Title
Asymptotic efficiencies of spacings tests for goodness of fit

Journal
Metrika, 36(1)

ISSN
0026-1335

Authors
Jammalamadaka, S Rao
Zhou, Xian
Tiwari, Ram C

Publication Date
1989-12-01

DOI
10.1007/bf02614112

Peer reviewed

Metrika (1989) 36:355–377

Asymptotic Efficiencies of Spacings Tests for Goodness of Fit

By S. R. Jammalamadaka[1], X. Zhou[1] and R. C. Tiwari[2]

Abstract: Tests based on higher-order or m-step spacings have been considered in the literature for the goodness of fit problem. This paper studies the asymptotic distribution theory for such tests based on non-overlapping m-step spacings when m, the length of the step, also increases with the sample size n, to inifinity. By utilizing the asymptotic distributions under a sequence of close alternatives and studying their relative efficiencies, we try to answer a central question about the choice of m in relation to n. Efficiency comparisons are made with tests based on overlapping m-step spacings, as well as corresponding chi-square tests.

Key words: m-step spacings, limit distributions, asymptotic efficiencies, goodness of fit, chi-square.

1 Introduction

Let $X_1, \ldots, X_{n-1}$ be independently and identically distributed (i.i.d) real-valued random variables (r.v.'s) with a common continuous distribution function (d.f.) F, and let $X_{(1)} \leqslant \ldots \leqslant X_{(n-1)}$ be the corresponding order statistics. For the goodness of fit problem of testing whether F is equal to some specified d.f. F_0 (say), without loss of any generality, we can use the transformation $x \to F_0(x)$ on the X_i's and, therefore, assume that F has support on $[0, 1]$ and F_0 is the uniform d.f. on $[0, 1]$.

Let $\{m_\nu\}$ and $\{n_\nu\}$ be the two nondecreasing sequences of positive integers with $m_\nu, n_\nu \to \infty$ as $\nu \to \infty$. The *non-overlapping* (or disjoint) m-step spacings (or m-spacings) are defined as

$$D_{im}^{(m)} = X_{(im+m)} - X_{(im)} \qquad i = 0, 1, \ldots, \lambda - 1 \tag{1.1}$$

[1] S. Rao Jammalamadaka and Xian Zhou, Statistics and Applied Probability Program, University of California at Santa Barbara, Santa Barbara, CA 93106, USA.
[2] Ram C. Tiwari, Department of Mathematics, University of North Carolina, Charlotte N.C. 28223, USA.

0026–1335/89/6/355–377 $2.50

where $\lambda = \lambda_\nu = n_\nu/m_\nu$ is assumed without loss of generality to be an integer and $X_{(0)} = 0, X_{(n)} = 1$.

Under the null hypothesis of uniform distribution on [0, 1], we use the special notation $U_{(i)}$ (replacing $X_{(i)}$) for uniform order statistics and $T_{im}^{(m)}$ (replacing $D_{im}^{(m)}$) for uniform spacings. For testing the null hypothesis

$$H_0 : F_0(x) = x \quad 0 \leqslant x \leqslant 1 \tag{1.2}$$

against a sequence of alternatives

$$H_{1\nu} : F_\nu(x) = x + r(\nu)L_\nu(x) \quad 0 \leqslant x \leqslant 1 \tag{1.3}$$

with $L_\nu(0) = L_\nu(1) = 0$, and $r(\nu) \to 0$ as $\nu \to \infty$, we consider the class of statistics

$$V_\nu = \sum_{i=0}^{\lambda-1} h_\nu(\lambda D_{im}^{(m)}) \tag{1.4}$$

where $h_\nu(\cdot)$ are real valued functions defined on $(0, \infty)$. The asymptotic distribution theory of V_ν for fixed m is given by del Pino (1977). In this paper, we derive the asymptotic distributions of V_ν under the null hypothesis H_0 as well as under the alternatives $H_{1\nu}$ when both m and n tend to infinity. Then, we compare the Pitman Asymptotic Relative Efficiencies (ARE's) for the pair

(i) $V_{\nu 1} = \sum_{i=0}^{\lambda-1} h_\nu^{(1)}(\lambda D_{im}^{(m)}), \quad V_{\nu 2} = \sum_{i=0}^{\lambda-1} h_\nu^{(2)}(\lambda D_{im}^{(m)})$

where $\{h_\nu^{(1)}\}$ and $\{h_\nu^{(2)}\}$ are two different sequences of functions. Also for two different sequences of integers $\{m_\nu\}$ and $\{m'_\nu\}$, we compare the Pitman efficiencies of

(ii) $V_\nu = \sum_{i=0}^{\lambda-1} h_\nu(\lambda D_{im}^{(m)}), \quad V'_\nu = \sum_{i=0}^{\lambda'-1} h_\nu(\lambda' D_{im'}^{(m')})$

where $m = cn^p(1 + o(1))$, $c > 0$, $0 < p < 1$, $m' = c'n^{p'}(1 + o(1))$, $c' > 0$, $0 < p' < 1$, and $\lambda' = n/m'$. Here $o(1)$ stands for a quantity going to zero as $\nu \to \infty$. We show in section 4 that the ARE in case (i) is $\text{ARE}(V_{\nu 1}, V_{\nu 2}) = 1$ and in case (ii),

$$\mathrm{ARE}(V_\nu, V_\nu') = \begin{cases} \infty & \text{if } p' < p \\ (c/c')^{1/(1+p)} & \text{if } p' = p \\ 0 & \text{if } p' > p. \end{cases}$$

Tests based on overlapping spacings have also been studied in the literature. The overlapping spacings are defined by

$$D_i^{(m)} = X_{(i+m)} - X_{(i)} \qquad i = 0, 1, \ldots, n-1.$$

where $X_{(k)} = 1 + X_{(n-k)}$ for $k \geqslant n$ circularly, for convenience. The corresponding statistic is

$$V_\nu^* = \sum_{i=0}^{n-1} h_\nu(\lambda D_i^{(m)}).$$

Kuo and Rao (1979, 1981) studied the general asymptotic theory and the ARE of this class of statistics V_ν^* for fixed m; see also Cressie (1976) for a special case. Beirlant and van Zuijlen (1985) consider the weak convergence of the empirical spacings process under the *null* hypothesis when $m \to \infty$. Hall (1986) obtained the limiting distributions of V_ν^* under a sequence of alternatives. In Section 4.3, we show that for the same m_ν, V_ν^* is more efficient than V_ν, which should be expected because V_ν^* uses more information from the data than V_ν does. However, by choosing the size of the step m_ν corresponding to V_ν to be larger than that corresponding to V_ν^*, V_ν can be made as efficient as or even more efficient than V_ν^*. On the other hand, V_ν involves considerably less calculations than V_ν^* does and may be preferable from a computational and practical point of view.

In Section 4.4, the efficiencies of V_ν are also compared with yet another test namely the chi-square test $\sum_{i=0}^{\lambda-1} O_{i\lambda}^2$, considered by Quine and Robinson (1985), where $\{O_{i\lambda}\}$ are the frequencies in the cells $\left[\frac{i}{\lambda}, \frac{i+1}{\lambda}\right]$, $i = 0, 1, \ldots, \lambda - 1$.

