normality (6.6) for $U_1 = S_{\phi_1}$ and any f considered in Theorem 6.2, with the parameters $\mu_1(f)$ and $\sigma_1^2(f)$ given in Tables 5.1 and 6.2, yielding the result

$$\frac{1}{\sqrt{n}}\left(S_{\phi_1} - n\left(1 - \gamma - \int_0^1 \ln f \, \mathrm{d}x\right)\right) \xrightarrow{\mathcal{D}} N\left(0, \int_0^1 \left(\frac{\pi^2}{3} - 2\right)\frac{\mathrm{d}x}{f} - 1\right) \quad \text{as } n \to \infty. \tag{7.5}$$

We see that the present Theorem 6.2 unifies and extends the results proved separately in the literature in three different situations for two particular statistics from the set (5.3). The formulas for all asymptotic parameters $\mu_\alpha(f)$ and $\sigma_\alpha^2(f)$ of the statistics U_α are shown to follow via the asymptotic equivalence of these statistics (cf Theorem 3.1) and the continuity of these parameters in α (cf Theorem 5.1) from Hall's formula (cf (7.4)) for the asymptotic parameters of $\tilde{U}_\alpha$ with $\alpha \in (-1/2, \infty)$ different from 0 and 1.

8. Acknowledgements

The research for this paper was supported by GAČR grant 102/07/1131, MŠMT grant 1M 0572 and K. U. Leuven project GOA/98/06.

9. References

Cressie, N. (1976) On the logarithms of high-order spacings. *Biometrika*, **63**, 345–355.

Cressie, N. (1978) Power results for tests based on high-order gaps. *Biometrika*, **65**, 214–218.

Cressie, N. (1979) An optimal statistic based on higher order gaps. *Biometrika*, **66**, 619–627.

Csiszár, I. (1963) Informationstheoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizität von Markoffschen Ketten. *Publ. Math. Inst. Hungarian Acad. Sci., Ser. A*, **8**, 85–108.

Darling, D. A. (1953) On a class of problems related to the random division of an interval. *Ann. Math. Statist.*, **24**, 239–253.

Del Pino, G. E. (1979) On the asymptotic distribution of k-spacings with applications to goodness of fit tests. *Ann. Statist.*, **7**, 1058–1065.

Dudewicz, E. J. and van der Meulen, E. C. (1981) Entropy-based tests of uniformity. *J. Amer. Statist. Assoc.*, **76**, 967–974.

Durbin, J. (1973) *Distribution theory for tests based on the sample distribution function.* Regional Conference Series in Applied Mathematics, Vol. 9. Philadelphia: SIAM.

Ekström, M. (1999) Strong limit theorems for sums of logarithms of high order spacings. *Statistics*, **33**, 153–169.

Greenwood, M. (1946) The statistical study of infectious diseases. *J. Roy. Statist. Soc., Ser. A*, **109**, 85–110.

Hall, P. (1984) Limit theorems for sums of general functions of m-spacings. *Math. Proc. Cambridge Philos. Soc.*, **96**, 517–532.

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

Jammalamadaka, S. R. and Tiwari, R. C. (1986) Efficiencies of some disjoint spacings tests relative to a χ^2 test. In M.L.Puri, J.P. Vilaplana and W.Wertz (eds), *New Perspectives in Theoretical and Applied Statistics*, 311-318. New York: Wiley.

Jammalamadaka, S. R., Zhou, X. and Tiwari, R. C. (1989) Asymptotic efficiency of spacings tests for goodness of fit. *Metrika*, **36**, 355-377.

Jimenez, R. and Shao, Y. (2009) Limit theorems for ϕ-divergences based on k-spacings. *Comm. Statist. Theory Methods*, **38**, 695–710.

Kimball, B. F. (1947) Some basic theorems for developing tests of fit for the case of nonparametric probability distribution functions. *Ann. Math. Statist.*, **18**, 540–548.

Kimball, B. F. (1950) On the asymptotic distribution of the sum of powers of unit frequency differences. *Ann. Math. Statist.*, **21**, 263–271.

Kuo, M. and Rao, J. S. (1981) Limit theory and efficiencies for tests based on higher order spacings. In *Statistics – Applications and New Directions.* Proc. Golden Jubilee Conference of the Indian Statistical Institute, 333–352. Calcutta: Statistical Publishing Society.

Liese, F. and Vajda, I. (1987) *Convex Statistical Distances.* Leipzig: Teubner.

Liese, F. and Vajda, I. (2006) On divergences and informations in statistics and information theory. *IEEE Trans. Inform. Theory*, **52,** 4394-4412.

Lindsay, G. G. (1994) Efficiency versus robustness. The case of minimum Hellinger distance and other methods. *Ann. Statist.*, **22**, 1081–1114.

Menéndez, M., Morales, D. Pardo, L. and Vajda, I. (1998) Two approaches to grouping of data and related disparity statistics. *Comm. Statist. Theory Methods*, **27**, 609–633.

Misra, N. and van der Meulen, E. C. (2001) A new test of uniformity based on overlapping simple spacings. *Comm. Statist. Theory Methods*, **30**, 1435–1470.

Morales, D. Pardo, M. C. and Vajda, I.(2003) Limit laws for disparities of spacings. *J. Nonparametr. Statist.*, **15**, 325–342.

Moran, P. A. P. (1951) The random division of an interval –Part II. *J. Roy. Statist. Soc. Ser. B*, **13**, 147–150.

Pyke, R. (1965) Spacings. *J. Roy. Statist. Soc. Ser. B*, **27**, 395–449.

Rao, J. S. and Kuo, M. (1984) Asymptotic results on Greenwood statistic and some of its generalizations. *J. Roy. Statist. Soc. Ser. B*, **46**, 228–237.

Read, T. R. C. and Cressie, N. (1988) *Goodness of Fit Statistics for Discrete Multivariate Data*. Berlin: Springer.

Shao, Y. and Hahn, M. G. (1995) Limit theorems for the logarithm of sample spacings. *Statist. Probab. Lett.*, **24**, 121–132.

Vajda, I. (2007) Asymptotic comparisons ofof divergence-based goodness-of-fit statistics. *Publ. Stastist. Inst. Univ. Paris* , **51**, 49-66.

Vajda, I. and van der Meulen, E. c. (2006) On estimation and testing based on m-spacings. *Prague Stochastics 2006* (Eds Hušková M., Janžura M.), Matfyzpress, Prague.

van Es, B. (1992) Estimating functionals related to a density by a class of statistics based on spacings. *Scand. J. Statist.*, **19**, 61–72.

Vasicek, O. (1976) A test for normality based on sample entropy. *J. Roy. Statist. Soc. Ser. B*, **38**, 54-59.

Weiss, L. (1957) The asymptotic power of certain test of fit based on sample spacings. *Ann. Math. Statist.*, **28**, 783–786.