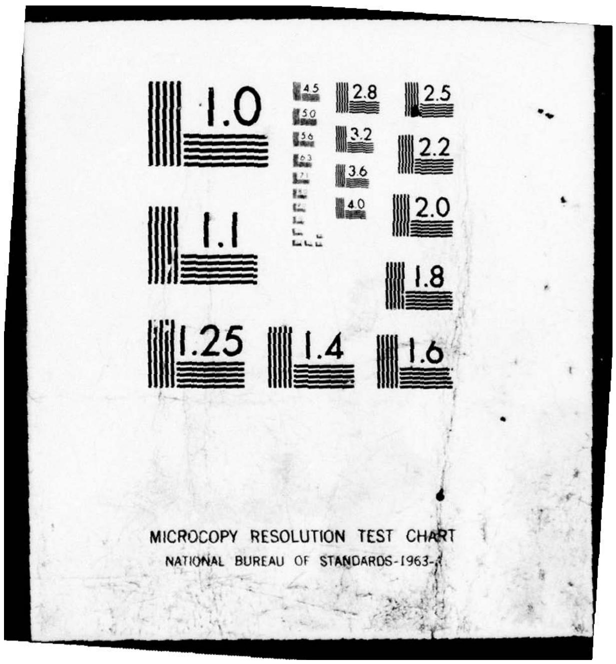
AD-A065 519 FLORIDA STATE UNIV TALLAHASSEE DEPT OF STATISTICS F/G 12/1
ASYMPTOTIC NORMALITY AND EFFICIENCIES OF TESTS BASED ON MODIFIE--ETC(U)
JAN 79 J S RAO, J SETHURAMAN DAAG29-76-G-0238
UNCLASSIFIED FSU-STATISTICS-M489 ARO-13888.15-MX NL

1 OF 1
AD
A065519

END
DATE
FILMED
5--79
DDC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

he Florida State University

Department

D

(18) ARO

(19) 13888.15-MX

LEVEL II (12)

AD A0 65519

(6)

ASYMPTOTIC NORMALITY AND EFFICIENCIES OF TESTS
BASED ON MODIFIED SPACINGS.

by (9) Technical rept.,

(10) J. S. Rao, University of California, Santa Barbara

and

J. Sethuraman, Florida State University, Tallahassee

FSU Technical Report No. M489
USARO-D Technical Report No. 30

(11) Jan[redacted]979 (12) 31p.

The Florida State University
Department of Statistics
Tallahassee, Florida 32306

(14) FSU-STATISTICS-M489, TR-30

DDC FILE COPY

[1]Research supported by the United States Army Research Office under Grant No. (15) DAAG29-76-G0238. The United States Government is authorized to reproduce and distribute reprints for governmental purposes.

NSF-MCS78-01881

THE VIEW, OPINIONS, AND/OR FINDINGS CONTAINED IN THIS REPORT ARE THOSE OF THE AUTHOR(S) AND SHOULD NOT BE CONSTRUED AS AN OFFICIAL DEPARTMENT OF THE ARMY POSITION, POLICY, OR DECISION, UNLESS SO DESIGNATED BY OTHER DOCUMENTATION.

ACCESSION for
NTIS White Section
DDC Buff Section
UNANNOUNCED
JUSTIFICATION
BY
DISTRIBUTION/AVAILABILITY CODES
Dist. AVAIL. and/or SPECIAL
A

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

DDC
RECEIVED
MAR 9 1979
RECEIVED
D

79 02 27 044

400 277

ASYMPTOTIC NORMALITY AND EFFICIENCIES OF TESTS BASED ON MODIFIED SPACINGS

J. S. Rao[1]
University of California, Santa Barbara
and
J. Sethuraman[2]
Florida State University, Tallahassee

ABSTRACT:

Tests based on adjusted or modified spacings are proposed for testing goodness of fit problems. The weak convergence of the empirical distribution function of such modified spacings is studied using some earlier results of the authors. The asymptotic theory under close alternative sequences is also given thus enabling one to calculate the asymptotic relative efficiencies of such tests.

[1] Research supported in part by the National Science Foundation under Grant No. MCS78-01881.

[2] Research supported by the United States Army Research Office, Durham, under Grant No. DAAG29-76-G0238.

AMS 1970 subject classification numbers: Primary 62E20, 62G20, 62G10; Secondary 60F05. Key words and phrases: Modified spacings, empirical distribution function, weak convergence, efficiencies of tests.

ASYMPTOTIC NORMALITY AND EFFICIENCIES OF TESTS BASED ON MODIFIED SPACINGS

J.S. Rao, University of California, Santa Barbara

and

J. Sethuraman, Florida State University, Tallahassee.

1. Introduction and Summary.

Let $X_1, X_2, \ldots, X_{n-1}$ be $(n-1)$ independently and identically distributed random variables with a common distribution function (d.f.) $G(x)$. The goodness-of-fit problem is to test whether $G(x)$ is a specified d.f. . When the latter d.f. is continuous, a simple probability transformation on the random variables would permit us to equate the pre-assigned d.f. to the uniform d.f. on $[0,1]$. Therefore, from now on we assume that this reduction has been effected and that under the hypothesis, $G(x)$ is the uniform d.f. on $[0,1]$.

Let $X_1' \le X_2' \le \ldots \le X_{n-1}'$ be the order statistics. The sample spacings $(D_1, \ldots, D_n)$ are defined by

$$D_i = X_i' - X_{i-1}', \quad i = 1, \ldots, n$$

where we put $X_o' = 0$, $X_n' = 1$. In order that this definition of the sample spacings be meaningful under any alternative, the d.f. $G(x)$ must have the carrier $[0,1]$. (The carrier of a d.f. is the smallest closed set with probability one.)

Under the null hypothesis, $E(D_i) = 1/n$ for all i. We will therefore call $\{nD_i,\ i = 1, \ldots, n\}$ as 'normalised' spacings. Further if $h_{n1}, \ldots, h_{nn}$ be some positive numbers, then we shall call $\{nD_i/h_{ni},\ i = 1, \ldots, n\}$ as 'modified' or 'adjusted' spacings. For example, one way of adjusting the spacings is to divide them by their expectations under some arbitrary d.f. . The rationale behind dealing with the modified spacings is that in some cases one may choose to adjust the spacings by their expectations under some alternative distribution to increase the efficiency or otherwise, to enlarge the class of statistics based on spacings. This may be thought of as being analogous to the use of Normal scores and other scores in rank tests (see Section 6 for details). Tests based on spacings have been proposed for the goodness-of-fit problem by several authors. See e.g., Greenwood (1946), Kimball (1950), Sherman (1950) and Darling (1953). Distribution theory of such statistics and their asymptotic efficiencies have been studied by Sethuraman and Rao (1970) and Rao and Sethuraman (1975). Since these two papers are closely related to the present work and are referred often, we will refer to them as SR and RS respectively.