The organization of the paper is as follows: In Section 2, the asymptotic distribution of V_ν under H_0 is derived while its distribution under $H_{1\nu}$ is given in Section 3. Section 4 contains a discussion of the ARE's. Finally in Section 5, the "asymptotic sufficiency" of m-spacings in this particular testing context and the comments on the choice of m in relation to n are discussed. Proofs are postponed to the last section, namely Section 6.

A few words about the notations: "$X \sim Y$" will mean that the r.v. X has the same d.f. as the r.v. Y, while "$X \sim F$" means that X has d.f. F. We write "$\xrightarrow{d}$" and "$\xrightarrow{P}$" to denote convergence in distribution and in probability, respectively. $N(\mu, \sigma^2)$ stands for a normal distribution with mean μ and variance σ^2. Throughout $Z_1, Z_2, \ldots$ are i.i.d. exponential r.v.'s with mean equal to 1, and

$$\bar{Z}_n = \frac{1}{n} \sum_{i=1}^{n} Z_i, \quad \bar{Z}_{im}^{(m)} = \frac{1}{m} \sum_{j=1}^{m} Z_{im+j}. \tag{1.5}$$

S_m is a gamma r.v. with shape parameter m and scale parameter 1. We assume that all the r.v.'s are defined on a common probability space. The notation $X_\nu = o_p(r_\nu)$ will mean that $|X_\nu/r_\nu| \xrightarrow{P} 0$ and $X_\nu = O_p(r_\nu)$ means that for any $\epsilon > 0$ there exist $M > 0$ and $\nu_0 > 0$ such that $P(|X_\nu/r_\nu| > M) < \epsilon$ for all $\nu > \nu_0$, while $O(r_\nu)$ is a sequence of numbers such that $O(r_\nu)/r_\nu$ is bounded. The first three derivatives of a function f are denoted by f', f'' and f''' respectively. The subscript ν will be omitted for notational convenience, when there is no possible confusion.

2 Asymptotic Null Distribution of V_ν when $\lambda_\nu \to \infty$

The asymptotic distribution of V_ν under H_0 as $\lambda_\nu \to \infty$ is given by

Theorem 2.1: Suppose the following assumptions hold:

(A.1) there exist M and $\delta > 0$ such that $|h_\nu'''(x)| \leqslant M$ for $x \in [1 - \delta, 1 + \delta]$ and for all ν,

(A.2) $h_\nu''(1) \to b$ where $|b| < \infty$, as $\nu \to \infty$, and

(A.3) $n/m^k \to 0$ for some positive integer k. Then,

$$\frac{m}{\sqrt{\lambda}} (V_\nu - \lambda \mu_\nu) \xrightarrow{d} N\left(0, \frac{1}{2} b^2\right),$$

where $\mu_\nu = E\left[h_\nu\left(\frac{1}{m} S_m\right)\right]$. □

The proof of Theorem 2.1 will be given in Section 6 and requires several lemmas.

The case where $h_\nu \equiv h$ is of special interest. In such a case, since $h_\nu''(1) \equiv h''(1)$, Assumption (A.2) of Theorem 2.1 can be removed. Note also that if h''' is bounded on $(0, \infty)$, then (A.3) is not required. Thus we get

Corollary 2.1: If either

(a) h''' is bounded in $[1-\delta, 1+\delta]$ for some $\delta > 0$ and $n/m^k \to 0$ for some k, or

(b) h''' is bounded in $(0, \infty)$, then

$$\frac{m}{\sqrt{\lambda}}(V_\nu - \lambda\mu_\nu) \xrightarrow{d} N\left(0, \frac{1}{2}\left[h''(1)\right]\right).$$

Example 1: Let $h(x) = x^2$, then the condition (b) of corollary 2.1 is trivially satisfied, hence we have

$$\frac{m}{\sqrt{\lambda}} \sum_{i=0}^{\lambda-1} \left[(\lambda T_{im}^{(m)})^2 - 1 - \frac{1}{m}\right] \xrightarrow{d} N(0, 2).$$

Example 2: Let $h(x) = x \log x$. Then $\mu_\nu = m\psi(m+1) - \log m$ where $\psi(m+1) = \sum_{i=1}^{m+1} \frac{1}{j} - \gamma$ and γ is the Euler's constant. Clearly the first part of condition (a) of Corollary 2.1 holds. Thus if $n/m^k \to 0$ for some k, then

$$\frac{m}{\sqrt{\lambda}} \sum_{i=0}^{\lambda-1} [(\lambda T_{im}^{(m)} \log(\lambda T_{im}^{(m)}) + \log m - \psi(m+1))] \xrightarrow{d} N\left(0, \frac{1}{2}\right).$$

Example 2.3: Let $h(x) = \log x$. Similar to Example 2.2, we conclude that if $n/m^k \to 0$ for some k, then

$$\frac{m}{\sqrt{\lambda}} \sum_{i=0}^{\lambda-1} [\log(\lambda T_{im}^{(m)}) + \log m - \psi(m)] \xrightarrow{d} N\left(0, \frac{1}{2}\right).$$

3 Asymptotic Distribution of V_ν Under Alternatives when $\lambda_\nu \to \infty$

Let $D_{im}^{(m)}$, $i = 0, \ldots, \lambda - 1$, be non-overlapping m-step spacings under alternatives

$$H_{1\nu} : F_\nu(x) = x + r(\nu)L_\nu(x). \quad 0 \leqslant x \leqslant 1$$

with $L_\nu(0) = L_\nu(1) = 0$. We shall make the following assumption:

(A.4) $L_\nu' = l_\nu$, $L_\nu'' = l_\nu'$ exist and l_ν converges uniformly to a function l on $[0, 1]$. Note that (A.4) implies that L_ν are uniformly bounded and l is continuous on $[0, 1]$, hence l^2 is Riemann-integrable on $[0, 1]$

Also define

$$T_{im}^{(m)} = F_\nu(X_{(im+m)}) - F_\nu(X_{(im)}) = D_{im}^{(m)} + r(\nu)l_\nu(\tilde{X}_{im})D_{im}^{(m)} \tag{3.1}$$

where $X_{(im)} \leqslant \tilde{X}_{im} \leqslant X_{(im+m)}$.

Theorem 3.1: If $r(\nu) = a_\nu(nm)^{-1/4}$, $a_\nu \to a$ as $\nu \to \infty$, and Assumptions (A.1), (A.2), (A.3) and (A.4) hold, then

$$\frac{m}{\sqrt{\lambda}} \sum_{i=0}^{\lambda-1} [h_\nu(\lambda D_{im}^{(m)}) - h_\nu(\lambda T_{im}^{(m)})] \xrightarrow{P} \frac{1}{2} a^2 b \|l\|_2^2.$$

The proof is given in Section 6.

