Title
A test of goodness-of-fit based on Gini's index of spacings

Journal
Statistics & Probability Letters, 68(2)

ISSN
0167-7152

Authors
Jammalamadaka, S Rao
Goria, MN

Publication Date
2004-06-01

DOI
10.1016/j.spl.2004.02.009

Peer reviewed

Available online at www.sciencedirect.com
SCIENCE DIRECT®

Statistics & Probability Letters 68 (2004) 177–187

STATISTICS & PROBABILITY LETTERS

www.elsevier.com/locate/stapro

A test of goodness-of-fit based on Gini's index of spacings

S. Rao Jammalamadaka[a,*], M.N. Goria[b]

[a]Department of Statistics and Probability, University of California, Santa Barbara, CA 93106, USA
[b]Trento University, Trento, Italy

Received October 2003; accepted February 2004

Abstract

This paper introduces a new test for goodness-of-fit based on the Gini index which is the sum over all pairs, of the absolute differences of the observed spacings. We derive its exact and asymptotic distributions under the null hypothesis, after showing that it is distributionally equivalent to the sum of uniform observations on the unit interval. After a discussion of local powers of this and related tests, we provide simulated power comparisons, which demonstrate that the Gini test is better than all the other competitors considered, against a wide variety of alternatives.

Keywords: Goodness-of-fit tests; Spacings; Gini index; Order statistics; Directional data

1. Introduction

Consider a random sample $X_1,\ldots,X_n$ from a continuous distribution function(d.f.) F^* on the real line and let $(X_{(1)},\ldots,X_{(n)})$ denote the order statistics from this sample. Considerable attention has been devoted in the literature to the problem of goodness-of-fit i.e., testing the hypothesis

$$\mathrm{H} : F^*(x) = F_0^*(x) \quad \text{against the alternative} \quad K : F^*(x) \neq F_0^*(x),$$

where F_0^* is a completely specified continuous d.f. Apart from various ad hoc tests, there are

* Corresponding author. Tel.: +1-805-8933119; fax: +1-805-8932334.
E-mail address: rao@pstat.ucsb.edu (S.R. Jammalamadaka).

0167-7152/$ - see front matter
doi:10.1016/j.spl.2004.02.009

3 general classes of tests for this problem, viz.

1. the χ^2 tests which depend on groupings,
2. tests based on empirical d.f.'s, and
3. tests based on spacings i.e. the differences between order statistics.

There have been suggestions in the literature that tests based on empirical d.f.'s such as the Kolmogorov–Smirnov and Cramer–von Mises statistics perform better in detecting differences between d.f.'s, whereas tests based on spacings are particularly effective in revealing the differences between densities. See for instance, the discussion in Pyke (1965), which also provides an excellent survey on the topic of spacings. Jammalamadaka and Tiwari (1987) show that comparable spacings tests are better than chi-squared tests, in terms of local power. It should be noted that, outside of the goodness-of-fit framework, spacings-based tests arise in many other contexts including reliability. See e.g. Soofi et al. (1995), and Jammalamadaka and Taufer (2003). In a survey paper Birnbaum (1953) discusses general methods of chi-square, Kolmogorov–Smirnov, Cramer–von Mises and spacings. The reader is also referred to Shorack and Wellner (1986) for a thorough treatment of distribution-free goodness-of-fit tests based on empirical d.f.'s.

Without loss of any generality, one can reduce the above problem of goodness-of-fit, to testing the hypothesis of uniformity on the unit interval, by means of the probability integral transformation $U = F^*(X)$. That is, on the basis of the transformed sample, $U_i = F_0^*(X_i)$, $i = 1, 2, \ldots, n$ the problem becomes one of testing uniformity i.e.,

$$\mathrm{H} : F(u) = u,\ 0 < u < 1 \quad \text{against the alternative} \quad K : F(u) \neq u,\ 0 < u < 1.$$