The present paper is devoted to the study of asymptotic distributions and asymptotic relative efficiencies (ARE's) of tests based on modified spacings. As pointed out in SR, for calculating Pitman efficiencies, it is enough to obtain the limiting distribution under a sequence of alternatives which converge to the hypothesis. This problem turns out to be somewhat simpler as can be seen from the cases treated by Cibisov (1961) and Weiss (1965). Se also RS and SR. We, therefore, choose the following sequence of alternatives

$$G(x) = x + L(x)/n^{\partial}, \quad 0 \leq x < 1$$

where ∂ is a number $\geq 1/4$ and $L(x)$ is twice continuously differentiable on $[0,1]$. These conditions imply our earlier requirement that the carrier of $G(x)$ be $[0,1]$. We shall say that this alternative is at a distance of order $n^{-\partial}$ from the hypothesis.

As in RS we obtain the limiting distributions of the empirical distribution functions of the adjusted or modified spacings in the sense of weak convergence of measures in $D[0,\infty]$ the space of functions on $[0,\infty]$ with no discontinuities of the second kind. This is done through some interesting results of independent interest concerning the empirical distribution functions of perturbed random variables and randomly scaled random variables which are given in Sections 3 and 4. Appealing to the invariance principle, we immediately have the limiting distributions of a large class of statistics which are symmetric in the modified spacings.

As shown in RS and SR, tests which are symmetric in the normalised spacings have limiting power greater than the test size only if $\partial = 1/4$ i.e., they cannot discriminate alternatives which are at a distance of order $n^{-\partial}$ from the hypothesis, for any $\partial > 1/4$. It is also shown there that among the many such standard tests due to Greenwood (1946), Kimball (1950), Sherman (1950) and Darling (1953), the one by Greenwood based on $\sum_{i=1}^{n} (nD_i)^2/n$ has the maximum ARE. We note another interesting feature of the symmetric spacings tests, namely their ARE's do not depend on the particular choice of the alternative sequence i.e., $L(x)$.

These features are not shared by tests symmetric in the modified spacings, which are discussed in Section 6. Here the efficiencies depend on the alternative sequence as may be seen from the expressions (6.5) and (6.6). Also these tests based on modified spacings can in general, distinguish alternatives at the more standard distance of $n^{-\frac{1}{2}}$, though there may be exceptions as shown by an example (see Section 6 for details).

The reader not interested in the detailed derivations may skip Sections 2, 3 and 4 and go over to Sections 5 and 6 where the efficiency comparisons have been made.

2. Preliminaries

Let $X_1, X_2, \ldots, X_{n-1}$ be $(n-1)$ independently and identically distributed random variables with a continuous d.f. $G_n(x)$, whose carrier is $[0,1]$, $n = 2,3,\ldots$. $X_1,\ldots,X_{n-1}$ may also depend on n, but we shall suppress this in our notations throughout. $G_n(x)$ is a sequence of alternative distributions, which converges to the uniform distribution on $[0,1]$, the distribution specified by the hypothesis.

Assumption (A):

We assume $G_n(x)$ to be of the form

$$G_n(x) = x + L(x)/n^{\partial} \tag{2.1}$$

for $x \in [0,1]$ where ∂ is a fixed constant $\geq 1/4$. We impose the following regularity condition on $L(x)$. $L(x)$ is twice continuously differentiable on $[0,1]$. If $\ell(x)$ and $\ell'(x)$ denote the first and *second derivatives respectively of* $L(x)$, then we note that there is a constant L_o such that

$$L(x)| \leq L_o,\ |\ell(x)| \leq L_o,\ |\ell'(x)| \leq L_o \tag{2.2}$$

for all $x \in [0,1]$.

The inverse function of $G_n(x)$ is denoted by $G_n^{-1}(p)$, $0 \leq p \leq 1$. We define

$$k_n(p) = g_n[G_n^{-1}(p)] = [dG_n^{-1}(p)/dp]^{-1}. \tag{2.3}$$

It may be verified that in our case

$$G_n^{-1}(p) = p - L(p)/n^{\partial} + o(1/n^{\partial}) \tag{2.4}$$

$$k_n(p) = 1 + \ell(p)/n^{\partial} - L(p)\ell'(p)/n^{2\partial} + o(1/n^{\partial+\partial^*}) \tag{2.5}$$

where $o(\cdot)$ is uniform in p and $\partial^* = \max(0, \frac{1}{2} - \partial)$.

We will obtain several limit distributions under the sequence of alternatives $G_n(x)$ satisfying assumption (A). It is clear, however, that the limit distributions under the hypothesis are obtained by putting $L(x) \equiv 0$. We will make some further remarks about these alternatives in Section 6. Let the random variables (r.v.'s) $X_1,\ldots,X_{n-1}$ be arranged in increasing order of magnitude thus

$$(2.6) \qquad 0 \le X_1' \le \cdots \le X_{n-1}' \le 1.$$

The sample spacings have been defined in Section 1 as

$$(2.7) \qquad D_i = X_i' - X_{i-1}', \quad i = 1,\ldots,n$$

where we put $X_0' = 0$, $X_n' = 1$.

We first relate these sample spacings D_i to the spacings based on uniformly distributed r.v.'s on $[0,1]$ (to be called uniform sample spacings). Let $U_1,\ldots,U_{n-1}$ be $(n-1)$ independently and identically distributed r.v.'s with a uniform distribution on $[0,1]$. These are arranged in increasing order of magnitude thus

$$0 \le U_1' \le \cdots \le U_{n-1}' \le 1.$$

The uniform sample spacings are defined by

$$(2.8) \qquad T_i = U_i' - U_{i-1}', \quad i = 1,\ldots,n$$

where again we put $U_0' = 0$, $U_n' = 1$.

For two r.v.'s X and Y, we write $X \sim Y$ to mean that X and Y are distributionally equivalent, that is, the distributions of X and Y are identical. We know that

$$(X_i', \; i = 0,\ldots,n) \sim (G_n^{-1}(U_i'), \; i = 0,\ldots,n)$$

and thus

$$(D_i,\ \ i = 1,\ldots,n) \sim (G_n^{-1}(U_i') - G_n^{-1}(U_{i-1}'),\ \ i = 1,\ldots,n)$$

$$= (T_i/k_n(\tilde{U}_i),\ \ i = 1,\ldots,n)$$

$$\text{where}\ \ U_{i-1}' \le \tilde{U}_i \le U_i'$$

$$= (T_i/\alpha_{ni}^*,\ \ i = 1,\ldots,n) \tag{2.9}$$

where

$$\alpha_{ni}^* = 1 + \beta_{ni}^*/n^{\partial} + \gamma_{ni}^*/n^{2\partial} + R_{ni}^* \tag{2.10}$$

with

$$\beta_{ni}^* = \ell(\tilde{U}_i) \tag{2.11}$$

$$\gamma_{ni}^* = -L(\tilde{U}_i)\ell'(\tilde{U}_i) \tag{2.12}$$

and

$$\sup_i \sqrt{n}\ |R_{ni}^*| \to 0 \ \text{ almost everywhere} \tag{2.13}$$

in view of (2.5). Also, from the existence of the limiting distribution of the Kolmogorov-Smirnov statistic,

