3iometrics, 43, 833–845.
ion for unbalanced designs. J.R. Statist.
h orthogonal block structure. Proc. Roy
ric. Science, 40, 375–380.
Iorth Carolina Agric. Exp. Station Tech.
iization problem in forming designs with
ierlin:Springer-Verlag.
. Int. Statist. Rev., 55, 261–277.
ntal Design. Oxford: Oxford University
s. Int. Statist. Rev., 52, 33–81.
valid restricted randomization schemes
. Statist. Soc. B, 43, 167–172.

Statistics **27** (1996) 325–338
Reprints available directly from the publisher
Photocopying permitted by license only

A NOTE ON SIGNED RANK TESTS FOR THE CHANGEPOINT PROBLEM

JAMES A. KOZIOL

The Scripps Research Institute, La Jolla

(Received 24 March 1993; in final form 19 April 1995)

Summary. We consider two classes of signed rank statistics to test for smooth or abrupt changepoints in sequences of independent random variables. We derive asymptotic null distributions and finite sample approximations for the two classes. We infer from a Monte Carlo power study that the signed rank statistics may compare favorably with parametric analogues in detecting abrupt changes in a sequence of independent normal random variables.

AMS 1980 classifications: 62G10, 65U05.

Key words: Changepoint, weak convergence of stochastic processes, boundary crossing probabilities, Bahadur efficiency.

1. INTRODUCTION AND SUMMARY

The problem of detecting a change in location or scale in a sequence of independent observations has been extensively studied in the statistical literature; see Shaban (1980) and Zacks (1983) for bibliographies. Rank tests have been proposed for this problem, in order to provide procedures that may be robust against deviations from parametric distributional assumptions, or insensitive to spurious outliers; such procedures have been reviewed by Wolfe and Schechtman (1984) and Csörgő and Horváth (1987, 1988). The purpose of this note is to investigate two classes of signed rank statistics in testing for a single abrupt change, or a smooth change (Lombard, 1987), in the center of symmetry. We introduce these two classes, along with standard rank statistic terminology, in Section 2. The first class in an extension of Pettitt (1979), and the second class is a modification of the first, chosen primarily to provide a natural estimate of the position of the changepoint. We derive limiting null distributions, and provide finite sample approximations in Section 3, and compare relative performance under abrupt and smooth change alternatives in Section 4. Interestingly, the two classes of statistics have the same exact and limiting null distributions, and limiting Bahadur or Pitman efficiency equalling one under abrupt or smooth changepoint alternatives, yet can be distinguished on the basis of power with finite sample sizes. Also in Section 4, we compare both classes to a parametric score-like statistic suggested by Pettitt (1980) and further studied by James, James, and Siegmund (1987). We briefly discuss two-sided versions of the test statistics in Section 5, and conclude with a number of remarks.

The changepoint formulation assumed in this note, namely, that the initial center of symmetry is known, commonly occurs in statistical process control. More particularly, the motivation for this work arose from a series of experiments recently undertaken in the Division of Sleep Disorders of Scripps Clinic to assess the effectiveness of low-energy emission therapy (LEET), a non-pharmacologic approach, in the treatment of persistent psycho-physiologic insomnia (Erman *et al.*, 1990a, 1990b; Pasche *et al.*, 1990). In typical experiments, patients with chronic insomnia were randomly assigned to active treatment or placebo groups; treatment consisted of the intrabuccal emission of low intensity 27 mHz amplitude-modulated electromagnetic fields, given as a 20-minute treatment in the late afternoon three times per week for a four-week study period. Statistically independent assessments of sleep latency were performed on each group at baseline, and nightly over the course of the study. A plausible alternative to lack of efficacy of LEET is a changepoint, whereby LEET would induce an improvement in sleep latency at some time over the course of the study. One means of analysis of the data arising from these experiments is the methodology described in this note, as knowledge of baseline differences between the active and placebo groups may be restructured into knowledge of the initial level of symmetry. See Koziol *et al.* (1993) for further details.

A more generalized formulation of the changepoint problem would lead to tests for parameter changes at unknown times in linear regression models. In the nonparametric context, this has been studied by P. K. Sen (1977, 1978, 1980) and Hušková (1986, 1989) among others. Kim and Siegmund (1989) and MacNeill and colleagues (1991, 1993) also consider this problem, form different perspectives. Indeed, the asymptotic distribution theory advanced in this paper can be readily established through weak convergence arguments (e.g., P. K. Sen) or strong convergence results (Csörgő and Horváth).

2. SIGNED RANK STATISTICS FOR THE CHANGEPOINT PROBLEM

Suppose $X_1, X_2, \ldots, X_n$ is a sequence of independent random variables, with $X_i \sim F(x - \theta_i)$, where $F(\cdot)$ is absolutely continuous with density function f symmetric about 0. The changepoint problem may be formulated as a test of the null hypothesis

$$H_{0n}: \theta_1 = \theta_2 = \cdots \theta_n$$

versus either a one-sided alternative of a single abrupt change,

$$H_{1n}: \theta_1 = \cdots = \theta_k < \theta_{k+1} = \cdots = \theta_n, \tag{2.1}$$

or a one-sided alternative of a smooth change,

$$\begin{aligned} H_{2n}: \theta_1 = \cdots = \theta_{k_1} &< \theta_{k_1} + \Delta < \theta_{k_1} + 2\Delta \\ &< \cdots < \theta_{k_1} + (k_2 - k_1)\Delta = \theta_{k_2} = \cdots = \theta_n, \quad k_1 < k_2, \end{aligned} \tag{2.2}$$

where

$$\Delta = (\theta_{k_2} - \theta_{k_1})/(k_2 - k_1),$$

CH

and the changepoints k, k_1, and
known, and without loss of ge
changepoint hypothesis on the s

V_j

where $\operatorname{sgn}(X_i) = +1$ (-1) if
$|X_1|, |X_2|, \ldots, |X_n|$, and the score
ly continuous, skew-symmetric g