From here on, we shall assume that such a transformation has been made and define

$$D_i = U_{(i)} - U_{(i-1)},\ i = 1, 2, \ldots, n+1, \text{ with the notation } U_{(0)} \equiv 0,\ U_{(n+1)} \equiv 1.$$

Although there has been significant work prior to it, it is fair to say that the impetus for work on the theory of spacings came from the paper by Greenwood (1946), who proposed the statistic $\sum_{i=1}^{n+1} D_i^2$ to test whether certain events such as the spread of disease occur at random (or follow a Poisson process with fixed rate) on the time axis. Since then, many tests based on spacings have been proposed in literature; see Pyke (1965) for a good review. Sethuraman and Rao (1970) and Rao and Sethuraman (1975) provide a unified treatment of the asymptotic distribution theory for symmetric spacings statistics, by expressing such statistics in terms of the empirical spacings process.

The $\{D_i\}$ defined above are called one-step or simple spacings. One can generalize this idea of spacings by defining, for any fixed integer m, $(1 \leqslant m < n)$, the m-step (or higher-order) spacings. The *overlapping* m-step spacings are defined by

$$D_i^{(m)} = \begin{cases} U_{(i+m)} - U_{(i)}, & i = 0, 1, \ldots, n+1-m, \\ 1 + U_{(i+m-n-1)} - U_{(i)}, & i = n+2-m, \ldots, n, \end{cases}$$

whereas the *non-overlapping* m-step spacings are defined by

$$\mathbf{D}_i^{(m)} = U_{(i+1)m} - U_{(i)m}, \quad i = 0, 1, \ldots, [n/m] - 1.$$

Jammalamadaka and Kuo (1984) consider tests based on these higher-order spacings and show that tests based on m-step spacings have higher efficiency compared to their counterparts based on simple or one-step spacings.

The spacings $\{D_i,\ i=1,2,\ldots,n+1\}$, under the null hypothesis of uniformity, form an exchangeable set of random variables with an expected value of $1/(n+1)$. Since they add up to one, their arithmetic mean is also $\bar{D}=1/(n+1)$. Thus, tests of uniformity can be constructed by measuring how different $\{D_i,\ i=1,2,\ldots,n+1\}$ are from their average value of $1/(n+1)$. In Section 2, we show that many of the spacings tests considered in the literature can be viewed in terms of various choices of measures of dispersion of these $\{D_i\}$.

An alternate way to construct a general class of such tests is to use the Csiszár divergence measure between two probability distributions $F_1(\cdot)$ and $F_2(\cdot)$

$$I_h(F_1,\ F_2)=\int_{\Re} h\left(\frac{\mathrm{d}F_1(x)}{\mathrm{d}F_2(x)}\right)\,\mathrm{d}F_2(x),$$

where $h:(0,\ \infty)\to\Re$ is a convex function with $h(1)=0$. See e.g., Ekström (1997) and Ghosh and Jammalamadaka (2001). A special case of this measure is the Kullback–Liebler distance, which takes the simple form $\sum_{i=1}^{n+1}\log D_i$ in this particular instance. Darling (1953) has studied this statistic and its first order approximation $\sum_{i=1}^{n+1} 1/D_i$.

In Section 2, We propose a new test for goodness-of-fit based on the Gini's measure of dispersion and in Section 3, we discuss its exact and asymptotic distributions under the null hypothesis. Some comments on the local powers of this and related statistics are provided in Section 4 and numerical comparison of powers is provided in Section 5. Finally, in Section 6, we provide conclusions and discuss possible future research.

2. Spacings tests as measures of dispersion

As stated earlier, a test of uniformity based on spacings corresponds to a dispersion measure for D_i and one can use any dispersion measure to serve that purpose.

Common among the dispersion measures are the variance, mean absolute deviation, range and interquartile distance. While measures such as the mean absolute deviation and Gini's index are less frequently used, perhaps due to the computational difficulties involved, these are known to be more robust to the outliers than the others mentioned.