Combining Theorem 2.1 with Theorem 3.1, we get

Theorem 3.2: Under the same conditions as in Theorem 3.1,

$$\frac{m}{\sqrt{\lambda}} \sum_{i=0}^{\lambda-1} [h_\nu(\lambda D_{im}^{(m)}) - \mu_\nu] \xrightarrow{d} N\left(\frac{1}{2} a^2 b \|l\|_2^2, \frac{1}{2} b^2\right).$$

Corollary 3.1: If h''' is bounded in $[1-\delta, 1+\delta]$ for some $\delta > 0$, and $n/m^k \to 0$ for some k, then under alternatives $H_{1\nu}$ with assumption (A.4) and $r(\nu) = a_\nu (nm)^{-1/4}$, $a_\nu \to a$,

$$\frac{m}{\sqrt{\lambda}} \sum_{i=0}^{\lambda-1} \left[h(\lambda D_{im}^{(m)}) - Eh\left(\frac{1}{m} S_m\right)\right] \xrightarrow{d} N\left(\frac{1}{2} a^2 h''(1) \|l\|_2^2, \frac{1}{2} [h''(1)]^2\right).$$

In particular, we have

$$\frac{m}{\sqrt{\lambda}} \sum_{i=0}^{\lambda-1} [\log (\lambda D_{im}^{(m)}) + \log m - \psi(m)] \xrightarrow{d} N\left(\frac{1}{2} a^2 \|l\|_2^2, \frac{1}{2}\right)$$

and

$$\frac{m}{\sqrt{\lambda}} \sum_{i=0}^{\lambda-1} \left[(\lambda D_{im}^{(m)})^2 - 1 - \frac{1}{m}\right] \xrightarrow{d} N(a^2 \|l\|_2^2, 2).$$

Remark 3.1: The asymptotic normality of $V_\nu = \sum_{i=0}^{\lambda-1} h_\nu(\lambda D_{im}^{(m)})$ no longer holds when m increases in proportion to n. If $\lambda = n/m$ denote this constant proportion, then in this case, we have the results: (proofs of which are omitted for brevity)

(i) Under the hull hypothesis, if Assumption (A.1) and (A.2) hold, then

$$m \sum_{i=0}^{\lambda-1} [h_\nu(\lambda T_{im}^{(m)}) - h_\nu(1)] \xrightarrow{d} \frac{1}{2} b \chi_{\lambda-1}^2$$

where $\chi_{\lambda-1}^2$ is a chi-square distribution with degrees of freedom $(\lambda - 1)$.

(ii) Under alternatives $H_{1\nu}$ (see (1.3)) with $r(\nu) = a_\nu n^{-1/2}$ and $a_\nu \to a > 0$. If (A.1) and (A.2) hold, and $L_\nu(x)$ converges uniformly to a continuous function $L(x)$ on $[0, 1]$, then

$$m \sum_{i=0}^{\lambda-1} [h_\nu(\lambda D_{im}^{(m)}) - h_\nu(1)] \xrightarrow{d} \frac{1}{2} b \chi_{\lambda-1}^2 (a^2 \lambda \Delta^2)$$

where $\Delta^2 = \sum_{i=0}^{\lambda-1}\left[L\left(\frac{i+1}{\lambda}\right) - L\left(\frac{i}{\lambda}\right)\right]^2$ and $\chi^2_{\lambda-1}(a^2\lambda\Delta^2)$ is a non-central chi-square distribution with degree of freedom $\lambda - 1$ and non-centrality parameter $(a^2\lambda\Delta^2)$.

4 Efficiencies of Spacings Tests

4.1 Pitman Relative Efficiency

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 a sequence of alternatives converging to the null hypothesis. Let T_n be a test statistic for testing the hypothesis H_0 against the alternatives

$$H_{1n} : F_n(x) = x + r_n L(x), \quad 0 \leqslant x \leqslant 1 \tag{4.1}$$

with $r_n = a_n n^{-\delta}$ $(a_n \geqslant 0,\ \delta > 0)$, $a_n \to a < \infty$. Suppose $T_n \xrightarrow{d} N(a^2\mu, \sigma^2)$ under H_{1n}. Then we define the "efficacy" of T_n by $\mathrm{Eff}(T_n) = (\mu^2/\sigma^2)^{1/4\delta}$. It can be shown that the ARE of a test T_{1n} relative to another test T_{2n} is

$$\mathrm{ARE}(T_{1n}, T_{2n}) = \mathrm{Eff}(T_{1n})/\mathrm{Eff}(T_{2n}) \tag{4.2}$$

provided $\mathrm{Eff}(T_{2n})$ and $\mathrm{Eff}(T_{1n})$ are not both zeros.

4.2 Efficiencies of Non-Overlapping Spacings Tests

Now we consider the ARE of the following pairs of test statistics:

(i) $V_{\nu 1} = \sum_{i=0}^{\lambda-1} h_\nu^{(1)}(\lambda D_{im}^{(m)})$ and $V_{\nu 2} = \sum_{i=0}^{\lambda-1} h_\nu^{(2)}(\lambda D_{im}^{(m)})$

(ii) $V_\nu = \sum_{i=0}^{\lambda-1} h_\nu(\lambda D_{im}^{(m)})$ and $V'_\nu = \sum_{i=0}^{\lambda'-1} h_\nu(\lambda' D_{im'}^{(m')})$

where $m = cn^p(1 + o(1)), m' = c'n^{p'}(1 + o(1))$, and $\lambda' = n/m'$ (4.3)

with $c, c' > 0, p, p' \in (0, 1)$.

From Theorem 3.2, it is easy to see that

$$\mathrm{Eff}(V_\nu) = \left[\frac{\frac{1}{4}c\|l\|_2^4 b^2}{\frac{1}{2}b^2}\right]^{1/(1+p)} = \left(\frac{1}{2}c\|l\|_2^4\right)^{1/(1+p)}.$$

Thus we get

Theorem 4.1: For m and m' given by (4.3)

(i) $\mathrm{ARE}(V_{\nu 1}, V_{\nu 2}) = 1$ for any $h_\nu^{(1)}, h_\nu^{(2)}$ satisfying the conditions of Theorem 3.2;

$$\text{(ii) } \mathrm{ARE}(V_\nu, V'_\nu) = \begin{cases} \infty & \text{if } p' < p \\ (c/c')^{1/(1+p)} & \text{if } p' = p \\ 0 & \text{if } p' > p. \end{cases}$$

4.3 Efficiency Comparison Between V_ν and V_ν^*

Now consider the ARE between the non-overlapping spacings test V_ν and the overlapping spacings test V_ν^*. First we state a theorem for the limiting distribution of V_ν^*.

Theorem 4.2: With the same conditions as in Theorem 3.1,

$$\frac{1}{\sqrt{\lambda}} \sum_{i=0}^{n-1} [h_\nu(\lambda D_i^{(m)}) - \mu_\nu] \xrightarrow{d} N\left(\frac{1}{2}a^2 b\|l\|_2^2, \frac{1}{3}b^2\right).$$

Because this theorem does not differ significantly from that proved by P. Hall (1986) and the proof is, in any case, similar to that of our Theorem 3.2 (except for the part corresponding to Lemma 2.4), we omit the proof.

It follows that under the same conditions as in Theorem 3.1, for m and H_{1n} considered above,

$$\mathrm{Eff}(V_\nu^*) = \left(\frac{3}{4} c \|l\|_2^4\right)^{1/(1+p)}$$

Thus we conclude:

(i) If V_ν and V_ν^* are based on the same sequence $m = cn^p(1 + o(1))$, $0 < p < 1$, $c > 0$, then

$$\mathrm{ARE}(V_\nu^*, V_\nu) = (3/2)^{1/(1+p)} > 1$$

i.e. V_ν^* is more efficient than V_ν.

(ii) However, if we increase m to $m' = c'n^p(1 + o(1))$, and call the resulting spacings test V_ν', then

$$\mathrm{ARE}(V_\nu^*, V_\nu') = \left(\frac{3c}{2c'}\right)^{1/(1+p)} \leqslant 1 \quad \text{if } c' \geqslant (3/2)c.$$

This shows that a non-overlapping spacings test with step size of $(3/2)m$ or larger is more efficient than a corresponding overlapping spacings test with step size of m. The fact that the non-overlapping spacings tests are less complicated and easier to compute is a decided advantage.