$a_n(i) = \varphi($

Under H_{0n}, we may readily sh

and

$\operatorname{Cov}(V_{in},$

where

$\omega_{an}^2 = n^{-1} \sum_{i=1}^{n}$

Define

$Z_n(t) = (n\omega_{an}^2$

and extend Z_n to a functional Z_N

$Z_N(t) = Z_{[Nt]}$

From Sen (1981, Theorem 5.3.1),
topology on $D[0, 1]$ to the Brow
In a generalization of Pettitt
defined as

for assessing H_{0n}. The limiting nul
and is given by

$Pr\left(\sup_{0 \le t \le 1}\right.$

here, Φ denotes the standard nor

and the changepoints k, k_1, and k_2 are unknown. Suppose that the initial level θ_1 is known, and without loss of generality is equal to zero. We then base tests of the changepoint hypothesis on the signed rank statistics

$$V_{jn} = \sum_{i=1}^{j} \operatorname{sgn}(X_i) a_n(R_i^+),$$

where $\operatorname{sgn}(X_i) = +1 \ (-1)$ if $X_i \geqslant (<) 0$, R_i^+ is the rank of $|X_i|$ among $|X_1|, |X_2|, \ldots, |X_n|$, and the score function a_n is related to a square integrable, absolutely continuous, skew-symmetric generating function φ by

$$a_n(i) = \varphi\left(\frac{1}{2} + \frac{1}{2}\frac{i}{n-1}\right), \quad i = 1, 2, \ldots, n.$$

Under H_{0n}, we may readily show (Hájek and Šidák, 1967) that

$$\mathbf{E}(V_{jn}) = 0,$$

$$\operatorname{Var}(V_{jn}) = \omega_{an}^2 j,$$

and

$$\operatorname{Cov}(V_{in}, V_{jn}) = \operatorname{Var}(V_{in}), \quad 1 \leqslant i \leqslant j \leqslant n,$$

where

$$\omega_{an}^2 = n^{-1} \sum_{i=1}^{n} [a_n(i)]^2 \to \int_0^1 \varphi^2\left(\frac{1}{2} + \frac{1}{2}t\right) \mathrm{d}t = \omega^2.$$

Define

$$Z_n(t) = (n\omega_{an}^2)^{-1/2} V_{jn}, \quad t = j/n, \quad j = 1, \ldots, n,$$

$$Z_n(0) = 0,$$

and extend Z_n to a functional Z_N on $[0, 1]$ by

$$Z_N(t) = Z_{[Nt]} + (Nt - [Nt]) Z_{[Nt]+1}, \quad 0 \leqslant t \leqslant 1.$$

From Sen (1981, Theorem 5.3.1), Z_N converges weakly under $\{H_{0n}\}$ in the Skorohod topology on $D[0, 1]$ to the Brownian motion process Z on $[0, 1]$.

In a generalization of Pettitt (1979), Koziol (1987) introduced the statistic L_n^+, defined as

$$L_n^+ = \max_{1 \leqslant j \leqslant n} V_{jn},$$

for assessing H_{0n}. The limiting null distribution of $(n\omega_{an}^2)^{-1/2} L_n^+$ is that of $\sup_{0 \leqslant t \leqslant 1} Z(t)$, and is given by

$$Pr\left(\sup_{0 \leqslant t \leqslant 1} Z(t) \leqslant c\right) = 2\Phi(c) - 1, \quad c \geqslant 0;$$

here, Φ denotes the standard normal cumulative distribution function.

Let us next introduce a different test statistic, denoted M_n^+, for assessing H_{0n}, as follows:

$$M_n^+ = \max_{1\leq j\leq n} (V_{nn} - V_{jn}) = V_{nn} - \min_{1\leq j\leq n} V_{jn}.$$

The motivation for M_n^+ should be clear: under H_{0n}, the rank process Z_N determined by the V_{jn} is centered about 0, whereas, under either alternative H_{1n} or H_{2n}, the rank process will begin to drift positively subsequent to the first changepoint. Moreover, M_n^+ should give a better indication of the location of the changepoint than L_n^+.

Under the null hypothesis, note that L_n^+ and M_n^+ have the same null distribution. This follows from their distributional invariance with respect to reversal of the time order of the data under H_{0n}.

From the continuous mapping theorem (Billingsley, 1968), the limiting null distribution of $(n\omega_{an}^2)^{-1/2} M_n^+$ is that of $Z(1) - \inf_{0\leq t\leq 1} Z(t)$, and is given in the following lemma.

Lemma.

$$Pr\left\{Z(1) - \inf_{0\leq t\leq 1} Z(t) \leq c\right\} = 2\Phi(c) - 1, \quad c \geq 0.$$

Proof. This lemma is well-known, and may be proved by a number of techniques. Koziol (1991) gives an elementary proof based upon first principles from Billingskey (1968, Section 11). Alternatively, note that $Z(1) - Z(t)$ has the same distribution as $Z(1-t)$, and so the maxima of $Z(1) - Z(t)$ and $Z(t)$ have the same distribution.

L_n^+ and M_n^+ thus have the same limiting null distributions. We turn to comparison of the operating characteristics of these two statistics in the following sections.

3. FINITE SAMPLE NULL DISTRIBUTION RESULTS

Note that the null distributions of L_n^+ and M_n^+ can be determined exactly by enumeration of all possible configurations of signs and signed ranks; however, given $2^n \cdot n!$ such configurations with sample size n, the cost of this approach becomes prohibitive as n increases. We thus carried out a simulation experiment in order to determine how large the sample size must be before the asymptotic critical values provide reasonable approximations.

The simulation experiment, conducted on a VAX 11/750, consisted of drawing 10000 random samples of size n, for $n = 10(10)200$, from the uniform $(-1, 1)$ distribution using IMSL subroutine GGUBT, then finding the 50th, 75th, 90th, 95th and 99th percentiles of the empirical distributions of L_n^+ and M_n^+. We used both Wilcoxon scores, with generating function $\varphi(u) = 2u - 1$, and quantile normal scores, with generating function $\varphi(u) = \Phi^{-1}(i/n+1)$ (van Eeden, 1963). The empirical percentiles of M_n^+ are depicted in Figure 1. The related results for L_n^+ are quite similar, as one might expect, hence are not shown.

Clearly, even for reasonably large sample sizes, the use of the asymptotic critical values will result in tests that are somewhat conservative. As in Joe, Koziol, and Petkau (1981), we therefore fit equations of the form $\mathbf{E}_{n,p} = A_p - B_p \exp(-C_p n^{1/2})$ using

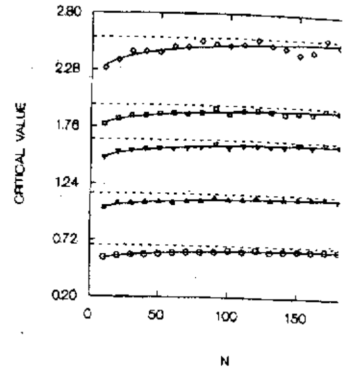


Figure 1 Estimated critical values for M
scores (Panel B). The dashed lines ind
empirical values, as explained in the tex

Table 1 Coefficients for finite sample cr

Statistic	Percentile	S
L_n^+	.50	W N
	.75	W N
	.90	W N
	.95	W N
	.99	W N
M_n^+	.50	W N
	.75	W N
	.90	W N
	.95	W N
	.99	W N

ited M_n^+, for assessing H_{0n}, as

min V_{jn}.
$\leqslant j \leqslant n$

rank process Z_N determined by
ernative H_{1n} or H_{2n}, the rank
e first changepoint. Moreover,
he changepoint than L_n^+.
ave the same null distribution.
respect to reversal of the time

1968), the limiting null distribu-
and is given in the following

$-1, \quad c \geqslant 0.$

ed by a number of techniques.
irst principles from Billingskey
) has the same distribution as
ive the same distribution.
tions. We turn to comparison of
he following sections.