We remark that the test based on the variance

$$V_n=\sum_{i=1}^{n+1}(D_i-1/(n+1))^2 \tag{1}$$

corresponds to the Greenwood statistic (see also Kimball (1947)). A generalized Greenwood statistic based on m-step spacings (overlapping, as well as non-overlapping) and its asymptotic optimality, has been studied in detail by Rao and Kuo (1984).

An alternate dispersion measure, namely the mean absolute deviation leads to the statistic

$$R_n=\sum_{i=1}^{n+1}|D_i-1/(n+1)| \tag{2}$$

which was suggested by Sherman (1950). This latter statistic was independently introduced in the context of circular data by Rao (1969) with Russell and Levitin (1995) providing an extensive table of critical values for it. More generally, it is clear that any convex function of the spacings or of higher-order spacings can be used for this purpose.

The range, the interquartile range and the Gini's index of the observed spacings can also be used for goodness-of-fit testing. This has not been discussed in literature, although Rao and Sobel (1980) provide a pertinent and useful discussion of tests based on ordered spacings. Our goal here is to study the test statistic based on the Gini's index, while the others are deferred to a later study.

Gini, in introducing his measure of dispersion, raised objection to using the variance or the mean absolute deviation, since they measure the deviation of individual observations from the "center," thus interlinking the concept of location with variability. According to him, these are two distinct properties, needing distinct measures which do not depend on each other. Therefore he proposed the sum of pairwise distances between the observations, as a measure of dispersion. Gini measure can be related to the measure of concentration and has thus found considerable appeal in economics and received renewed attention.

Applying this idea to $\{D_i\}$ in our context, we may define the Gini index of dispersion for spacings, as

$$G_n = \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} |D_i - D_j|. \tag{3}$$

Our goal in this paper is to study this Gini statistic G_n, its distributional properties and how it compares with other competing tests of uniformity.

Remark 1. One can consider a generalized Gini statistic defined by

$$G_n(r) = \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} |D_i - D_j|^r, \quad r > 0.$$

Clearly the special case $r = 1$ corresponds to the Gini statistic G_n that we discuss here, while the special case $r = 2$ corresponds to the Greenwood statistic.

Remark 2. If we consider the vector of spacings, $\mathbf{D}=(D_1, \ldots, D_{n+1})$, the Greenwood statistic $V=\mathbf{D}'\mathbf{D}$ is a quadratic form which gives equal weights to all the spacings. Since $\{D_i\}$ are correlated, it is tempting to ask whether one gets different (and possibly better) test by taking into account the covariance-structure between these spacings. It is known (see e.g. Pyke (1965)) that

$$\mathrm{Cov}(\mathbf{D}) = \Sigma = (\sigma_{ij}), \quad i, j = 1, 2, \ldots, n+1,$$

where

$$\sigma_{ij} = -\frac{1}{(n+1)^2(n+2)}, \quad \sigma_{ii} = \frac{n}{(n+1)^2(n+2)}.$$

Since the covariance matrix is singular, we may take its generalized inverse Σ^-. Consider the statistic

$$V^* = \mathbf{D}'\Sigma^-\mathbf{D}$$

which introduces appropriate weights into the quadratic form. It can be checked that the matrix

$$\Sigma^- = ((\sigma^{ij})),$$

with

$$\sigma^{ij} = (n+1)(n+2), \quad \sigma^{ii} = 2(n+1)(n+2)$$

is indeed the Penrose inverse of Σ. The resulting statistic V^* is exactly identical to the unweighted Greenwood statistic, V. Thus there is no gain in such a weighting scheme.

Our objective in the next few sections is to discuss the test statistic G_n based on Gini's index and compare its performance with competitors like the Greenwood statistic.