$$\sup_i \sqrt{n}\ |U_i' - i/n| = O_p(1). \tag{2.14}$$

Thus from the continuity of L, ℓ and ℓ',

$$\sup_i n^{\partial^*} |\beta_{ni}^* - \beta(i/n)| = o_p(1) \tag{2.15}$$

$$\sup_i |\gamma_{ni}^* - \gamma(i/n)| = o_p(1) \tag{2.16}$$

where

(2.17) $$\beta(p) = \ell(p)$$

(2.18) $$\gamma(p) = -L(p)\ell'(p), \quad 0 \le p \le 1.$$

Now, let $W_1,\ldots,W_n$ be n independently and identically distributed exponential r.v.'s with density function e^{-w}, $w \ge 0$. Let $W_n^* = (W_1 + \cdots + W_n)$ and let $\overline{W}_n = W_n^*/n$. Then it is well known that

$$(T_i, \ i = 1,\ldots,n) \sim (W_i/W_n^*, \ i = 1,\ldots,n).$$

Thus (2.9) may be rewritten as

(2.19) $$(D_i, \ i = 1,\ldots,n) \sim (W_i/\alpha_{ni}^{**} W_n^*, \ i = 1,\ldots,n)$$

where

(2.20) $$(\alpha_{ni}^{**}, \ i = 1,\ldots,n) \sim (\alpha_{ni}^*, \ i = 1,\ldots,n).$$

In view of (2.20), we save on notation by writing α_{ni}^* for α_{ni}^{**} and retain its structure defined in (2.10) and will later on utilise the properties (2.13), (2.15) and (2.16).

The empirical d.f., $H_n(x)$ of the normalised spacings is defined as follows

(2.21) $$H_n(x) = \sum_1^n I(nD_i;\, x)/n, \quad x \ge 0$$

where

(2.22) $$I(z;\, x) = \begin{cases} 1 & \text{if} \quad z \le x \\ 0 & \text{if} \quad z > x \end{cases}.$$

Using the equivalence (2.19), we note that

$$\{H_n(x), \ x \ge 0\} \sim \{ \sum_1^n I(W_i/\alpha_{ni}^* \overline{W}_n;\, x)/n, \quad x \ge 0\}$$

(2.23) $$= \{F_n^*(x\overline{W}_n), \ x \ge 0\}$$

where

$$F_n^*(x) = \sum_1^n I(W_i/\alpha_{ni}^*;\ x)/n. \tag{2.24}$$

Relation (2.23) says that the distributions of the stochastic processes $\{H_n(x), x \geq$ and $\{\sum_1^n I(W_i/\alpha_{ni}^* \overline{W}_n;\ x)/n,\ \ x \geq 0\}$ in $D[0,\infty]$ coincide and this distributional equivalence is stronger than the distributional equivalence of the finite dimensional marginals. We refer to $F_n^*(x)$ as the empirical d.f. of $W_1,\ldots,W_n$ with random perturbations and a random scale factor $\overline{W}_n$.

If $(h_{n1}, h_{n2},\ldots,h_{nn})$, $n = 1,2,\ldots$ be a triangular array of positive constants, define

$$D_i^* = nD_i/h_{ni},\quad i = 1,\ldots,n. \tag{2.25}$$

We shall call $(D_1^*,\ldots,D_n^*)$ modified spacings modified by $h_{n1},\ldots,h_{nn}$ and the empirical d.f. of these modified spacings is defined by $H_n^*(x)$ where

$$H_n^*(x) = \sum_1^n I(D_i^*;\ x)/n. \tag{2.26}$$

From (2.19), it follows that

$$\{H_n^*(x),\ x \geq 0\} \sim \{F_n^*(x\overline{W}_n),\ \ x \geq 0\} \tag{2.27}$$

where the α_{ni}^*'s used in the definition (2.24) of $F_n^*(x)$ here are distributionally equivalent as follows:

$$\{\alpha_{ni}^*,\ \ i = 1,\ldots,n\} \sim \{h_{ni}(1+\beta_{ni}^*/n^{\partial} + \gamma_{ni}^*/n^{2\partial} + R_{ni}^*),\ \ i = 1,\ldots,n\} \tag{2.28}$$

where β_{ni}^*, γ_{ni}^* and R_{ni}^* satisfy the conditions laid down in (2.15), (2.16) and (2.13). As in the remark after (2.20), we replace the symbol '$\sim$' in (2.28) by '$=$' in order to avoid introducing new notations.

3. Asymptotic distribution of the empirical d.f. of random variables subject to perturbations.

Let $Z_1, Z_2, \ldots$ be independently and identically distributed r.v.'s with a common d.f. $F(x)$ with $F(0) = 0$. We assume that $F(x)$ is thrice differentiable. Let $f(x)$, $f'(x)$ and $f''(x)$ denote the first, second and third derivatives respectively of $F(x)$. We impose the following blanket condition (B) on $F(x)$.

Assumption (B):

$xf(x)$, $x^2 f'(x)$ and $x^3 f''(x)$ are bounded on $[0,\infty]$.

Let $\{\alpha_{ni1},\ i = 1,\ldots,n;\ n = 1,2,\ldots,\}$ be a triangular array of constants. Then the random variables

$$Z_{ni} = Z_i/\alpha_{ni1}, \quad i = 1,\ldots,n$$

are said to be perturbed random variables, $n = 1,2,\ldots$. Let

$$(3.1) \qquad F_{n1}(x) = \sum_1^n I(Z_{ni};\ x)/n$$

$$= \sum_1^n I(Z_i/\alpha_{ni1};\ x)/n.$$

We refer to $F_{n1}(x)$ as the empirical d.f. of $(Z_1,\ldots,Z_n)$ under a perturbation by the non-random quantities $\{\alpha_{ni1},\ i = 1,\ldots,n\}$. The following structure is assumed of $\{\alpha_{ni1},\ i = 1,\ldots,n\}$. There exist continuous functions $\beta(p)$ and $\gamma(p)$ on $[0,1]$ such that

$$(3.2) \qquad \alpha_{ni1} = 1 + \beta(i/n)/n^{\partial} + \gamma(i/n)/n^{2\partial} + R_{ni}$$

where ∂ is a constant $\geq 1/4$ and

$$(3.3) \qquad \sup_i \sqrt{n}\,|R_{ni}| \to 0 \text{ as } n \to \infty.$$

If $\partial > 1/2$, then the second and third terms on the right hand side (RHS) of (3.2) can be absorbed into R_{ni} and if $1/4 < \partial \le \frac{1}{2}$, then the third term of the RHS can be absorbed into R_{ni}. We note that $\alpha_{ni1} \to 1$ uniformly in i, so that without loss of generality we may assume

$$1/2 \le \alpha_{ni1} \le 2 \tag{3.4}$$

for all n and i. Let

$$F^+_{n1}(x) = E(F_{n1}(x))$$

$$= E\left(\sum_1^n I(Z_i/\alpha_{ni1};\ x)/n\right)$$

$$= \sum_1^n F(x\alpha_{ni1})/n \tag{3.5}$$

and

$$\eta_{n1}(x) = \sqrt{n}\,(F_{n1}(x) - F^+_{n1}(x)), \quad x \ge 0. \tag{3.6}$$

It is easy to see that $F^+_{n1}(x)$ tends to $F(x)$ uniformly in x.