ULTS

an be determined exactly by
signed ranks; however, given 2^n.
approach becomes prohibitive
nent in order to determine how
tical values provide reasonable

750, consisted of drawing 10000
iform $(-1, 1)$ distribution using
, 90th, 95th and 99th percentiles
ed both Wilcoxon scores, with
al scores, with generating func-
pirical percentiles of M_n^+ are
ite similar, as one might expect,

use of the asymptotic critical
e. As in Joe, Koziol, and Petkau
$= A_p - B_p \exp(-C_p n^{1/2})$ using

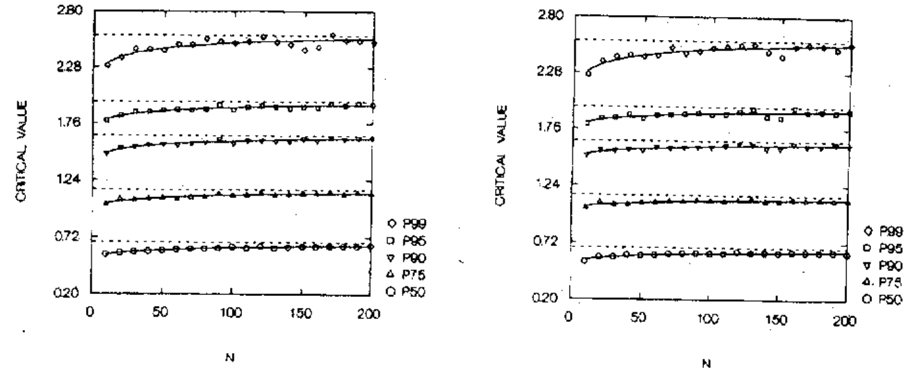


Figure 1 Estimated critical values for M_n^+ under H_{0n}, using Wilcoxon scores (Panel A) or quantile normal scores (Panel B). The dashed lines indicate the asymptotic critical values; the solid lines are smoothed empirical values, as explained in the text.

Table 1 Coefficients for finite sample critical value smoothing equations $E_{n,p} = A_p - B_p \exp(-C_p N^{1/2})$

Statistic	Percentile	Scores*	A_p	B_p	C_p
L_n^+	.50	W	.6745	.1648	.1174
		N		.1737	.1174
	.75	W	1.1503	.1994	.1454
		N		.1617	.1132
	.90	W	1.6445	.2578	.1662
		N		.1982	.1279
	.95	W	1.9600	.2993	.1807
		N		.2255	.1342
	.99	W	2.5758	.4272	.1679
		N		.4560	.1708
M_n^+	.50	W	.6745	.1676	.1190
		N		.1610	.1116
	.75	W	1.1503	.1898	.1354
		N		.1390	.1007
	.90	W	1.6445	.2603	.1622
		N		.1754	.1212
	.95	W	1.9600	.2670	.1550
		N		.2155	.1273
	.99	W	2.5758	.4691	.1924
		N		.5096	.1897

nonlinear least squares to the empirical critical values $\mathbf{E}_{n,p}$ approximating the asymptotic critical values A_p for each percentile p, in order to smooth the simulation results and to provide an interpolation formula for other sample sizes. [One might alternatively use Siegmund's (1985) approximation $\mathbf{E}_{n,p} = A_p - D_p/n^{1/2}$. However, since the residual sums of squares from the Joe-Koziol-Petkau approximation were typically smaller than the corresponding residual sums of squares from the Seigmund approximation {except at $p = 0.99$}, only the JKP approximation is reported herein]. For reference, these parameter values are given in Table 1.

4. POWER COMPARISONS

4.1. Asymptotics

Our basis for asymptotic comparison of L_n^+ and M_n^+ is the following theorem, which establishes the nature of the drift of the limiting Brownian motion process under a sequence of contiguous alternatives.

Theorem. *Let X_i, $i = 1, \ldots, n$, be independent random variables with continuous distributions $F(x - \theta_i)$. Assume that F possesses almost everywhere an absolutely continuous density function f which is symmetric about 0, and with finite Fisher information $I(f)$, given by*

$$I(f) = \int_0^1 \varphi_0^2(t, f)\mathrm{d}t,$$

where

$$\varphi_0(t, f) = \frac{-f'[F^{-1}(t)]}{f[F^{-1}(t)]}, \quad 0 < t < 1.$$

Consider the sequence of local alternatives $\{H_{1n}\}$ and $\{H_{2n}\}$, where $\theta_n = n^{-1/2}\theta$, $\theta > 0$, and $k/n \to \tau$, $0 < \tau < 1$ under $\{H_{1n}\}$, $k_1/n \to \tau_1$, $k_2/n \to \tau_2$, $0 < \tau_1 < \tau_2 < 1$ under $\{H_{2n}\}$, as $n \to \infty$. Then under $\{H_{in}\}$,

$$Z_N(t) \xrightarrow{D} Z(t) + \gamma_i(t), \quad t \in [0, 1], \quad i = 1, 2,$$

where

$$\gamma_1(t) = \begin{cases} 0 & t \leqslant \tau \\ \dfrac{(t-\tau)\theta}{\omega}\displaystyle\int_0^1 \varphi(u)\varphi_0(u, f)\mathrm{d}u, & \tau < t \leqslant 1 \end{cases} \tag{4.1}$$

and

$$\gamma_2(t) = \begin{cases} 0 & t \leqslant \tau_1 \\ \dfrac{(t-\tau_1)^2\theta}{2(\tau_2-\tau_1)\omega}\displaystyle\int_0^1 \varphi(u)\varphi_0(u, f)\mathrm{d}u, & \tau_1 < t \leqslant \tau_2 \\ \left[\dfrac{(\tau_2-\tau_1)}{2} + (t-\tau_2)\right]\dfrac{\theta}{\omega}\displaystyle\int_0^1 \varphi(u)\varphi_0(u, f)\mathrm{d}u, & \tau_2 < t \leqslant 1. \end{cases} \tag{4.2}$$

Proof. The proof is a slight r
If we denote the joint distril
that $P_n^{(0)}$ refers to the null case
that $\{P_n^{(\theta)}\}$ is contiguous to $\{P$
from Sen (1981, Theorem 5.3
from Sen (1981, Theorem 4.3,
distributions of $\{Z_N\}$ under
Theorem 5.3.3), and Hájek ar
(1970) for a more general
weak convergence under a wide
herein.