3. Exact and asymptotic distribution of G_n under the hypothesis

In this section, we derive the exact as well as the asymptotic distribution of G_n under the null hypothesis of uniformity. The Gini's measure of dispersion, G_n can be rewritten in terms ordered spacings as follows:

$$\begin{aligned} G_n &= \sum_{i=1}^{n+1}\sum_{j=1}^{n+1} |D_{(i)} - D_{(j)}| \\ &= 2\sum_{i=1}^{n+1}\sum_{j>i}^{n+1} (D_{(j)} - D_{(i)}) \\ &= 2\sum_{i=1}^{n+1}\left[\left(\sum_{j=i+1}^{n+1} D_{(j)}\right) - (n-i+1)D_{(i)}\right]. \end{aligned} \tag{4}$$

On writing the above explicitly and simplifying, we obtain

$$G_n = 2\sum_{i=1}^{n+1}(2i-n-2)D_{(i)} = 4\sum_{i=1}^{n+1} iD_{(i)} - 2(n+2) = 4L_n - 2(n+2), \tag{5}$$

where

$$L_n = \sum_{i=1}^{n+1} iD_{(i)}. \tag{6}$$

Now define

$$E_i = (n-i+2)[D_{(i)} - D_{(i-1)}], \quad i = 1, 2, \ldots, n+1,$$

or, conversely

$$D_{(i)} = \sum_{j=1}^{i} E_j/(n-j+2).$$

Then L_n defined in Eq. (6) can be expressed as

$$L_n = \sum_{i=1}^{n+1} i \sum_{j=1}^{i} E_j/(n-j+2).$$

On explicitly writing the above double sum, we get

$$L_n = \sum_{i=1}^{n+1} \left[\frac{(n+1)(n+2)}{2} - \frac{(i-1)i}{2} \right] E_i/(n-i+2).$$

On further simplification and using the fact that $\sum_{i=1}^{n+1} E_i = 1$, we get

$$L_n = 1/2 \sum_{i=1}^{n+1} (n+i+1)E_i = (n+1)/2 + 1/2 \sum_{i=1}^{n+1} iE_i.$$

Starting with the joint density of $U_{(i)}$, Durbin (1961) shows by repeated transformation that $(E_1, \ldots, E_{n+1})$ are distributionally equivalent to $(D_1, \ldots, D_{n+1})$ *under the null hypothesis of uniformity*. This result can also be checked directly by representing uniform spacings in terms of scaled exponential random variables and using a very similar property that holds for exponential order statistics. Thus, under the null hypothesis, the distribution of L_n is the same as that of

$$L_n \sim (n+1)/2 + 1/2 \sum_{i=1}^{n+1} iD_i = (n+1)/2 + L_n^*/2, \tag{7}$$

where $L_n^* = \sum_{i=1}^{n+1} iD_i$ is related to the sum $S_n = \sum_{i=1}^{n} U_i$ of n uniform random variables on the unit interval. Indeed

$$\begin{aligned} S_n &= \sum_{i=1}^{n} U_i = \sum_{i=1}^{n} U_{(i)} \\ &= \sum_{i=1}^{n} \sum_{j=1}^{i} D_j = \sum_{i=1}^{n+1} (n-i+1)D_i \\ &= (n+1) - \sum_{i=1}^{n+1} iD_i = (n+1) - L_n^*. \end{aligned} \tag{8}$$

Combining Eqs. (5), (7) and (8), we get

$$G_n \sim 2(n - S_n) \tag{9}$$

under the null hypothesis. Note that this is a distributional equivalence that holds under the hypothesis of uniformity and should not be interpreted any more broadly. Again, a uniform observation U_i has the same distribution as that of $(1 - U_i)$, it clear that the distribution of $n - S_n$ is the same as that of S_n which has the following well-known density (see e.g. Wilks (1962, p. 204))

$$f_n(x) = \frac{1}{(n-1)!} \sum_{k=0}^{\infty} (-1)^k \binom{n}{k} \langle x - n \rangle^{n-1}, \ \ 0 < x < n.$$

From this, one can write the exact null density of the Gini statistic G_n, in the range $0 < y < 2n$, as

$$g_n(y) = \frac{1}{(n-1)!2} \sum_{k=0}^{\infty} (-1)^k \binom{n}{k} \left\langle \frac{y}{2} - n \right\rangle^{n-1}.$$

As for the large sample distribution under the null hypothesis, note that by the Central Limit Theorem for iid variables,

$$S_n \simeq N(n/2, n/12)$$

under the null hypothesis of uniformity. Therefore by using the relationship (9) of G_n to S_n, we can claim the asymptotic normality

$$G_n \simeq N(n, n/3)$$

under the null. More precisely, as n goes to infinity,

$$(3/n)^{1/2}(G_n - n)$$

converges to a standard Normal distribution.