Remark 1: When condition (B) holds, we can replace $F^+_{n1}(x)$ which enters the definition of $\eta_{n1}(x)$ in (3.6) by

$$F^+{}_{n1}(x) = \begin{cases} F(x) & \text{if } \partial > 1/2 \\ F(x) + xf(x)\int_0^1 \beta(p)dp/n^{\partial} & \text{if } 1/4 < \partial \le 1/2 \\ F(x) + xf(x)\int_0^1 \beta(p)dp/n^{1/4} & \\ \quad + [xf(x)\int_0^1 \gamma(p)dp + x^2 f'(x)\int_0^1 \beta^2(p)dp/2]n^{\frac{1}{2}} & \text{if } \partial = \frac{1}{4} \end{cases} \tag{3.7}$$

after omitting terms which are of smaller order than $n^{-1/2}$ uniformly in x. The most general conditions under which $F^+_{n1}(x)$ can be replaced as above must depend on ∂. However since we are contemplating only the application with $F(x) = 1 - \exp(-x)$ in Section 6, we will content ourselves by imposing the blanket condition (B).

The following theorem which establishes the weak convergence of $\{\eta_{n1}(x),\ x \geq 0\}$ is taken from RS, where the weak convergence on the space $D[0,\infty]$ is also briefly explained

Theorem 3.1 (RS): Let condition (B) hold. The sequence $\{\eta_{n1}(x),\ x \geq 0\}$ considered as a stochastic process in $D[0,\infty]$ converges weakly to a Gaussian process $\{\eta_1(x),\ x \geq 0\}$ with mean zero and covariance function

$$K_1(x,y) = K_1(y,x) = F(x)(1 - F(y)) \quad \text{for } x \leq y. \tag{3.8}$$ ▮

To allow for perturbations by constants which are more general than given in (3.2), we consider a triangular sequence $\{\alpha_{ni2},\ i = 1,\ldots,n\}$, $n = 1,2,\ldots$ with the following structure

$$\alpha_{ni2} = \theta(i/n)[1 + \beta(i/n)/n^{\partial} + \gamma(i/n)/n^{2\partial} + R_{ni}] \tag{3.9}$$

where $\beta(p)$ and $\gamma(p)$ are continuous functions on $[0,1]$ and R_{ni} satisfies (3.3). We put the following condition (C) on $\theta(p)$.

Assumption (C):

$\theta(p)$ is continuous on $[0,1]$ except at a finite number of points and, for each x, the integrals of $F(x\theta(p))$, $\theta(p)f(x\theta(p))$ and $\theta^2(p)f'(x\theta(p))$ as functions of p on $[0,1]$ exist and are finite.

We shall see later that this generalisation gives results which enable us to obtain the limiting distributions of statistics based on modified spacings. Define the empirical d.f. $F_{n2}(x)$ of the Z's perturbed by the $\{\alpha_{ni2}\}$ given in (3.9), by a formula similar to (3.1). Let $F^+_{n2}(x)$ and $\eta_{n2}(x)$ be as defined in (3.5) and (3.6), with the perturbation constants $\{\alpha_{ni2}\}$ instead of $\{\alpha_{ni1}\}$. The following theorem is proved exactly as Theorem 3.1 and is stated without proof.

Theorem 3.2: Let assumptions (B) and (C) hold. The sequence of stochastic processes $\{\eta_{n2}(x),\ x \geq 0\}$ in $D[0,\infty]$ converges to the Gaussian stochastic process $\{\eta_2(x),\ x \geq 0\}$ with mean zero and covariance function $K_2(x,y)$ defined by

$$K_2(x,y) = K_2(y,x) = \int_0^1 F(x\theta(p))[1 - F(y\theta(p))]dp \quad \text{for} \quad x \leq y. \tag{3.10}$$

Remark 2: Under conditions (B) and (C), $F^+_{n2}(x)$, which is defined by (3.5) through the constants $\{\alpha_{ni2}\}$ of (3.9) can be replaced by

$$F^+_{n2}(x) = \begin{cases} \int_0^1 F(x\theta(p))dp & \text{if} \quad \partial > 1/2 \\ \int_0^1 F(x\theta(p))dp + \int_0^1 x\beta(p)\theta(p)f(x\theta(p))dp/n^{\partial} & \text{if} \quad \frac{1}{4} < \partial \leq \frac{1}{2} \\ \int_0^1 F(x\theta(p))dp + \int_0^1 x\beta(p)\theta(p)f(x\theta(p))dp/n^{1/4} & \\ \quad + \left[\int_0^1 x\gamma(p)\theta(p)f(x\theta(p))dp + \int_0^1 x^2\beta^2(p)\theta^2(p)f'(x\theta(p))dp/2\right]/n^{1/2} & \\ & \text{if} \quad \partial = 1/4 \end{cases} \tag{3.11}$$

up to terms of smaller order than $n^{-1/2}$ uniformly in x.

We now proceed to establish a limit theorem for the empirical d.f. of randomly perturbed r.v.'s . Let $\{\alpha^*_{ni1},\ i = 1,\ldots,n\}$ $n = 1,2,\ldots$ be a triangular scheme of random variables with the form

$$\alpha^*_{ni1} = 1 + \beta^*_{ni}/n^{\partial} + \gamma^*_{ni}/n^{2\partial} + R^*_{ni} \tag{3.12}$$

where

$$\sup_i \sqrt{n}\ |R^*_{ni}| = o_p(1) \tag{3.13}$$

and there are continuous functions $\beta(p)$ and $\gamma(p)$ on $[0,1]$ such that

$$\sup_i n^{\partial^*}|\beta^*_{ni} - \beta(i/n)| = o_p(1)$$

(3.14)

$$\sup_i |\gamma^*_{ni} - \gamma(i/n)| = o_p(1)$$

where

(3.15)
$$\partial^* = \begin{cases} \frac{1}{2} - \partial & \text{if} \quad \partial < \frac{1}{2} \\ 0 & \text{if} \quad \partial \geq \frac{1}{2}. \end{cases}$$

Let

(3.16)
$$F^*_{n1}(x) = \sum_1^n I(Z_i/\alpha^*_{ni1};\ x)/n.$$

This is the empirical d.f. of $Z_1,\ldots,Z_n$ perturbed by the random quantities $\{\alpha^*_{ni1},\ i = 1,\ldots,n\}$. Let the non-random quantities $\{\alpha_{ni1}\}$ be defined as

(3.17)
$$\alpha_{ni1} = 1 + \beta(i/n)/n^{\partial} + \gamma(i/n)/n^{2\partial} + R_{ni}$$

in terms of the new $\beta(p)$ and $\gamma(p)$ and $F^+_{n1}(x)$ be defined by the relation (3.5) with the new α_{ni1}'s. We have the following result from RS on the process

(3.18)
$$\left\{\eta^*_{n1}(x) = \sqrt{n}\,(F^*_{n1}(x) - F^+_{n1}(x)),\quad x \geq 0\right\}.$$

Theorem 3.3 (RS): Let condition (B) hold. The sequence of processes $\{\eta^*_{n1}(x),\ x \geq 0\}$ in $D[0,\infty]$ converges weakly to the Gaussian process $\{\eta_1(x),\ x \geq 0\}$ with mean zero and covariance function given by (3.8).