We may compare the asymp
classes of alternatives by using a
is straightforward to establish
in Bahadur's terminology: The
maximal drift of $\gamma_i(t)$ under $\{H_{in}$
have the same approximate slop
emphasized that the most impo
vicinity of a null hypothesis. In
(1976) condition, so that the lin
namely one, equals its limiting P
the two statistics on the basis c
their performances with reason
would be similar.

[In related work, Praagman
of max- and sum-type linear
problem. Of particular note, he
a max-type statistic that is at lea
changepoint alternatives. We ref
details.]

B. Finite samples

We conducted a limited simul
of L_n^+ and M_n^+ with finite sam
(2.1), with $X_i \sim N(\theta_i, 1)$ θ_n
$n = 10(10)200$, and $\tau = k/n = 0.1$(
scores with both L_n^+ and M_n^+.
statistic,

where $S_j = X_i + \cdots + X_j$. James,
a score-like test statistic, and I
compared with the (modified) li

es $\mathbf{E}_{n,p}$ approximating the asym-
to smooth the simulation results
iple sizes. [One might alternative-
$, -D_p/n^{1/2}$. However, since the
iu approximation were typically
ares from the Seigmund approxi-
mation is reported herein]. For
1.

is the following theorem, which
Brownian motion process under

variables with continuous distribu-
ywhere an absolutely continuous
th finite Fisher information $I(f)$,

$< t < 1$.

$\{H_{2n}\}$, where $\theta_n = n^{-1/2}\theta$, $\theta > 0$,
$_2, 0 < \tau_1 < \tau_2 < 1$ under $\{H_{2n}\}$, as

$i = 1, 2,$

$t \leqslant \tau$

$u, \quad \tau < t \leqslant 1$ (4.1)

$t \leqslant \tau_1$

$du, \quad \tau_1 < t \leqslant \tau_2$ (4.2)

$(u, f)du, \quad \tau_2 < t \leqslant 1.$

Proof. The proof is a slight modification to that of Theorem 5.3.3 of Sen (1981). If we denote the joint distribution of $(X_1, \ldots, X_n)$ under H_{in} as $P_n^{(i)}$, $i = 1, 2$, so that $P_n^{(0)}$ refers to the null case, then it follows from Hájek and Šidák (1967, VI.1.3) that $\{P_n^{(i)}\}$ is contiguous to $\{P_n^{(0)}\}$. The tightness of $\{Z_N\}$ under $\{P_n^{(0)}\}$ is established from Sen (1981, Theorem 5.3.1), and the tightness of $\{Z_N\}$ under $\{P_n^{(i)}\}$ follows from Sen (1981, Theorem 4.3.4). Lastly, the convergence of the finite-dimensional distributions of $\{Z_N\}$ under $\{P_n^{(i)}\}$ to those of $\{Z + \gamma_i\}$ follows from Sen (1981, Theorem 5.3.3), and Hájek and Šidák (1967, Theorem VI.2.5). See also Hušková (1970) for a more general representation, which may be used to establish weak convergence under a wider class of contiguous alternatives than that considered herein.

We may compare the asymptotic performances of L_n^+ and M_n^+ against particular classes of alternatives by using a criterion of efficiency introduced by Bahadur (1960). It is straightforward to establish that both $\{L_n^+\}$ and $\{M_n^+\}$ are standard sequences in Bahadur's terminology: They have the same limiting null distribution, and the maximal drift of $\gamma_i(t)$ under $\{H_{in}\}$ is realized at $\gamma_i(1)$. Hence, $\{L_n^+\}$ and $\{M_n^+\}$ in general have the same approximate slopes under $\{H_{1n}\}$ and $\{H_{2n}\}$. Bahadur (1967) has further emphasized that the most important property of a slope is its value in the immediate vicinity of a null hypothesis. In this regard, both $\{L_n^+\}$ and $\{M_n^+\}$ satisfy Wieand's (1976) condition, so that the limiting Bahadur efficiency of $\{L_n^+\}$ relative to $\{M_n^+\}$, namely one, equals its limiting Pitman efficiency. Hence, we cannot distinguish between the two statistics on the basis of this asymptotic comparison, and might expect that their performances with reasonable sample sizes typically encountered in practice would be similar.

[In related work, Praagman (1988) investigated the exact Bahadur efficiencies of max- and sum-type linear rank statistics for the two-sample changepoint problem. Of particular note, he established that for each sum-type statistic, there is a max-type statistic that is at least as efficient in the Bahadur sense, uniformly over all changepoint alternatives. We refer the interested reader to Praagman (1988) for further details.]

B. *Finite samples*

We conducted a limited simulation study to compare the relative performances of L_n^+ and M_n^+ with finite sample sizes. We chose the abrupt changepoint setting (2.1), with $X_i \sim N(\theta_i, 1)$ $\theta_n = \theta =$ magnitude of location shift $= 0.1(0.1)2.0$, $n = 10(10)200$, and $\tau = k/n = 0.1(0.1)0.9$. We used both Wilcoxon and quantile normal scores with both L_n^+ and M_n^+. For further comparison, we included Pettitt's (1980) statistic,

$$P_n^+ = \max_{1 \leqslant k \leqslant n} \left(\frac{kS_n}{n - S_k} \right),$$

where $S_j = X_1 + \cdots + X_j$. James, James, and Siegmund (1987) have shown that P_n^+ is a score-like test statistic, and has reasonable power in assessing H_{0n} versus H_{1n}, compared with the (modified) likelihood ratio tests investigated therein. [Note that

Pettitt's procedure does not require knowledge of the initial center of symmetry in its formulation. Nevertheless, it is not necessarily true that procedures which utilize the most prior information will outperform those that do not. See Wells (1990) for further discussion.]