(iii) If we consider the comparison based on the *same number* of spacings, the non-overlapping spacings will yield more powerful tests. This is not surprising because non-overlapping spacings need a larger sample size than the same number of overlapping spacings. In fact, if we let V_ν and V_ν^{**} respectively denote the non-overlapping spacings test and the corresponding overlapping spacings test with the same number λ of spacings, then V_ν will be able to detect alternatives at a rate of $((\lambda m)m)^{-1/4} = (\lambda m^2)^{-1/4}$, while V_ν^{**} can only detect a rate of $(\lambda m)^{-1/4}$. Thus $\mathrm{ARE}(V_\nu, V_\nu^{**}) = \infty$, i.e., the non-overlapping spacings tests are substantially superior but this comes at the expense of a larger sample size.

4.4 Efficiencies of Spacings Tests Relative to Chi-Square Test

Jammalamadaka and Tiwari (1986) discuss the efficiencies of the well-known Greenwood statistic based on m-spacings relative to a chi-square test for fixed m. Quine and Robinson (1985) considered the ARE of the chi-square test $\sum_{i=0}^{\lambda-1} O_{i\lambda}^2$ relative to the likelihood-ratio test $\sum_{i=0}^{\lambda-1} O_{i\lambda} \log O_{i\lambda}$ where $O_{i\lambda}$ is the frequency of the observations in the interval $\left[\frac{i}{\lambda}, \frac{i+1}{\lambda}\right]$, $i = 0, \ldots, \lambda - 1$. We now consider the ARE of m-spacings tests relative to the chi-square test with $m \to \infty$.

From Quine and Robinson (1985), after a slight generalization, we can get $\frac{1}{\sqrt{\lambda}}\left\{\frac{\lambda}{n}\sum_{i=0}^{\lambda-1} O_{i\lambda}^2 - \lambda - n\right\} \xrightarrow{d} N(a^2c^{1/2}\|l\|_2^2, 2)$ if $\lambda = \frac{1}{2}nq(1 + o(1))$ $(0 < q < 1)$ and $r_n = a_n\lambda^{1/4}n^{-1/2}$, $a_n \to a > 0$. Thus, if $m = cn^p(1 + o(1))$ $(0 < p < 1)$ so that $\lambda = \frac{n}{m} = \frac{1}{c}n^{1-p}(1 + o(1))$, and $r_n = a_n n^{-(1+p)/4}$, then

$$\frac{1}{\sqrt{\lambda}}\left\{\frac{\lambda}{n}\sum_{i=0}^{\lambda-1} O_{i\lambda}^2 - \lambda - n\right\} \xrightarrow{d} N(a^2c^{1/2}\|l\|_2^2, 2),$$

giving $\mathrm{Eff}\left(\sum_{i=0}^{\lambda-1} O_{i\lambda}^2\right) = \left(\frac{1}{2}c\|l\|_2^4\right)^{1/(1+p)}$, hence $\mathrm{ARE}\left(V_\nu, \sum_{i=0}^{\lambda-1} O_{i\lambda}^2\right) = 1$.

5 The Asymptotic Sufficiency of Spacings

In this section, we consider the asymptotic sufficiency of non-overlapping m-step spacings $\{D_{im}^{(m)}\}$. Since clearly the sufficiency of spacings $\{D_{im}^{(m)} : i = 0, \ldots, \lambda - 1\}$ is equivalent to that of $\{X_{(\mathrm{im})} : i = 1, \ldots, \lambda - 1\}$, we will show that $\{X_{(im)} : i = 1, \ldots, \lambda - 1\}$ are asymptotically sufficient when the observations arise from the alternative $H_{1\nu}$ in (1.3) with $r(\nu) = (nm)^{-1/4}$. This asymptotic sufficiency implies that there is no big loss of information in using $\{D_{im}^{(m)}\}$ to test H_0 vs $H_{1\nu}$ when n is large. More precisely we show the following:

Let $X_i \sim F_\nu(x) = x + (nm)^{-1/4}L_\nu(x)$, $m = cn^p(1 + o(1))$ with $c > 0$, $0 < p < 1$, and let f_ν be the joint density of $\{X_{(1)}, \ldots, X_{(n-1)}\}$. If $L_\nu' = l_\nu$ are "sufficiently smooth" (which will be made precise later), then, with $Y_{im} = X_{(im)}$, $i = 1, \ldots, \lambda - 1$, we can generate Y_j for $j \notin \{m, 2m, \ldots, (\lambda-1)m\}$, from the observed $\{Y_m, \ldots, Y_{(\lambda-1)m}\}$ only, such that the joint density g_ν of $\{Y_1, \ldots, Y_{n-1}\}$ satisfies

$$\log \frac{g_\nu(Y_1, \ldots, Y_{n-1})}{f_\nu(Y_1, \ldots, Y_{n-1})} \xrightarrow{P} 0 \quad \text{as } n \to \infty.$$

The construction of g_ν follows from Weiss (1974):

Given $\{Y_m, \ldots, Y_{(\lambda-1)m}\}$, the joint conditional distributions of $\{Y_{im+j} : j = 1, \ldots, m-1\}$ are the same as the distributions of ordered values of $m-1$ i.i.d. uniform r.v.'s on (Y_{im}, Y_{im+m}), $i = 0, \ldots, \lambda - 1$, with $Y_0 = 0$, $Y_n = 1$.

For such a joint density g_ν, F_ν and m given above, we can prove the following theorem:

Theorem 5.1: Suppose there exist an integer $r \geqslant 3$ such that $p < \dfrac{4r-3}{4r-1}$ and $l_\nu(x)$ have up to the r-th continuous and uniformly bounded derivatives on $[0, 1]$. Then

$$\log \frac{g_\nu(Y_1, \ldots, Y_{n-1})}{f_\nu(Y_1, \ldots, Y_{n-1})} \xrightarrow{P} 0 \quad \text{as } n \to \infty$$

and the convergence rate of $\log(g_\nu/f_\nu)$ to zero is determined by $n^{-3(1-p)/2}$ if $p > 3/5$, or by n^{-p} if $p \leqslant 3/5$.

Proof: The arguments are similar to that of Weiss (1974). The details are omitted here.

Corollary 5.1: The rate at which $\log(g_\nu/f_\nu)$ converges to zero is maximized when $p = 3/5$.

Proof: From Theorem 5.1 we see that maximizing the convergence rate of $\log(g_\nu/f_\nu)$ is equivalent to maximizing

$$\varphi(p) \stackrel{\text{def}}{=} \begin{cases} 3(1-p)/2 & \text{if } p > 3/5 \\ p & \text{if } p \leqslant 3/5 \end{cases}.$$

Clearly $\varphi(p)$ reaches its maximum value when $p = 3/5$.

Conclusion About the Choice of m

1. We know that for fixed m, spacings tests, V_ν have no power against the alternatives which are at a distance of the hypothesis by $n^{-\delta}$ when $\delta > 1/4$ (see e.g., Kuo and Rao 1981). Our results indicate that this can be improved by allowing m to increase with n, to infinity. Section 3 shows that V_ν in (1.4) can always distinguish the alternatives which are apart from the hypothesis by $(nm)^{-1/4}$. Thus, if the alternatives are $F_\nu(x) = x + n^{-\delta} L_\nu(x)$ for $\delta \in \left(\frac{1}{4}, \frac{1}{2}\right)$, the choice $m = 0(n^{4\delta - 1})$ for the length of the step will keep the power away from the significance level and 1.