Now suppose that $\{\alpha^*_{ni2},\ i = 1,\ldots,n\}$, $n = 1,2\ldots$ is a triangular scheme of r.v.'s with the form

$$(3.19) \qquad \alpha^*_{ni2} = \theta^*_{ni}\left[1 + \beta^*_{ni}/n^{\partial} + \gamma^*_{ni}/n^{2\partial} + R^*_{ni}\right]$$

where R^*_{ni}, β^*_{ni} and γ^*_{ni} satisfy the assumptions (3.13) and (3.14) and further there is a function $\theta(p)$ on $[0,1]$ such that

$$(3.20) \qquad \sup_i \sqrt{n}\, |\theta^*_{ni} - \theta(i/n)| = o_p(1)$$

and $\theta(p)$ satisfies condition (C). Now define $F^*_{n2}(x)$, $F^+_{n2}(x)$ and $\eta^*_{n2}(x)$ similar to the expressions in (3.16), (3.5) and (3.18), respectively with $\{\alpha^*_{ni2}\}$ instead of $\{\alpha^*_{ni1}\}$. The following theorem then follows from Theorem 3.2 in exactly the same way as Theorem 3.3 follows from Theorem 3.1.

<u>Theorem 3.4</u>: Let assumptions (B) and (C) hold. The sequence of stochastic processes $\{\eta^*_{n2}(x),\ x \geq 0\}$ converges weakly to the Gaussian process $\{\eta_2(x),\ x \geq 0\}$ in $D[0,\infty]$ where $\{\eta_2(x),\ x \geq 0\}$ has mean zero and covariance function $K_2(x,y)$ given in (3.10).

A note of clarification regarding the notation may be in order. There are four different sets of perturbation constants in all, namely: $\{\alpha_{ni1}\}$ of Theorem 3.1 satisfying conditions (3.2) and (3.3); $\{\alpha_{ni2}\}$ of Theorem 3.2 satisfying (3.3) and (3.9); $\{\alpha^*_{ni1}\}$ of Theorem 3.3 satisfying (3.12)-(3.14); and finally $\{\alpha^*_{ni2}\}$ of Theorem 3.4 satisfying (3.13), (3.14), (3.19) and (3.20). It may be noted that a subscript 2 is attached to denote perturbation constants with modifying term $\theta(i/n)$ in them as opposed to the use of subscript 1 where such a term is absent. Similarly the random perturbation constants, as well as any quantities like empirical distribution functions based on those, are starred (as opposed to the analogous unstarred version based on non-random perturbations). Thus $\eta_{n2}(x)$ is the process based on the non-random $\{\alpha_{ni2}\}$ which have the modifying term $\theta(i/n)$ in them while $\eta^*_{n1}(x)$ is based on the random $\{\alpha^*_{ni1}\}$ without the $\theta(i/n)$ term in them. The four theorems of the next section also correspond to these four cases.

4. Asymptotic distribution of the empirical d.f. when the random variables are subject to perturbations and a random scale factor.

We retain the notations of the earlier sections. Let $\eta_{n1}(x)$ be defined as in (3.6) through $F_{n1}(x)$ and $F^+_{n1}(x)$ which are in turn defined as in (3.1) and (3.5) and the α_{ni1}'s have the structure (3.2).

Let Z^*_n be a r.v. and let $\xi_n = \sqrt{n}(Z^*_n - 1)$. We now make the following assumption (D) on the Stochastic process $\{\eta_{n1}(x),\ x \ge 0\}$ and ξ_n.

Assumption (D):

For any finite collection $(x_1, \ldots, x_k)$, the distribution of $\{\eta_{n1}(x_1), \eta_{n1}(x_2), \ldots, \eta_{n1}(x_k), \xi_n\}$ converges weakly to the distribution of $\{\eta_1(x_1), \ldots, \eta_1(x_k), \xi\}$, which is multivariate normal with zero means and covariances given by

$$\text{(4.1)} \qquad \operatorname{cov}(\eta_1(x_i), \eta_1(x_j)) = K_1(x_i, x_j), \quad 1 \le i, \quad j \le k$$

where $K_1(x,y)$ is as defined in (3.8) and

$$\text{(4.2)} \qquad \operatorname{cov}(\eta_1(x_i), \xi) = a_1(x_i), \quad i = 1, \ldots, k$$

and

$$\text{(4.3)} \qquad \operatorname{var}(\xi) = 1.$$

We add the following to the assumption (B) made on $F(x)$ in Section 3.

Assumption (B*):

There is an $\alpha > 0$ such that $x^\alpha(1 - F(x)) \to 0$ and $xf(x) \to 0$ as $x \to \infty$. Again from RS we have

<u>Theorem 4.1</u> (RS): Let the assumptions (B), (B^*) and (D) hold. Let

(4.4) $$\zeta_{n1}(x) = \sqrt{n}\,(F_{n1}(xZ_n^*) - F_{n1}^+(x)).$$

Then

(4.5) $$\sup_{0 \le x \le \infty} |\zeta_{n1}(x) - \eta_{n1}(x) - \xi_n x f(x)| = o_p(1).$$

Thus $\{\zeta_{n1}(x),\ x \ge 0\}$ converges weakly to the Gaussian process $\{\zeta_1(x) = \eta_1(x) + xf(x)\xi,\ \ x \ge 0\}$ in $D[0,\infty]$ which has mean zero and covariance function

(4.6) $$\begin{aligned} K_3(x,y) &= K_3(y,x) \\ &= K_1(x,y) + xyf(x)f(y) + xf(x)a_1(y) + yf(y)a_1(x). \end{aligned}$$

We now extend this result to the more general non-random perturbation factors $\{\alpha_{ni2}\}$ defined in (3.8). Let $F_{n2}(x)$, $F_{n2}^+(x)$ and $\eta_{n2}(x)$ be as defined and used in Theorem 3.2. If Z_n^* be a r.v., we assume that $\xi_n = \sqrt{n}\,(Z_n^* - 1)$ satisfies the following assumption (D^*) with the process $\{\eta_{n2}(x),\ x \ge 0\}$.