For each choice of n, we determined approximate $\alpha = 0.05$ level critical values for L_n^+ and M_n^+ from the smoothing equations in Table 1. We similarly used the James, James, and Siemund (1987) approximation for the exceedance probability of P_n^+, namely,

$$Pr\{P_n^+ \geqslant b\} \doteq \exp\{-2n^{-1}(b+\rho)^2\}, \quad \rho \doteq 0.583,$$

for its approximate $\alpha = 0.05$ level critical values. Then, for each set of n, θ, and τ, we used IMSL routine GGNML on a VAX 11/750 to generate 1000 random samples of size n from the appropriate abrupt changepoint model, and thereafter determined the empirical powers of L_n^+, M_n^+, and P_n^+ against that particular alternative. We briefly summarize our findings here.

From (4.1), the drift $\gamma_1(t)$ of the limiting Brownian motion process $Z(t)$ under the abrupt changepoint model increases with θ and $(1-\tau)$. We utilize this fact to amalgamate some of the simulation results in Figures 2–4, where we depict the empirical powers of L_n^+ and M_n^+ with uniform scores (LU, MU) or normal scores (LN, MN) corresponding to pooled combinations of $\theta(1-\tau)$, for sample sizes $n = 20$, 50, 150. Results for $n = 100$ and 200 were qualitatively quite similar, hence are omitted.

As expected, power does increase with $\theta(1-\tau)$, and n. One might also anticipate from (4.1) that the normal score versions of the test statistics should outperform the uniform score versions, albeit slightly. [Recall from van Eeden (1963) that the efficiency of the Wilcoxon signed rank statistic relative to van der Waerden's normal score statistic in the one sample setting with normally distributed data is $3/\pi$.] This is indeed the case, as may be seen in Figures 2–4. In each of these figures, approximating curves or smoothers (Cleveland, 1981) were fit separately to the normal scores and uniform scores empirical powers; from these curves, the normal scores test always outperformed its uniform counterpart, though the difference was slight for small n. Lastly, one might infer from comparison of the A and B panels of each figure that M_n^+ seems somewhat more powerful than L_n^+ in this hypothesis testing setting. Let us address this issue in a slightly different manner.

In Figures 5 and 6, we depict the empirical powers of L_n^+ and M_n^+, both using normal scores, and P_n^+, under the abrupt changepoint setting described earlier, for representative choice of θ, τ, and n. One might conclude from these figures that M_n^+ is indeed preferable to L_n^+ in this setting, with the power of L_n^+ evidently dropping off more rapidly than that of M_n^+ as τ increases. As James, James and Siegmund (1987) had pointed out, the power function of P_n^+ is not monotone in τ, but is maximal at about $\tau = 0.5$; in contrast, the powers of L_n^+ and M_n^+ monotonically decrease with τ. Nevertheless, L_n^+ seems to dominate P_n^+ in terms of power up to about $\tau = 0.5$, and M_n^+ dominates P_n^+ up to about $\tau = 0.6$ or 0.7. In an absolute sense, none of the statistics demonstrates reasonable power in detecting late changepoints ($\tau = 0.8$). If earlier changepoints are anticipated, M_n^+ would be the statistic of choice.

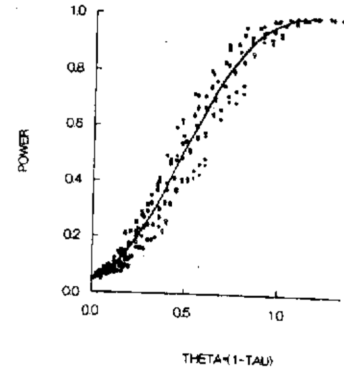


Figure 2 Empirical powers of L_n^+ (Pan
alpha level .05 with $n = 20$. The empirica
□, and using normal scores by ○. See t
The solid lines are smoothers of the emp
tension = 0.5. Throughout Figures 2–4
lower smoother to the Wilcoxon score

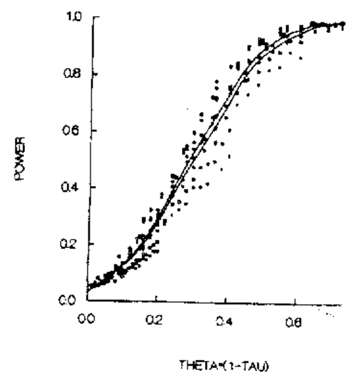


Figure 3 Observed and smoothed empi
$\alpha = 0.05$, $n = 50$, using Wilcoxon scores (

5. TWO-SIDED ALTERNATI

Suppose the direction θ of the
changepoint alternatives is unk
introduced for these alternatives

$$L_n = \max_{1 \leqslant j \leqslant n}$$

$$M_n = \max\Big[$$

: initial center of symmetry in its
hat procedures which utilize the
not. See Wells (1990) for further

$\alpha = 0.05$ level critical values for
1. We similarly used the James,
exceedance probability of P_n^+,

}, $\rho \doteq 0.583$,

for each set of n, θ, and τ, we used
ite 1000 random samples of size
, and thereafter determined the
particular alternative. We briefly

n motion process $Z(t)$ under the
t). We utilize this fact to amalga-
, where we depict the empirical
U) or normal scores (LN, MN)
for sample sizes $n = 20$, 50, 150.
similar, hence are omitted.
nd n. One might also anticipate
statistics should outperform the
n Eeden (1963) that the efficiency
in der Waerden's normal score
buted data is $3/\pi$.] This is indeed
se figures, approximating curves
the normal scores and uniform
l scores test always outperformed
ght for small n. Lastly, one might
figure that M_n^+ seems somewhat
ting. Let us address this issue in

ers of L_n^+ and M_n^+, both using
point setting described earlier,
ht conclude from these figures
ig, with the power of L_n^+ evi-
as τ increases. As James, James
unction of P_n^+ is not monotone
owers of L_n^+ and M_n^+ monotoni-
ominate P_n^+ in terms of power
out $\tau = 0.6$ or 0.7. In an absolute
power in detecting late change-
ed, M_n^+ would be the statistic of