2. In section 4, we say that the larger m gives higher efficiency (in the Pitman sense) even when m depends on n. Therefore, for high efficiency, one should choose $m = 0(n^p)$ with p close to 1. As we demonstrate in a companion paper, non-zero power is achieved by tests symmetric in spacings (cf. equation (1.4)) against alternatives at a distance of $n^{-1/2}$, only when $m = 0(n)$ i.e., only for tests based on a finite number of spacings. See Remark 3.1. See also Del Pino (1979) for similar comments. Section 5 shows that such a choice still retains the "asymptotic sufficiency" of the spacings tests under an appropriate sequence of alternatives as long as $p < 1$.

3. However, Corollary 5.1 indicates that under suitable alternaitves, the rate of convergence of $\log (g_\nu / f_\nu)$ to zero is maximized by taking $p = 3/5$, that is, taking $m = 0(n^{3/5})$. Since the rapid convergence of $\log (g_\nu/f_\nu)$ to zero may be interpreted to mean that the loss of information is minimized, this suggests the optimal choice of m to be $m = 0(n^{3/5})$ in this sense. For finite sample sizes, one can do simulation studies, as in Dudewicz and van der Meulen (1981), to find an appropriate step size m.

6 Proofs of Theorems 2.1 and 3.1

Before proving Theorem 2.1, we need Lemmas 2.1–2.4. First observe the distributional equivalence of the normalized uniform spacings

$$\{\lambda T_{im}^{(m)} : i = 0, \ldots, \lambda - 1\} \sim \{\bar{Z}_{im}^{(m)} / \bar{Z}_n : i = 0, \ldots, \lambda - 1\}. \tag{6.1}$$

A Taylor expansion of $h_\nu(\bar{Z}_{im}^{(m)}/\bar{Z}_n)$ around $\bar{Z}_{im}^{(m)}$ yields

$$\frac{m}{\sqrt{\lambda}}(V_\nu - \lambda\mu_\nu) \sim \frac{m}{\sqrt{\lambda}} \sum_{i=0}^{\lambda-1} [h_\nu(\bar{Z}_{im}^{(m)}/\bar{Z}_n) - \mu_\nu] = W_\nu + R_1 + R_2 \tag{6.2}$$

where

$$W_\nu = \frac{m}{\sqrt{\lambda}} \sum_{i=0}^{\lambda-1} [h_\nu(\bar{Z}_{im}^{(m)}) - \mu_\nu - h'_\nu(1)(\bar{Z}_{im}^{(m)} - 1)],$$

$$R_1 = \frac{(\bar{Z}_n - 1)^2}{2\bar{Z}_n^2} \frac{m}{\sqrt{\lambda}} \sum_{i=0}^{\lambda-1} (\bar{Z}_{im}^{(m)})^2 h''_\nu(\alpha_{i\nu}) \tag{6.3}$$

with $\alpha_{i\nu}$ between $\bar{Z}_{im}^{(m)}/\bar{Z}_n$ and $\bar{Z}_{im}^{(m)}$, and

$$R_2 = \sqrt{n}(\bar{Z}_n - 1)\sqrt{m}[h'_\nu(1) - \frac{1}{\lambda\bar{Z}_n} \sum_{i=0}^{\lambda-1} h'_\nu(\bar{Z}_{im}^{(m)})\bar{Z}_{im}^{(m)}]. \qquad \square \tag{6.4}$$

Lemma 2.1: Let $\{\tau_\nu\}$ be a sequence of r.v.'s and $\{\theta_{i\nu}, i = 0, 1, \ldots, \lambda - 1\}$ be a triangular array of r.v. such that $\max_{0 \leqslant i \leqslant \lambda-1} |\theta_{i\nu} - 1| \xrightarrow{P} 0$ as $\nu \to \infty$. If for every $\epsilon > 0$, there exists $\delta > 0$ and M such that

$$P(|\tau_\nu| \geqslant \epsilon, \max_{0 \leqslant i \leqslant \lambda-1} |g_\nu(\theta_{i\nu})| \leqslant M) \to 0 \quad \text{as } \nu \to \infty$$

and $|g_\nu(x)| \leqslant M$ for $x \in [1 - \delta, 1 + \delta]$ and all ν, then $\tau_\nu \xrightarrow{P} 0$.

Proof:

$$P(|\tau_\nu| \geqslant \epsilon) \leqslant P(|\tau_\nu| \geqslant \epsilon, \max_{0 \leqslant i \leqslant \lambda-1} |g_\nu(\theta_{i\nu})| \leqslant M)$$

$$+ P(\max_{0 \leqslant i \leqslant \lambda-1} |g_\nu(\theta_{i\nu})| > M) \to 0.$$

The proof follows from the assumption and the fact that

$$P(\max_{0 \leqslant i \leqslant \lambda-1} |g_\nu(\theta_{i\nu})| > M) \leqslant P(\max_{0 \leqslant i \leqslant \lambda-1} |\theta_{i\nu} - 1| > \delta). \qquad \square$$

Lemma 2.2: If $n/m^k \to 0$ for some $k \geqslant 2$, then

(a) $\max\limits_{0 \leqslant i \leqslant \lambda-1} |\bar{Z}_{im}^{(m)} - 1| \overset{P}{\to} 0$,

(b) $\max\limits_{0 \leqslant i \leqslant \lambda-1} \left| \dfrac{\bar{Z}_{im}^{(m)}}{\bar{Z}_n} - \bar{Z}_{im}^{(m)} \right| \overset{P}{\to} 0.$

Proof:

$$P\{\max_{0 \leqslant i \leqslant \lambda-1} |\bar{Z}_{im}^{(m)} - 1| \geqslant \epsilon\} \leqslant \sum_{i=0}^{\lambda-1} P\{|\bar{Z}_{im}^{(m)} - 1| \geqslant \epsilon\}$$

$$= \lambda P\{(S_m - m)^2 \geqslant m^2\epsilon^2\} \leqslant \frac{\lambda}{(\epsilon m)^{2k-2}} E[(S_m - m)^{2k-2}] = \frac{n}{m^k} 0(1) \to 0$$

(note that $E(S_m - m)^{2k-2} = 0(m^{k-1})$). This proves (a). (b) follows from

$$P\left\{\max_{0 \leqslant i \leqslant \lambda-1} \left| \frac{\bar{Z}_{im}^{(m)}}{\bar{Z}_n} - \bar{Z}_{im}^{(m)} \right| \geqslant \epsilon \right\} \leqslant \sum_{i=0}^{\lambda-1} P\left\{\left| \frac{\bar{Z}_{im}^{(m)}}{\bar{Z}_n} - \bar{Z}_{im}^{(m)} \right| \geqslant \epsilon \right\}$$

$$\leqslant \lambda P\left\{|\bar{Z}_n - 1| \geqslant \frac{\epsilon}{2}\right\} + \sum_{i=0}^{\lambda-1} P\left\{|\bar{Z}_{im}^{(m)} - 1| \geqslant \frac{1}{2}\right\} + \lambda P\left\{|\bar{Z}_n - 1| > \frac{1}{4}\right\} = \frac{n}{m^k} 0(1).$$

Lemma 2.3: If h_ν''' is uniformly bounded in a neighborhood of 1, and if $n/m^k \to 0$ for some $k \geqslant 2$, then $R_1 \overset{P}{\to} 0$, and $R_2 \overset{P}{\to} 0$.