Note that the exact as well as the asymptotic distributions of both L_n^* and L_n can be easily written down from that of S_n using Eqs. (8) and (7), respectively.

Since $E(D_{(i)}) \approx i/(n+1)^2$, it is reasonable to consider yet another test statistic for goodness-of-fit, viz.

$$T_n = \sum_{i=1}^{n+1} [D_{(i)} - i/(n+1)^2]^2. \tag{10}$$

When expanded, this can be written as a linear combination of the Gini Statistic G_n (or its equivalent L_n) and V_n, which is equivalent to the Greenwood statistic i.e.,

$$\begin{aligned} T_n &= V_n + 1/(n+1) - 2L_n/(n+1)^2 + (n+2)(2n+3)/\{6(n+1)^3\} \\ &= V_n - 2L_n/(n+1)^2 + (8n^2 + 19n + 12)/\{6(n+1)^3\} \end{aligned} \tag{11}$$

or

$$(n+1)^{3/2}T_n - 5/6(n+1)^{1/2} = [(n+1)^{3/2}V_n - 2(n+1)^{1/2}] - 2(n+1)^{-1/2}[L_n - 3/4(n+1)].$$

The exact null distribution of Greenwood statistic has been tabulated (see e.g. Burrows (1979)) and for large samples,

$$(n+1)^{3/2}V_n - 2(n+1)^{1/2}$$

is asymptotically normal with zero mean and variance 4 (see e.g. Darling (1953)). However since G_n and V_n are not necessarily independent, the exact and asymptotic null distribution of T_n needs further investigation.

4. Some comments on power for specific alternatives

Here we offer some comments on the powers of the tests S_n, L_n and L_n^* under various alternatives. It is well known (see e.g.. Section 3 of Rao and Sethuraman (1975) and references therein) that tests which are symmetric in spacings, have lower asymptotic powers in the Pitman sense. However, their small sample powers are reasonable (see the small simulation in the last section) and they are especially relevant in the field of directional data, where spacings form the maximal invariant. We should also remark that although the Gini test is equivalent to S_n under the null hypothesis, its behavior and performance can be very different under alternatives. We consider a close sequence of the so-called "smooth" alternatives of the form

$$H_n : F_n(x) = x + n^{-1/2}M_n(x), \quad M_n(0) = 0 \quad \text{and} \quad M_n(1) = 0. \tag{12}$$

Here we assume that $M_n(x)$ is continuously differentiable with derivative $m_n(x)$ and that it converges uniformly to $M(x)$ on $[0, 1]$, a twice continuously differentiable function with first and second derivatives $m(x)$, $m'(x)$, respectively, so that

$$\sup_{0 \leqslant x \leqslant 1} |m_n(x) - m(x)| = o(1).$$

Against such a sequence of alternatives, Holst and Jammalamadaka (1981) show that (see their Theorem 3.1 and Eq. (3.6)), the test based on

$$T_n = \sum_{i=1}^{n+1} m(i/(n+1))D_i.$$

is locally most powerful (LMP)for testing uniformity. The following remark is also taken from there:

Remark 3. Consider testing the hypothesis that a random sample is from a logistic distribution with $G(t) = 1/(1 + e^{-t})$ against local translation alternatives, $G(t - \theta/\sqrt{n})$. After transforming the data to the interval $[0, 1]$ through $x = G(t)$, the null hypothesis translates to testing uniformity, while $M_n(x)$ described in Eq. (12) converges to $M(x) = G'(G^{-1}(x))$ with derivative

$$m(x) = (2x - 1), \qquad 0 \leqslant x \leqslant 1$$

(cf. Holst and Jammalamadaka (1981)). Thus the LMP test for this context reduces to one based on L_n^*.