<u>Assumption (D^*)</u>:

For any finite collection $(x_1,\ldots,x_k)$, the distribution of $\{\eta_{n2}(x_1),\ldots,\eta_{n2}(x_k),\xi\}$ converges weakly to that of $\{\eta_2(x_1),\ldots,\eta_2(x_k),\xi\}$ which is a multivariate normal distribution with zero means and covariances given by

(4.7) $$\operatorname{cov}(\eta_2(x_i),\eta_2(x_j)) = K_2(x_i,x_j), \quad 1 \le i,\ \ j \le k$$

where $K_2(x,y)$ is as defined in (3.10) and

(4.8) $$\operatorname{cov}(\eta_2(x_i), \xi) = a_2(x_i), \quad i = 1,\ldots,k$$

and

$$\operatorname{var}(\xi) = 1.$$

Then we have the following theorem whose proof follows on the lines of the proof of Theorem 4.1 and is omitted.

Theorem 4.2: Let conditions (B), (B^*), (C) and (D^*) hold. Let

(4.9) $$\zeta_{n2}(x) = \sqrt{n}\,[F_{n2}(xZ_n^*) - F_{n2}^+(x)].$$

Then

(4.10) $$\sup_{0 \le x \le \infty} |\zeta_{n2}(x) - \eta_{n2}(x) - \xi_n x(\int_0^1 \theta(p)f(x\theta(p))dp)| = o_p(1).$$

Thus $\{\zeta_{n2}(x),\ x \ge 0\}$ converges weakly in $D[0,\infty]$ to the Gaussian process $\{\zeta_2(x) = \eta_2(x) + \xi\, x(\int_0^1 \theta(p)f(x\theta(p))dp),\ x \ge 0\}$ which has mean zero and covariance function

$$K_4(x,y) = K_4(y,x)$$
$$= K_2(x,y) + xy(\int_0^1 \theta(p)f(x\theta(p))dp)(\int_0^1 \theta(p)f(y\theta(p))dp)$$

(4.11) $$+ xa_2(y)(\int_0^1 \theta(p)f(x\theta(p))dp) + ya_2(x)(\int_0^1 \theta(p)f(y\theta(p))dp)$$

with $K_2(x,y)$ as in (3.10).

Now coming to the case of random perturbation factors, let $\{\alpha_{ni1}^*\}$ be as in (3.12) and $F_{n1}^*(x)$, $F_{n1}^+(x)$ and $\eta_{n1}^*(x)$ be as defined and used in Theorem 3.3 . Let $\{\alpha_{ni1}\}$, the non-random constants generated from $\{\alpha_{ni1}^*\}$ be as in (3.17). Let Z_n^*, ξ_n be as used in Theorem 4.1 and satisfy the condition (D) with the process $\{\eta_{n1}^*(x),\ x \ge 0\}$. Then we have the following extension of Theorem 4.1 to the case of random perturbations from RS.

Theorem 4.3 (RS): Let the conditions (B), (B^*) and (D) hold. Let

(4.12) $$\zeta_{n1}(x) = \sqrt{n}\,[F^*_{n1}(xZ^*_n) - F^+_{n1}(x)]$$

then

(4.13) $$\sup_{0 \le x \le \infty} |\zeta_{n1}(x) - \eta^*_{n1}(x) - \xi_n x f(x)| = o_p(1).$$

Thus $\{\zeta_{n1}(x),\ x \ge 0\}$ converges weakly to the Gaussian process $\{\zeta_1(x) = \eta_1(x) + \xi x f(x),\ x \ge 0\}$ defined in Theorem 4.1.

Finally let $\{\alpha^*_{ni2}\}$ be the more general random perturbation factors defined in (3.19) and satisfy the conditions stipulated there. Let $\{\alpha_{ni2}\}$, the non-random constants 'generated' by $\{\alpha^*_{ni2}\}$ be as defined in (3.9). Let $F^*_{n2}(x)$, $F^+_{n2}(x)$ and $\eta^*_{n2}(x)$ be as defined and used in Theorem 3.4. Let Z^*_n, ξ_n be as used in Theorem 4.2 and satisfy the condition (D^*) with the process $\{\eta^*_{n2}(x),\ x \ge 0\}$. Then the following theorem can be deduced from Theorem 4.2 in the usual way.

Theorem 4.4: Let the conditions (B), (B^*), (C) and (D^*) hold. Let

(4.14) $$\zeta_{n2}(x) = \sqrt{n}[F^*_{n2}(x\,Z^*_n) - F^+_{n2}(x)].$$

Then

(4.15) $$\sup_{0 \le x \le \infty} |\zeta_{n2}(x) - \eta^*_{n2}(x) - \xi_n x (\int_0^1 \theta(p) f(x\theta(p))dp)| = o_p(1).$$

Thus the process $\{\zeta_{n2}(x),\ x \ge 0\}$ converges weakly in $D[0,\infty]$ to the Gaussian process $\{\zeta_2(x),\ x \ge 0\}$ defined in Theorem 4.2 with mean zero and covariance function given by (4.11).

5. Asymptotic distributions of the empirical d.f.'s of normalised and modified spacings and tests based on them.

In this section, we relate the results of the last two sections to the spacings statistics. First we give the asymptotic distributions of the empirical d.f.'s of the normalised spacings $H_n(x)$ and of the modified spacings $H_n^*(x)$, defined in (2.21) and (2.26) respectively using the distributional equivalences (2.23) and (2.27). We then establish the asymptotic normality of some classes of test statistics based on these spacings.

We shall first consider the empirical d.f. of the normalised spacings $H_n(x)$, which from (2.23) is distributionally equivalent to $F_n^*(x\overline{W}_n)$. The r.v.'s $W_1, W_2, \ldots$ have the exponential d.f.

(5.1) $$F(x) = 1 - e^{-x}, \quad x \ge 0$$

which satisfies all the regularity conditions of Theorem 4.1 and the assumptions (B) and (B^*). Further the $\{\alpha_{ni1}^*\}$ used in the definition of $F_{n1}^*(x)$ satisfy the conditions (3.12), (3.13) and (3.14) with $\beta(p)$ and $\gamma(p)$ given by (2.17) and (2.18). Hence we have

(5.2) $$\int_0^1 \beta(p)dp = \int_0^1 \ell(p)dp = 0$$

$$\int_0^1 \gamma(p)dp = -\int_0^1 L(p)\ell'(p)dp = \int_0^1 \ell^2(p)dp = \int_0^1 \beta^2(p)dp.$$

Let

(5.3) $$\zeta_{n1}^*(x) = \sqrt{n}\,[H_n(x) - F_{n1}^+(x)]$$

$$\sim \sqrt{n}\,[F_{n1}^*(x\overline{W}_n) - F_{n1}^+(x)]$$

where

$$(5.4) \qquad F_{n1}^{+}(x) = \begin{cases} (1-e^{-x}) & \text{for } \partial > 1/4 \\ (1-e^{-x}) + \left(\int_0^1 \ell^2(p)dp\right)e^{-x}(x-x^2/2)/\sqrt{n} & \text{for } \partial = 1/4 \end{cases}$$

ignoring terms which are of smaller order than $n^{-1/2}$ uniformly in x. Further since the random scale factor here is $\overline{W}_n$, assumption (D) is satisfied and $a_1(x)$, defined in (4.2) is easily seen to be $(-xe^{-x})$. In view of these remarks we have the following theorem as a consequence of Theorem 4.3 as in RS.