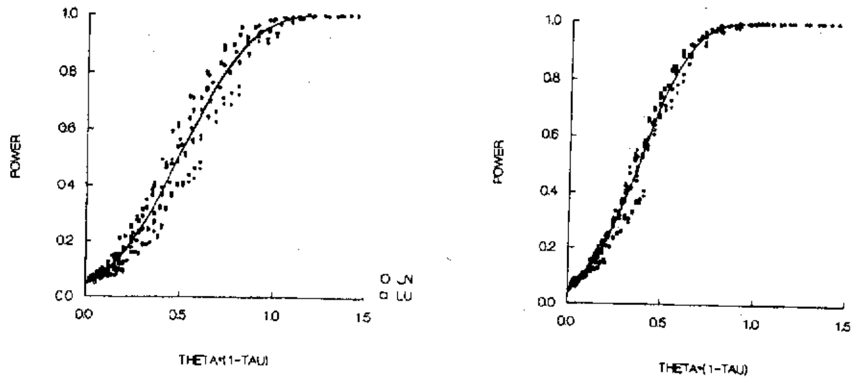


Figure 2 Empirical powers of L_n^+ (Panel A) and M_n^+ (Panel B) under the abrupt changepoint model (2.1) at alpha level .05 with $n = 20$. The empirical power values of the statistics using Wilcoxon scores are denoted by □, and using normal scores by ○. See the text for the combinations of θ and τ used in the simulation study. The solid lines are smoothers of the empirical power values, using Cleveland's (1981) LOWESS method, with tension = 0.5. Throughout Figures 2–4, the upper smoother corresponds to the normal score values, the lower smoother to the Wilcoxon score values.

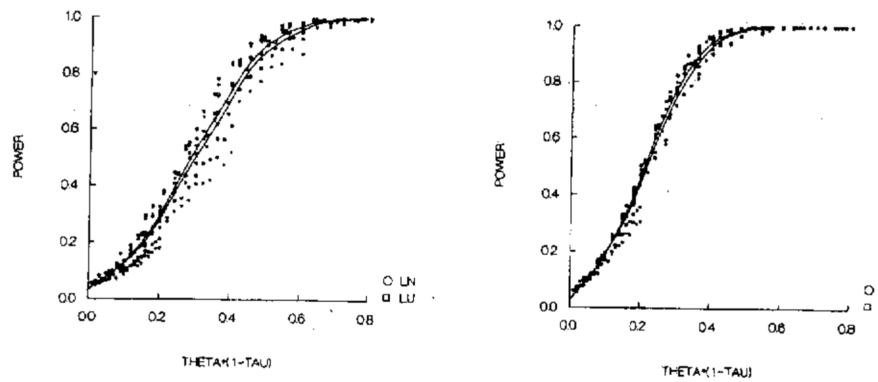


Figure 3 Observed and smoothed empirical powers of L_n^+ (Panel A) and M_n^+ (Panel B) under H_{1a}, with $\alpha = 0.05$, $n = 50$, using Wilcoxon scores (□) or normal scores (○).

5. TWO-SIDED ALTERNATIVES

Suppose the direction θ of the location shift in the abrupt changepoint or smooth changepoint alternatives is unknown. Two-sided versions of L_n^+ and M_n^+ can be introduced for these alternatives, as follows:

$$L_n = \max_{1 \leqslant j \leqslant n} |V_{jn}|,$$

$$M_n = \max\left[V_{nn} - \min_{1 \leqslant j \leqslant n} V_{jn},\ \max_{1 \leqslant j \leqslant n} V_{jn} - V_{nn}\right]$$

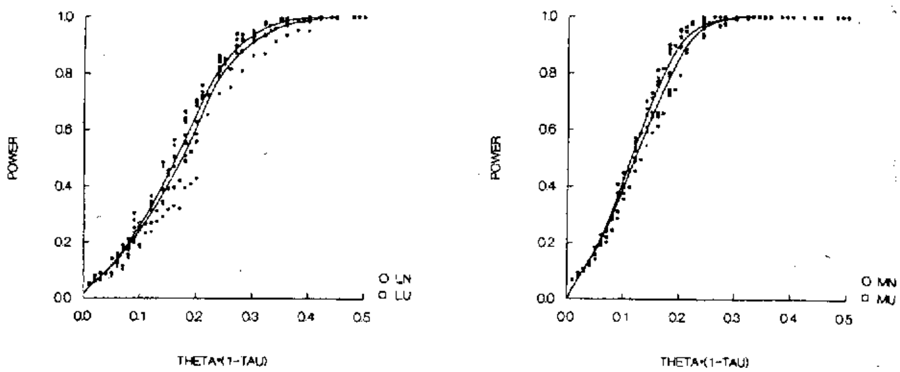


Figure 4 Observed and smoothed empirical powers of L_n^+ (Panel A) and M_n^+ (Panel B) under H_{1n}, with $\alpha = 0.05$, $n = 150$, using Wilcoxon scores (□) or normal scores (○).

The limiting null distribution of L_n is well known, being that of $\sup_{0 \leq t \leq 1} |Z(t)|$, and given by (Billingsley [1968], 11.13)

$$Pr\left\{\sup_{0 \leq t \leq 1} |Z(t)| \leq c\right\} = \sum_{k=-\infty}^{\infty} (-1)^k \{\Phi[(2k+1)c] - \Phi[(2k-1)c]\}, \quad c \geq 0.$$

The limiting null distribution of M_n is that of max $[Z(1) - \inf_{0 \leq t \leq 1} Z(t), \sup_{0 \leq t \leq 1} Z(t) - Z(1)]$, and is identical to that of lim L_n, as shown in the following lemma.

Lemma.

$$Pr\left\{\max[Z(1) - \inf_{0 \leq t \leq 1} Z(t), \sup_{0 \leq t \leq 1} Z(t) - Z(1)] \leq c\right\}$$
$$= \sum_{k=-\infty}^{\infty} (-1)^k \{\Phi[(2k+1)c] - \Phi[(2k-1)c]\}, \quad c > 0$$

Again, this lemma may be proved by a variety of techniques. Koziol (1991) provides an elementary proof based on first principles from Billingsley (1968). Alternately, we may argue as follows:
Set

$$A(t) = Z(t) - Z(1).$$

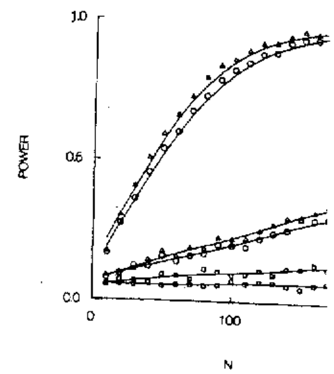
Then

$$\max[\sup A(t), -\inf A(t)] = \max[\sup A(t), \sup -A(t)]$$
$$= \sup \max[A(t), -A(t)] = \sup|A(t)|,$$

which yields the lemma directly.

Analogously, we would further anticipate that the relative merits of L_n^+ and M_n^+ in assessing the one-sided changepoint models (2.1) and (2.2) should similarly pertain to L_n and M_n in the two-sided setting.