Proof: If $\max\limits_{0 \leqslant i \leqslant \lambda-1} |h_\nu''(\alpha_{i\nu})| \leqslant M$, then from (6.3)

$$|R_1| \leqslant \frac{M}{2} \frac{1}{\bar{Z}_n^2} [\sqrt{n}(\bar{Z}_n - 1)]^2 \frac{1}{\lambda\sqrt{\lambda}} \sum_{i=0}^{\lambda-1} (\bar{Z}_{im}^{(m)})^2 \overset{P}{\to} 0$$

since $[\sqrt{n}(\bar{Z}_n - 1)]^2 \to \chi_1^2$, a chi-square r.v. with 1 degree of freedom, $\bar{Z}_n \xrightarrow{P} 1$ and $\lambda \to \infty$. Thus for every $\epsilon > 0$,

$$P\{|R_1| \geqslant \epsilon, \max_{0 \leqslant i \leqslant \lambda - 1} |h_\nu''(\alpha_{i\nu})| \leqslant M\} \to 0.$$

Note that $\alpha_{i\nu}$ is between $\bar{Z}_{im}^{(m)}/\bar{Z}_n$ and $\bar{Z}_{im}^{(m)}$, hence by Lemma 2.2

$$\max_{0 \leqslant i \leqslant \lambda-1} |\alpha_{i\nu} - 1| \leqslant \max_{0 \leqslant i \leqslant \lambda - 1} \{|\bar{Z}_{im}^{(m)}/\bar{Z}_n - \bar{Z}_{im}^{(m)}| + |\bar{Z}_{im}^{(m)} - 1|\} \xrightarrow{P} 0.$$

Now it follows from Lemma 2.1 that $R_1 \xrightarrow{P} 0$. To show $R_2 \xrightarrow{P} 0$, use the expansion

$$h_\nu'(\bar{Z}_{im}^{(m)}) = h_\nu'(1) + h_\nu''(1)(\bar{Z}_{im}^{(m)} - 1) + \frac{1}{2} h_\nu'''(\beta_{i\nu})(\bar{Z}_{im}^{(m)} - 1)^2,$$

where $\beta_{i\nu}$ is between $\bar{Z}_{im}^{(m)}$ and 1, and a similar argument as above. □

Lemma 2.4:

$$\frac{1}{\sqrt{\lambda}} \sum_{i=0}^{\lambda-1} [m(\bar{Z}_{im}^{(m)} - 1)^2 - 1] \xrightarrow{d} N(0, 2).$$

Proof: Note that $\{\bar{Z}_{im}^{(m)} : i = 0, \ldots, \lambda - 1\}$ are independent,

$$E[m(\bar{Z}_{im}^{(m)} - 1)^2 - 1] = 0,$$

$$\sum_{i=0}^{\lambda-1} \text{var}\,[m(\bar{Z}_{im}^{(m)} - 1)^2] = \lambda\left(2 + \frac{b}{m}\right), \quad \text{and}$$

$$\sum_{i=0}^{\lambda-1} E[m(\bar{Z}_{im}^{(m)} - 1) - 1]^4 = 0(\lambda).$$

Hence the lemma follows from the Lindberg Central Limit Theorem. □

Proof of Theorem 2.1: By (6.2) and Lemmas 2.3 and 2.4

$$\frac{m}{\sqrt{\lambda}} \sum_{i=0}^{\lambda-1} [h_\nu(\bar{Z}_{im}^{(m)}/\bar{Z}_n) - \mu_\nu] = W_\nu + o_p(1) \tag{6.5}$$

where

$$W_\nu = \frac{m}{\sqrt{\lambda}} \sum_{i=0}^{\lambda-1} [h_\nu(\bar{Z}_{im}^{(m)}) - \mu_\nu - h_\nu'(1)(\bar{Z}_{im}^{(m)} - 1)]$$

$$= \frac{1}{2} h_\nu''(1) \frac{1}{\sqrt{\lambda}} \sum_{i=0}^{\lambda-1} [m(\bar{Z}_{im}^{(m)} - 1)^2 - 1] + \frac{1}{6} \frac{m}{\sqrt{\lambda}} \sum_{i=0}^{\lambda-1} [h_\nu'''(\gamma_{i\nu})(\bar{Z}_{im}^{(m)} - 1)^3]$$

$$+ \frac{m}{\sqrt{\lambda}} \sum_{i=0}^{\lambda-1} [h_\nu(1) - \mu_\nu - \frac{1}{2m} h_\nu''(1)]$$

$$= W_\nu^{(1)} + W_\nu^{(2)} + C_\nu, \quad \text{say,} \tag{6.6}$$

with $\gamma_{i\nu}$ between $\bar{Z}_{im}^{(m)}$ and 1. It is easy to see that $E(W_\nu) = E(W_\nu^{(1)}) = 0$ hence $E(W_\nu^{(2)} + C_\nu) = 0$, and if $\max_{0 \leq i \leq \lambda-1} |h_\nu'''(\gamma_{i\nu})| \leq M$ for all ν, then $\text{Var}(W_\nu^{(2)}) \to 0$, implying $W_\nu^{(2)} + C_\nu \xrightarrow{P} 0$ and so from (6.6),

$$W_\nu = W_\nu^{(1)} + o_p(1) \tag{6.7}$$

Finally, since $W_\nu^{(1)} \xrightarrow{d} N\left(0, \frac{1}{2} b^2\right)$ by Lemma 2.4, the theorem follows from (6.5) and (6.7). □

For Theorem 3.1, we again prove Lemma 3.1–3.3 before proving the theorem.

Lemma 3.1: Under the assumptions of Theorem 3.1.

$$\frac{1}{\lambda} \sum_{i=0}^{\lambda-1} l_\nu^2(\bar{X}_{im})(\lambda D_{im}^{(m)})^2 h_\nu''(\lambda T_{im}^{(m)}) \xrightarrow{P} b \|l\|_2^2$$

Proof: From (3.1),

$$D_{im}^{(m)} = T_{im}^{(m)}[1 + r(\nu)l_\nu(\tilde{X}_{im})]^{-1} \tag{6.8}$$

and so

$$\lambda D_{im}^{(m)} h_\nu''(\lambda T_{im}^{(m)}) - b = \{[\lambda T_{im}^{(m)} h_\nu''(\lambda T_{im}^{(m)}) - b] + o(1). \tag{6.9}$$

Now

$$\max_{0 \leqslant i \leqslant \lambda-1} |\lambda T_{im}^{(m)} h_\nu''(\lambda T_{im}^{(m)}) - b| \leqslant \max_{0 \leqslant i \leqslant \lambda-1} |\lambda T_{im}^{(m)} [h_\nu''(\lambda T_{im}^{(m)}) - h_\nu''(1)]|$$

$$+ \max_{0 \leqslant i \leqslant \lambda-1} |h_\nu''(1)(\lambda T_{im}^{(m)} - 1)| + |h_\nu''(1) - b|$$

$$\leqslant \max_{0 \leqslant i \leqslant \lambda-1} |h_\nu'''(\theta_{i\nu}) \lambda T_{im}^{(m)} (\lambda T_{im}^{(m)} - 1)|$$

$$+ |h_\nu''(1)| \max_{0 \leqslant i \leqslant \lambda-1} |\lambda T_{im}^{(m)} - 1| + o(1) \xrightarrow{P} 0$$