Theorem 5.1 (RS): The sequence of stochastic processes $\{\zeta_{n1}^*(x),\ x \geq 0\}$ converges weakly to a Gaussian process $\{\zeta_1(x),\ x > 0\}$ with mean zero and covariance function

$$(5.5) \qquad K_3(x,y) = e^{-y}(1-e^{-x}-xye^{-x}),\quad x \leq y.$$

The invaviance principle may be invoked to obtain the limit distributions of various functionals of $\zeta_{n1}(x)$ and their ARE's computed, as was done in SR.

Consider now the modified spacings

$$(5.6) \qquad D_i^* = nD_i/h_{ni}$$

where h_{ni} satisfy

$$(5.7) \qquad \sup_i \sqrt{n}\,|h_{ni} - h(i/n)| = o(1)$$

where $h(p)$ is a function on $[0,1]$ having at most a finite number of discontinuities. Then the empirical d.f. $\{H_n^*(x),\ x \geq 0\}$ of $\{D_1^*,\ldots,D_n^*\}$

defined in (2.26) is distributionally equivalent to $\{F^*_{n2}(x\overline{W}_n),\ x \geq 0\}$ where $F^*_{n2}(x)$ is the empirical d.f. of the exponentially distributed r.v.'s $W_1,\ldots,W_n$ perturbed by the random factors $\{\alpha^*_{ni2},\ i = 1,\ldots,n\}$ which have the structure defined in (3.19), i.e.,

$$\alpha^*_{ni2} = \theta^*_{ni}(1 + \beta^*_{ni}/n^{\partial} + \gamma^*_{ni}/n^{2\partial} + R^*_{ni})$$

where θ^*_{ni}, β^*_{ni}, γ^*_{ni} and R^*_{ni} satisfy (3.20), (3.14), and (3.13) respectively with

$$\begin{aligned} &\theta(p) = h(p) \\ (5.8)\qquad &\beta(p) = \gamma(p) \\ &\gamma(p) = -L(p)\,\ell'(p), \quad 0 \leq p \leq 1. \end{aligned}$$

Let

$$(5.9)\qquad \zeta^*_{n2}(x) = \sqrt{n}\,[H^*_n(x) - F^+_{n2}(x)]$$

where

$$F^+_{n2}(x) = \int_0^1 (1 - e^{-xh(p)})dp \qquad \text{if } \partial > 1/2$$

$$= \int_0^1 (1 - e^{-xh(p)})dp + \left(\int_0^1 xe^{-xh(p)}\,\ell(p)h(p)dp\right)/n^{\partial}$$

$$(5.10)\qquad\qquad \text{if } 1/4 < \partial \leq 1/2$$

$$= \int_0^1 (1 - e^{-xh(p)})dp + \left(\int_0^1 xe^{-xh(p)}\,\ell(p)h(p)dp\right)/n^{1/4}$$

$$+ \int_0^1 \left[-xL(p)\,\ell'(p)h(p)e^{-xh(p)} - x^2h^2(p)\,\ell^2(p)e^{-xh(p)}/2\right]dp/n^{\frac{1}{2}}$$

$$\text{if } \partial = 1/4$$

to terms of order $n^{-1/2}$. As an immediate consequence of Theorem 4.4, since $a_2(x)$ defined in (4.8) is $\left(-x\int_0^1 h(p)e^{-xh(p)}dp\right)$, we have the following

<u>Theorem 5.4</u>: The sequence of stochastic processes $\{\zeta^*_{n2}(x), \ x \geq 0\}$ in $D[0,\infty]$ converges weakly to the Gaussian process $\{\zeta_2(x), \ x \geq 0\}$ with mean zero and covariance function

$$(5.11) \qquad K_4(x,y) = K_4(y,x) = \int_0^1 e^{-yh(p)}(1-e^{-xh(p)})dp$$

$$- xy\left(\int_0^1 h(p)e^{-xh(p)}dp\right)\left(\int_0^1 h(p)e^{-yh(p)}dp\right)$$

for $x \leq y$.

The invariance principle immediately gives

<u>Theorem 5.5</u>: Let $m(x)$ be an absolutely continuous function on $[0,\infty]$ with $m(0) < \infty$. Let $m'(x)$ be bounded on every finite interval and let the function on $D[0,\infty]$

$$y(\cdot) \to \int_0^\infty m'(x)y(x)dx$$

be almost everywhere continuous with respect to the Gaussian process $\{\zeta_2(x), \ x \geq 0\}$ defined in Theorem 5.4. Let

$$(5.12) \qquad T_n = \sum_1^n m(n\,D_i^*)/n.$$

Then the distribution of

$$(5.13) \qquad \sqrt{n}\,[T_n - \int_0^\infty m'(x)(1 - F^+_{n2}(x))\,dx + m(0)]$$

where $F^+_{n2}(x)$ is defined in (5.10), converges weakly to the normal distribution with mean zero and variance

$$(5.14) \qquad \int_0^\infty \int_0^\infty K_4(x,y)m'(x)m'(y)dx\,dy$$

where $K_4(x,y)$ is as defined in (5.11).

This theorem covers a very wide range of statistics based symmetrically on modified spacings.

6. Asymptotic relative efficiences of tests based on modified spacings.

The Pitman asymptotic relative efficiency (ARE) of a test relative to another test 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. This limiting power should be a value in between the limiting size, α and the maximum power 1, in order that it can give an insight into the power behaviour of the test. If the limiting power of a test at a sequence of alternatives is α, then its ARE with respect to any other test whose limiting power (with same size) is greater than α, is zero. On the other hand, if the limiting power of a test at a sequence of alternatives converges to a number in the interval $(\alpha, 1)$, then a measure of the rate of this convergence, called 'efficacy' can be computed. Under certain standard regularity assumptions (see, e.g., Fraser (1957)) which include a condition about the nature of alternatives, asymptotic normal distribution of the test statistic under these alternatives, etc., this 'efficacy' is given by

$$\text{efficacy} = \mu_\partial^4 / \sigma^4 . \tag{6.1}$$

Here μ_∂ and σ^2 are the mean and variance of the limiting normal distribution under the sequence of alternatives when the test-statistic has been normalised to have a limiting normal distribution with mean zero and finite variance under the hypothesis. In such a situation, the ARE of one test with respect to another is simply the ratio of their efficacies.