C



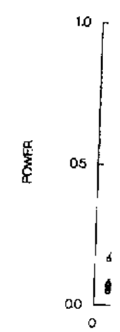
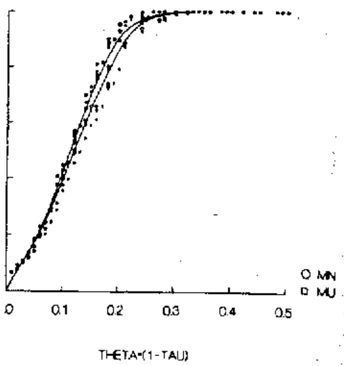


Figure 5 Empirical powers of L_n^+ (○ (2.1) at alpha level 0.05. In panel A, τ in Panel C, τ = 0.9, and θ = 0.5 or 1.5. T Cleveland's (1981) LOWESS method, w upper smoother for each statistic corres value.

FINAL REMARKS

Although we have not investigat against the smooth change altern would be quantitatively similar magnitude of the drift would be le (2.2) compared with (2.1), the po numerical calculations would be
In our limited simulation study, P_n^+ were similar to those reportec strength of L_n^- and M_n^+ is their e likelihood ratio tests of James, information concerning the expe



A) and M_n^+ (Panel B) under H_{1n}, with

:ing that of $\sup_{0 \le t \le 1} |Z(t)|$, and

$] - \Phi[(2k-1)c]\}, \quad c \geq 0.$

$[Z(1) - \inf_{0 \le t \le 1} Z(t), \sup_{0 \le t \le 1}$
wn in the following lemma.

$Z(1)] \le c \Big\}$

$-1)c]\}, \quad c > 0$

iques. Koziol (1991) provides an
;sley (1968). Alternately, we may

)]

$\sup |A(t)|,$

elative merits of L_n^+ and M_n^+ in
(2.2) should similarly pertain to

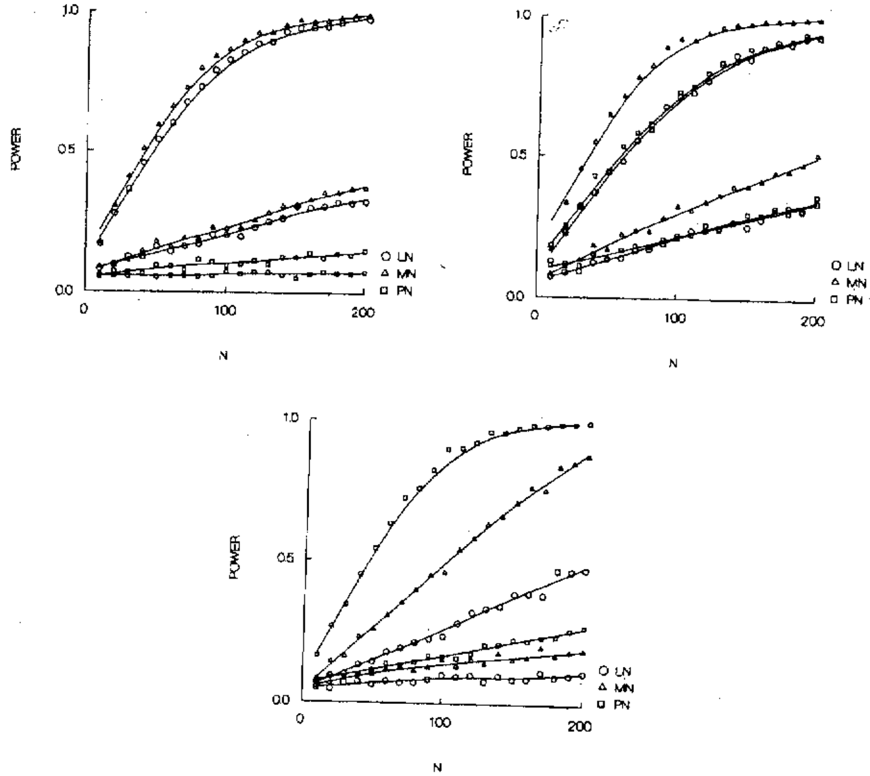


Figure 5 Empirical powers of L_n^+ (○), M_n^+ (△), and P_n^+(□) under the abrupt changepoint model (2.1) at alpha level 0.05. In panel A, $\tau = 0.1$, and $\theta = 0.1$ or 0.3; in Panel B, $\tau = 0.5$, and $\theta = 0.2$ or 0.5; in Panel C, $\tau = 0.9$, and $\theta = 0.5$ or 1.5. The solid lines are smoothers of the empirical power values, using Cleveland's (1981) LOWESS method, with tension = .5. Power increases with θ, so that in each panel, the upper smoother for each statistic corresponds to the larger value of θ, the lower smoother to the smaller value.

FINAL REMARKS

Although we have not investigated the relative performances of the signed rank tests against the smooth change alternative (2.2), we nevertheless anticipate that our findings would be quantitatively similar to those reported in Section 4. However, since the magnitude of the drift would be less for fixed terminal θ and initial shift position τ under (2.2) compared with (2.1), the powers of L_n^+, M_n^+, and P_n^+ would accordingly decline; numerical calculations would be useful for gauging the relative differences.

In our limited simulation study, M_n^+ appeared to dominate L_n^+; the characteristics of P_n^+ were similar to those reported by James, James, and Siegmund (1987). A particular strength of L_n^+ and M_n^+ is their enhanced power for τ near 0, where even the modified likelihood ratio tests of James, James, and Siegmund (1987) are weak. If *a priori* information concerning the expected position of the changepoint is available, it can