(where $\theta_{i\nu}$ is between $\lambda T_{im}^{(m)}$ and 1) by Lemma 2.2 and (6.1). Thus by (6.9),

$$\lambda D_{im}^{(m)} h_\nu''(\lambda T_{im}^{(m)}) \xrightarrow{P} b. \tag{6.10}$$

It follows that

$$\left| \frac{1}{\lambda} \sum_{i=0}^{\lambda-1} l^2(\tilde{X}_{im})(\lambda D_{im}^{(m)})^2 h_\nu''(\lambda T_{im}^{(m)}) - b \|l\|_2^2 \right|$$

$$\leqslant \max_{0 \leqslant i \leqslant \lambda-1} |\lambda D_{im}^{(m)} h_\nu''(\lambda T_{im}^{(m)}) - b| \left| \sum_{i=0}^{\lambda-1} l^2(\tilde{X}_{im}) D_{im}^{(m)} \right|$$

$$+ |b| \left| \sum_{i=0}^{\lambda-1} l^2(\tilde{X}_{im}) D_{im}^{(m)} - \|l\|_2^2 \right| \xrightarrow{P} 0$$

Thus,

$$\frac{1}{\lambda}\sum_{i=0}^{\lambda-1} l_\nu^2(\tilde{X}_{im})(\lambda D_{im}^{(m)})^2 h_\nu''(\lambda T_{im}^{(m)})$$

$$=\frac{1}{\lambda}\sum_{i=0}^{\lambda-1} l^2(\tilde{X}_{im})(\lambda D_{im}^{(m)})^2 h_\nu''(\lambda T_{im}^{(m)}) + o_p(1) \xrightarrow{P} b\|l\|_2^2. \qquad \square$$

Let $U_{(i)} = F_\nu(X_{(i)})$, $i = 0, 1, \ldots, n$, so that $U_{(1)} \leqslant \ldots \leqslant U_{(n-1)}$ are the order statistics from a uniform [0, 1] distribution.

Lemma 3.2: Under the assumptions of Theorem 3.1,

$$Y_\nu \overset{\text{def}}{=} \frac{m^{5/4}}{n^{3/4}}\sum_{i=0}^{\lambda-1}[h_\nu'(\lambda T_{im}^{(m)}) - h_\nu'(1)]\lambda T_{im}^{(m)} l_\nu(\tilde{U}_{im}) \xrightarrow{P} 0$$

where $U_{(im)} \leqslant \tilde{U}_{im} \leqslant U_{(im+m)}$.

Proof: In view of Lemmas 2.1 and 2.2, we can, without any loss of generality, assume that $|h_\nu''(\theta_{i\nu})| \leqslant M$ for all i and ν, where $\theta_{i\nu}$ is either between $\bar{Z}_{im}^{(m)}$ and 1 or between $\bar{Z}_{im}^{(m)}/\bar{Z}_n$ and 1. Also let $|l_\nu'| \leqslant M$, $|l_\nu''| < M$. Let $\zeta_{i\nu} \in \left[\frac{i}{\lambda}, \frac{i+1}{\lambda}\right]$ satisfy $l_\nu(\zeta_{i\nu}) = \lambda\int_{i/\lambda}^{(i+1)/\lambda} l_\nu(u)du$ so that $\sum_{i=0}^{\lambda-1} l_\nu(\zeta_{i\nu}) = 0$. Let G_n be the empirical d.f. of $U_{(1)}, \ldots, U_{(n)}$. Then from the limiting distribution of Kolmogorov-Smirnov statistic (c.f. e.g. Billingsley 1968) we obtain

$$|l_\nu(G_n(\tilde{U}_{im})) - l_\nu(\zeta_{i\nu})| \leqslant M/\lambda, \tag{6.11}$$

$$l_\nu(\tilde{U}_{im}) = l_\nu(G_n(\tilde{U}_{im})) + O_p(n^{-1/2}). \tag{6.12}$$

Hence we can write

$$Y_\nu = Y_{1\nu} + Y_{2\nu} + o_p(1), \tag{6.13}$$

where

$$Y_{1\nu} = \frac{m^{5/4}}{n^{3/4}} \lambda T_{im}^{(m)} l_\nu(\zeta_{i\nu})[h'_\nu(\lambda T_{im}^{(m)}) - h'_\nu(1)],$$

$$Y_{2\nu} = \frac{m^{5/4}}{n^{3/4}} \sum_{i=0}^{\lambda-1} \lambda T_{im}^{(m)}[l_\nu(G_n(\tilde{U}_m)) - l_\nu(\zeta_{i\nu})][h'_\nu(\lambda T_{im}^{(m)}) - h'_\nu(1)].$$

Furthermore,

$$Y_{1\nu} \sim \frac{1}{\bar{Z}_n} Y_{11\nu} - \frac{\sqrt{n}(\bar{Z}_n - 1)}{\bar{Z}_n^2} Y_{12\nu}$$

where

$$Y_{11\nu} = \frac{m^{5/4}}{n^{3/4}} \sum_{i=0}^{\lambda-1} l_\nu(\zeta_{i\nu})\bar{Z}_{im}^{(m)} h''_\nu(\beta_{i\nu})(\bar{Z}_{im}^{(m)} - 1),$$

$$Y_{12\nu} = \left(\frac{m}{n}\right)^{5/4} \sum_{i=0}^{\lambda-1} l_\nu(\zeta_{i\nu})h''_\nu(\alpha_{i\nu})(\bar{Z}_{im}^{(m)})^2,$$

with $\alpha_{i\nu}$ between $\bar{Z}_{im}^{(m)}/\bar{Z}_n$ and $\bar{Z}_{im}^{(m)}$ and $\beta_{i\nu}$ between $\bar{Z}_{im}^{(m)}$ and 1. From our choice of $\zeta_{i\nu}$ and the independence of $\{\bar{Z}_{im}^{(m)} : i = 0, \ldots, \lambda - 1\}$ we have $E(Y_{11\nu}) = 0$ and $\mathrm{Var}\,(Y_{11\nu}) = \lambda \frac{m^{5/2}}{n^{3/2}} O\left(\frac{1}{m}\right) = \left(\frac{1}{\lambda}\right)^{1/2} O(1) \to 0$, hence $Y_{11\nu} \xrightarrow{P} 0$. Also, it is easy to see $E|Y_{12\nu}| \to 0$. Thus $Y_{1\nu} \xrightarrow{P} 0$. By (6.11) we have

$$|Y_{2\nu}| \leqslant \frac{m^{5/4}}{n^{3/4}} \frac{M}{\lambda} \sum_{i=0}^{\lambda-1} |\lambda T_{im}^{(m)}|\, |h'_\nu(\lambda T_{im}^{(m)}) - h'_\nu(1)|$$

$$\leqslant M^2 \frac{m^{5/4}}{n^{3/4}} \frac{1}{\lambda} \sum_{i=0}^{\lambda-1} |\lambda T_{im}^{(m)}||\lambda T_{im}^{(m)} - 1|$$

hence

$$E|Y_{2\nu}| \leqslant M^2 \frac{m^{5/4}}{n^{3/4}} \frac{1}{\lambda} \sum_{i=0}^{\lambda-1} E\left[\frac{\bar{Z}_{im}^{(m)}}{\bar{Z}_n}\left|\frac{\bar{Z}_{im}^{(m)}}{\bar{Z}_n} - 1\right|\right] = \frac{m^{5/4}}{n^{3/4}} O(m^{-1/2}) \to 0.$$