Using Theorem 5.5, we can now compute the ARE's of tests which are symmetric in modified spacings. We defined $\{D_i^* = nD_i/h_{ni},\ i = 1,\dots,n\}$ as the modified spacings where the factors $\{h_{ni}\}$ satisfy the condition

$$\sup_i \sqrt{n}\, | h_{ni} - h(i/n) | = o(1). \tag{6.2}$$

If m be any function on $[0, \infty]$ satisfying the conditions of Theorem 5.5 we define a symmetric statistic based on the modified spacings

$$T_n^* = \sum_{i=1}^{n} m(D_i^*)/n \tag{6.3}$$

The mean under the hypothesis of this T_n^* is, say

$$\mu_{on}^* = \int_0^\infty \int_0^1 m'(x) e^{-xh(p)}\, dx\, dp \tag{6.4}$$

and under the alternatives (2.1) say

$$\begin{aligned} \mu_{1n}^* &= \mu_{on}^* \quad \text{if } \partial > 1/2 \\ &= \mu_{on}^* + A(m, L, h)/n^{\partial}, \quad \text{say if } 1/4 < \partial \le 1/2 \\ &= \mu_{on}^* + A(m, L, h)/n^{1/4} + B(m, L, h)/n^{1/2}, \quad \text{say if } \partial = 1/4 \end{aligned} \tag{6.5}$$

where

$$A(m, L, h) = -\int_0^\infty \int_0^1 m'(x) x \ell(p) h(p) e^{-xh(p)}\, dx\, dp$$

$$\begin{aligned} B(m, L, h) = \int_0^\infty \int_0^1 m'(x) e^{-xh(p)} [&x L(p) \ell'(p) h(p) \\ &+ (x^2/2) \ell^2(p) h^2(p)]\, dx\, dp. \end{aligned} \tag{6.6}$$

If $A(m, L, h) \neq 0$, then the sequence of tests based on T_n^* can distinguish alternatives of the form (2.1) at a distance of order $n^{-1/2}$ from the hypothesis. This shows that such tests have a better performance than tests considered earlier which are symmetric in the normalised spacings. However there is no surety that $A(m, L, h) \neq 0$ for all L. Consider the following example. Let

$$h_{ni} = n/(n-i+1)$$

$$h(p) = 1/(1-p), \quad 0 \le p < 1$$

$$D_i^* = (n-i+1)D_i$$

$$m(x) = x. \tag{6.7}$$

Then

$$T_n^* = \sum_1^n m(D_i^*)/n$$

$$= \sum_{i=1}^n [(n-i+1)D_i]/n$$

$$= \sum_{i=1}^n X_i/n + 1/n. \tag{6.8}$$

A simple computation shows

$$A(m, L, h) = \int_0^1 p\ell(p)dp \tag{6.9}$$

which is n^{∂} times the excess of the mean under the alternative over that under the hypothesis and is zero for alternatives under which T_n^* has a mean $1/2$. But if this excess is non-zero, the test based on $\sum_1^n m(nD_i^*)/n$ has a better performance than symmetric normalised spacings statistics considered in Theorem 5.3. However if $A(m, L, h) = 0$, this test statistic T_n^* discriminates such alternatives, if at all, only when they are at a distance of $n^{-\frac{1}{4}}$, which puts this on par with the symmetric spacings tests.

But it should be remarked that one can always construct tests based symmetrically on modified spacings which have the ability to detect alternatives at a distance of $n^{-\frac{1}{2}}$. This is because in testing the hypothesis of uniformity against the fixed sequence of alternatives $G_n(x) = x + L(x)/n^{\frac{1}{2}}$, the test statistic

$$T_n = \sum_{i=1}^{n} \ell(i/n+1) D_i \tag{6.10}$$

with $h(p) = 1/\ell(p)$ and $m(x) = x$, has $A(m,L,h) = \int_0^1 \ell^2(p)dp \neq 0$. Some recent investigations by Holst and Rao (1978) indicate that tests of the form (6.10) provide the locally most powerful spacings tests for uniformity against the fixed sequence of alternatives $G_n(x)$.

BIBLIOGRAPHY

Chentsov, N. N. (1956): Weak convergence of stochastic processes whose trajectories have no discontinuities of the second kind and the heuristic approach to the Kolmogorov-Smirnov tests, Theor. of Prob. and its Appl., 1, 140-144.

Cibisov, D. M. (1961): On the tests of fit based on sample spacings, Theoria vero. i. Prim., 6, 354-358.

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

Fraser, D. A. S. (1957): Non parametric methods in statistics, John Wiley and Sons, New York.

Greenwood, Major (1946): The statistical study of infectious diseases, Jour. Roy. Stat. Soc., 109, 85-103

Holst, L. and Rao, J.S.: (1978) Some results on asymptotic spacings theory, in preparation.

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

Pyke, R. (1965): Spacings, Jour. Roy. Stat. Soc. (Ser. B), 27, 395-449.

Rao, J. S. (1969): Some contributions to the analysis of circular data. Thesis submitted to the Indian Statistical Institute, Calcutta.

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

Sethuraman, J. (1965): Limit theorems for stochastic processes. Tech. Report No.10, Stanford University.

Sethuraman, J. and J. S. Rao (1970): Pitman efficiencies of tests based on spacings, Non-parametric Techniques in Statistical Inference, Cambridge University Press, 405-415.

Sherman, B. (1950): A random variable related to the spacing of sample values, Ann. Math. Statist. 21, 339-361.

Skorohod, A. V. (1956): Limit theorems for stochastic processes, Theor. of Prob. and its appl. 1, 261-290.

Weiss, L. (1965): On asymptotic sampling theory for distributions approaching the uniform distribution, Z. Wahrscheinlichkeitstheorie Verw. Geb. 4, 217-221.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1. REPORT NUMBER USARO-D No. 30	2. GOVT. ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE Asymptotic Normality and Efficiencies of Tests Based on Modified Spacings	5. TYPE OF REPORT Technical Report 6. PERFORMING ORGANIZATION REPORT NUMBER FSU Statistical Report M489
7. AUTHOR(s) J. S. Rao, J. Sethuraman	8. CONTRACT OR GRANT NUMBERS USARO DAAG29-76-G0238
9. PERFORMING ORGANIZATION NAME AND ADDRESS Florida State University Department of Statistics Tallahassee, Florida 32306	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Research Office-Durham P.O. Box 12211 Research Triangle Park, NC 27709	12. REPORT DATE January, 1979 13. NUMBER OF PAGES 27
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS

Modified spacings, Empirical distribution function, Weak convergence, Efficiencies of tests

20. ABSTRACT

Tests based on adjusted or modified spacings are proposed for testing goodness of fit problems. The weak convergence of the empirical distribution function of such modified spacings is studied using some earlier results of the authors. The asymptotic theory under close alternative sequences is also given thus enabling one to calculate the asymptotic relative efficiencies of such tests.

79