Things are not so clear when one is dealing with tests based on ordered spacings like L_n or its equivalent G_n. This is mainly because of the fact that although spacings under general alternatives can be expressed in terms of uniform spacings for close alternatives (see for example Eq. (3.8) of Rao and Sethuraman (1975)), their ordering depends very much on what the alternative density is. Hence one has to resort to simulations.

Remark 4. When considering alternatives at a distance of $n^{-1/4}$ (i.e., Eq. (12) with $n^{-1/2}$ replaced by $n^{-1/4}$), it is known that the Greenwood test V_n is locally optimal among tests based symmetrically on the spacings. However, it has no power against alternatives at a distance of $n^{-1/2}$ i.e., its power is the same as the level and is not a competitor for L_n or L_n^* (see Sethuraman and Rao (1969)).

5. Simulated power comparisons of tests

In this section, we report on a small simulation study that was carried out to compare the Monte Carlo powers of the four tests of goodness-of-fit, namely the test based on unordered spacings L_n^*, the Gini test based on ordered spacings G_n, the Greenwood statistic V_n, and the hybrid statistic

T_n mentioned in Eq. (11). For this purpose we consider testing the null hypothesis of a standard Normal distribution against several alternatives, at 0.05 level. These alternatives include heavy-tailed distributions such as the Student $t(v)$, $v = 1, 2, 3$, thin long-tailed, Laplace, and short-tailed, Logistic. Samples of sizes 10, 20, 50, 100 with replications of 1000 were generated using the MATLAB package. The resulting simulated powers are reported in the table below.

Simulated powers of the tests at various alternatives. $\alpha = 0.05$.

n	*Test*	t_1	t_2	t_3	*Logistic*	*Laplace*
10	L_n^*	0.093	0.078	0.065	0.089	0.068
	G_n	0.462	0.273	0.172	0.398	0.196
	V_n	0.291	0.190	0.122	0.288	0.127
	T_n	0.260	0.175	0.107	0.269	0.112
20	L_n^*	0.074	0.064	0.058	0.096	0.071
	G_n	0.610	0.300	0.159	0.519	0.209
	V_n	0.356	0.184	0.112	0.374	0.145
	T_n	0.321	0.165	0.099	0.345	0.136
50	L_n^*	0.077	0.079	0.065	0.082	0.052
	G_n	0.890	0.502	0.281	0.762	0.262
	V_n	0.624	0.288	0.164	0.550	0.178
	T_n	0.589	0.275	0.163	0.537	0.175
100	L_n^*	0.097	0.071	0.064	0.092	0.073
	G_n	0.987	0.645	0.364	0.928	0.408
	V_n	0.837	0.383	0.202	0.741	0.260
	T_n	0.782	0.318	0.183	0.689	0.240

From an examination of this table, the following statements can be asserted:

1. It is clear that the Gini test G_n dominates all the other competing tests, against each of the alternatives considered. T_n and V_n have comparable powers and this is to be expected in view of the high weight given to V_n in Eq. (11). Put differently, T_n ignores the better statistic G_n in its weighting, at its own peril!
2. It may be observed that for the Student's $t(v)$, $v = 1, 2, 3$ alternatives, the power decreases with increasing degrees of freedom. Again this is to be expected since the larger degrees of freedom represent closeness to normality (although with d.f. this small, we are quite far from normality in an absolute sense).
3. The power performance of the new test we introduced, the Gini test G_n is quite decent even for moderate sample sizes like 20 and is nearly twice that of the other competitors considered. We have done more extensive simulations with several other alternatives with similar ordering in relative performance.