Finally it follows from (6.13) that $Y_\nu \xrightarrow{P} 0$. □

Lemma 3.3: Under the assumptions of Theorem 3.1,

$$r(\nu) \frac{m}{\sqrt{\lambda}} \sum_{i=0}^{\lambda-1} h_\nu'(\lambda T_{im}^{(m)})\lambda[L_\nu(X_{(im+m)}) - L_\nu(X_{(im)})] \xrightarrow{P} 0. \tag{6.14}$$

Proof: Since $\sum_{i=0}^{\lambda-1} [L_\nu(X_{(im+m)}) - L_\nu(X_{(im)})] = L_\nu(1) - L_\nu(0) = 0$, the LHS of (6.14) equals

$$a_\nu \frac{m^{5/4}}{n^{3/4}} \sum_{i=0}^{\lambda-1} [h_\nu'(\lambda T_{im}^{(m)}) - h_\nu'(1)]\lambda[L_\nu(X_{(im+m)}) - L_\nu(X_{(im)})]$$

$$= a_\nu \frac{m^{5/4}}{n^{3/4}} \sum_{i=0}^{\lambda-1} [h_\nu'(\lambda T_{im}^{(m)}) - h_\nu'(1)] l_\nu(\tilde{X}_{im})\lambda D_{im}^{(m)}$$

where $X_{(im)} \leqslant \tilde{X}_{im} \leqslant X_{(im+m)}$. In view of (6.8) and Lemma 3.2, it suffices to show that

$$\frac{m^{5/4}}{n^{3/4}} \sum_{i=0}^{\lambda-1} [h_\nu'(\lambda T_{im}^{(m)}) - h_\nu'(1)][l_\nu(\tilde{X}_{im}) - l_\nu(\tilde{U}_{(im)})]\lambda T_{im}^{(m)} \xrightarrow{P} 0 \tag{6.15}$$

where $\tilde{U}_{im} = F_\nu(\tilde{X}_{im}) \in [U_{(im)}, U_{(im+m)}]$. But since $\tilde{U}_{im} = F_\nu(\tilde{X}_{im}) = \tilde{X}_{im} + r(\nu)L_\nu(\tilde{X}_{im})$, $|l_\nu(\tilde{X}_{im}) - l_\nu(\tilde{U}_{im})| \leqslant M a_\nu (nm)^{-1/4}$, thus (6.15) holds if

$$\frac{1}{\lambda} \sum_{i=0}^{\lambda-1} [h_\nu'(\lambda T_{im}^{(m)}) - h_\nu'(1)]\lambda T_{im}^{(m)} \xrightarrow{P} 0. \tag{6.16}$$

Because of Lemmas 2.1 and 2.2, we need only to prove (6.16) when $|h_\nu''(\alpha_{i\nu})| \leqslant M$ for $\alpha_{i\nu}$ between $\lambda T_{im}^{(m)}$ and 1. But in that case it follows immediately that

$$\left|\frac{1}{\lambda}\sum_{i=0}^{\lambda-1}[h_\nu'(\lambda T_{im}^{(m)}) - h_\nu'(1)]\lambda T_{im}^{(m)}\right|$$

$$\leqslant M\frac{1}{\lambda}\sum_{i=0}^{\lambda-1}|(\lambda T_{im}^{(m)} - 1)\lambda T_{im}^{(m)}| \sim \frac{M}{\bar{Z}_n^2}\frac{1}{\lambda}\sum_{i=0}^{\lambda-1}|\bar{Z}_{im}^{(m)} - \bar{Z}_n| \xrightarrow{P} 0. \qquad \square$$

Proof of Theorem 3.1: From (3.1)

$$\lambda D_{im}^{(m)} - \lambda T_{im}^{(m)} = -r(\nu)\lambda[L_\nu(X_{(im+m)}) - L_\nu(X_{(im)})] = -r(\nu)l_\nu(\tilde{X}_{im})\lambda D_{im}^{(m)}.$$

Hence

$$\frac{m}{\sqrt{\lambda}}\sum_{i=0}^{\lambda-1}|h_\nu(\lambda D_{im}^{(m)}) - h_\nu(\lambda T_{im}^{(m)})|$$

$$= \frac{m}{\sqrt{\lambda}}\sum_{i=0}^{\lambda-1}\{-h_\nu'(\lambda T_{im}^{(m)})\lambda[L_\nu(X_{(im+m)}) - L_\nu(X_{(im)})]r(\nu)$$

$$+\frac{1}{2}h_\nu''(\lambda T_{im}^{(m)})r^2(\nu)l_\nu^2(\tilde{X}_{im})(\lambda D_{im}^{(m)})^2 + o_p(r^2(\nu))\}$$

$$= -a_\nu\frac{m^{5/4}}{n^{3/4}}\sum_{i=0}^{\lambda-1}h_\nu'(\lambda T_{im}^{(m)})\lambda[L_\nu(X_{(im+m)}) - L_\nu(X_{(im)})]$$

$$+\frac{1}{2}a_\nu^2\frac{1}{\lambda}\sum_{i=0}^{\lambda-1}l_\nu^2(\tilde{X}_{im})(\lambda D_{im}^{(m)})^2h_\nu''(\lambda T_{im}^{(m)}) + o_p(1) \xrightarrow{P} \frac{1}{2}a^2b\|l\|_2^2$$

by Lemma 3.2 and Lemma 3.3. $\square$

References

Beirlant J, van Zuijlen MCA (1985) The empirical distribution function and strong laws for functions of order statistics of uniform spacings. J Multivariate Analysis 16:300–317

Billingsley P (1968) Convergence of probability measures. John Wiley, NY

Cressie N (1976) On the logarithm of high-order spacings. Biometrika 63:343–355

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

Dudewicz EJ, van der Meulen EC (1981) Entropy-based tests of uniformity. J Amer Statist Assoc 76:967–974

Hall P (1986) On powerful distributional tests based on sample spacings. J multivariate Analysis 19:201–224

Holst L (1972) Asymptotic normality and efficiencies for certain goodness-of-fit tests. Biometrika 59:137–145

Jammalamadaka SR, Tiwari RC (1986) Efficiencies of some disjoint spacings tests relative to a χ^2 Test. In: Puri ML, Vilaplana J, Wertz W (eds) New perspectives in theoretical and applied statistics. John Wiley, New York, pp 311–318

Kuo M, Rao JS (1981) Limit theory and efficiencies for tests based on higher order spacings. In: Statistics – Applications and new Directions, Proceedings of the Golden Jubilee Conference of the India Statistical Institute. Statistical Publishing Society, Calcutta, pp 333–352

Pyke R (1965) Spacings. JR Statist Soc B 27:395–449

Quine MP, Robinson J (1985) Efficiencies of Chi-square and likelihood ratio goodness-of-fit tests. Ann of Statist 13:727–742

Weiss L (1974) The asymptotic sufficiency of a relatively small number of order statistics in tests of fit. Ann of Statist 2:795–802

Received 28. 9. 1988
Revised version 1. 2. 1989