6. Conclusion

This paper introduces a new test for goodness-of-fit based on the Gini index of spacings. Its exact and asymptotic distributions are provided under the null hypothesis and a simulation study shows that this test has higher powers than others at certain common alternatives. There are several interesting open questions that are worth further investigation. For instance, one could study the range and the inter-quartile range of spacings, the more general statistic $G_n(r)$ mentioned in Remark 1, or the statistic T_n defined in Eq. (10). One could also investigate each of these statistics using non-overlapping or overlapping m-step spacings. Finally, it would be of interest to theoretically derive the Pitman efficiency of such tests under smooth alternatives, as in Sethuraman and Rao (1969).

Acknowledgements

The authors wish to thank a referee whose meticulous reading and numerous constructive suggestions, considerably improved the presentation of this paper.

References

Birnbaum, Z.W., 1953. Distribution free tests for a continuous probability functions. Ann. Math. Statist. 24, 1–8.
Burrows, P.M., 1979. Selected percentage points of Greenwood's Statistic. J. R. Statist. Soc. Ser. A 142, 256–258.
Darling, D.A., 1953. On a class of problems related to the random division of an interval. Ann. Math. Statist. 24, 239–253.
Durbin, J., 1961. Some methods of constructing exact tests. Biometrika 48, 41–55.
Ekström, M., 1997. Maximum spacings methods and limit theorems for statistics based on spacings. Ph.D. Thesis, Umea University, Sweden.
Ghosh, K., Jammalamadaka, S.Rao., 2001. A general estimation method using spacings. J. Statist. Planning Inference 93, 71–82.
Greenwood, M., 1946. The statistical study of infectious diseases. J. R. Statist. Soc. Ser. A 109, 85–110.
Holst, L., Jammalamadaka, S.Rao., 1981. Asymptotic spacings theory with applications to the two-sample problem. Can. J. Statist. 9, 79–89.
Jammalamadaka, S.Rao., Taufer, E., 2003. Testing exponentiality by comparing the empirical distribution function of the normalized spacings with that of the original data. J. Nonparametric Statist. 15, 719–729.
Jammalamadaka, S.Rao., Tiwari, R.C., 1987. Efficiencies of some disjoint spacings tests relative to a chi-square test. In: Puri, M.L., Vilaplana, J., Wertz, W. (Eds.), New Perspectives in Theoretical and Applied Statistics. Wiley, New York, pp. 311–317.
Kimball, B.F., 1947. Some basic theorems for developing tests of fit for the case of the non-parametric probability distribution functions. 18, 540–548.
Pyke, R., 1965. Spacings (with discussion). J. R. Statist. Soc. 7, 395–449.
Rao, J.S., 1969. Some contributions to the analysis of circular data. Ph.D. Thesis, Indian Statistical Institute, Calcutta.
Rao, J.S., Kuo, M., 1984. Asymptotic results on the Greenwood statistic and some of its generalizations. J. R. Statist. Soc. Ser. B 46, 228–237.
Rao, J.S., Sethuraman, J., 1975. Weak convergence of empirical distribution functions of random variables subject to perturbations and scale factors. Ann. Statist. 3, 299–313.
Rao, J.S., Sobel, M., 1980. Incomplete Dirichlet integrals with applications to ordered uniform spacings. J. Multivariate Anal. 10, 603–610.
Russell, G.S., Levitin, D.J., 1995. An expanded table of probability values for Rao's spacing test. Commun. Statist.: Simulation Comput. 24 (4), 879–888.

Sethuraman, J., Rao, J.S., 1969. Pitman Efficiencies of tests based on spacings. In: Puri, M.L. (Ed.), Nonparametric Techniques in Statistical Inference. Cambridge University Press, Cambridge, pp. 405–415.
Sherman, B., 1950. A random variable related to the spacing of sample values. Ann. Math. Statist. 21, 339–361.
Shorack, G.R., Wellner, J.A., 1986. Empirical Processes, with Applications to Statistics. John Wiley, NY.
Soofi, E.S, Ebrahimi, N., Habibullah, M., 1995. Information distinguishability with application to Analysis of failure data. J. Amer. Statist. Assoc. 90, 657–668.
Wilks, S.S., 1962. Mathematical Statistics. Wiley, New York.