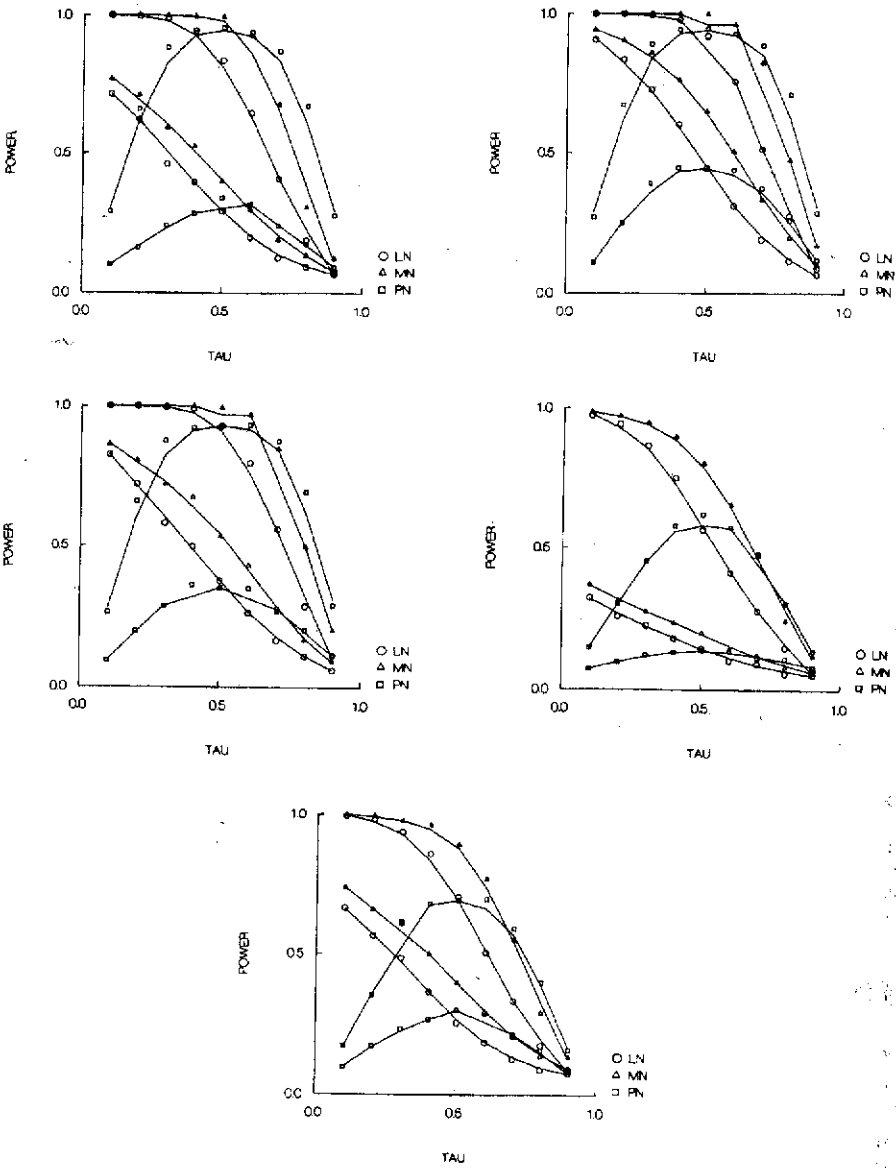
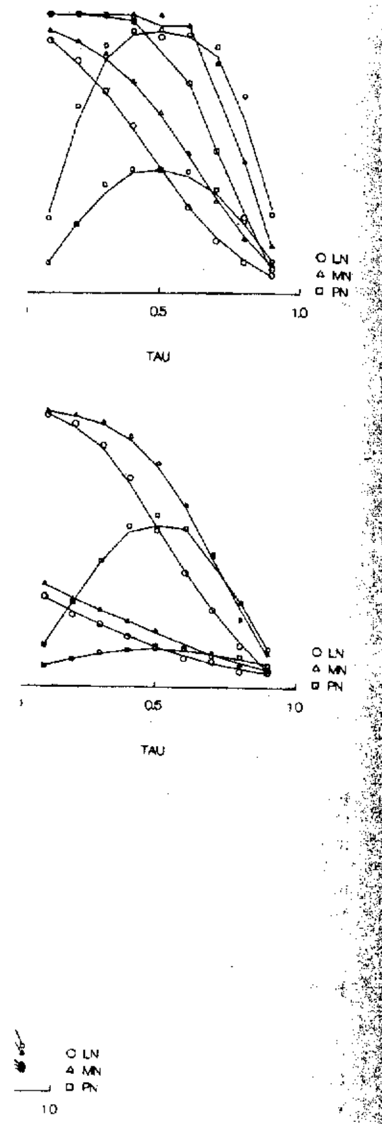


Figure 6 Empirical powers of L_n^+ (○), M_n^+ (△), and P_n^+(□) under the abrupt changepoint model (2.1) at alpha level 0.05. In panel A, $n = 0.20$, and $\theta = 0.6$ or 1.6; in Panel B, $n = 50$, and $\theta = 0.5$ or 1.0; in Panel C, $n = 100$, and $\theta = 0.3$ or 0.7; in Panel D, $n = 150$, and $\theta = 0.2$ or 0.4; in Panel E, $n = 200$, and $\theta = 0.1$ or 0.3. The solid lines are smoothers of the empirical power values, using Cleveland's (1981) LOWESS method, with tension = 0.5. Power increases with θ, so that in each panel, the upper smoother for each statistic corresponds to the larger value of θ, the lower smoother to the smaller value.

profitably be utilized for sel
A further advantage of M_n^+
confidence sets for the chang
1988), and James, James, and

References

[1] Bahadur R. R. (1960). Stochasti
[2] Bahadur R. R. (1967). Rates of
303–324.
[3] Billingsley, P. (1968). *Convergen*
[4] Cleveland, W. S. (1981). LOWE
regression. *American Statisticia*
[5] Csörgő, M. and Horváth, L. (198
Inf. **17**, 1–9.
[6] Csörgő, M. and Horváth, L. (
Krishnaiah and CR Rao, eds. *H*
[7] Erman, M., Hajdukovic, R., Coh
and objective sleep variables. *Sl*
[8] Erman, M., Hajdukovic, R., Coh
LEET in the treatment of insom
[9] Hájek, J. and Šidák, Z. (1967). *T*
[10] Hušková, M. (1970). Asymptoti
Z. Wahrscheinlichkeitstheorie ver
[11] Hušková, M. (1988). Adaptive p
[12] Hušková, M. (1989). Nonparame
In: P Hackl, ed. *Analysis and F*
71–85.
[13] James, B. James, K. L. and Siegn
[14] James, B. James, K. L. and Sieg
applications to changepoint prol
[15] Jandhyala, V. K. and MacNeill,
regression models. *J. Statist. Pla*
[16] Joe, H., Koziol, J. A. and Petka
survival distributions. *Biometrics*
[17] Kim, H. J. and Siegmund, D. (1
regression. *Biometrika* **76**, 409–4
[18] Koziol, J. A. (1987). A note on no
[19] Koziol, J. A. (1991). A note on si
report.
[20] Koziol, J. A., Erman, M., Pasche,
in a sequence of repeated measure
J. Appl. Statist. **20**, 393–400.
[21] Lombard, F. (1987). Rank tests f
[22] Pasche, B., Erman, M. and Mitler
323, 486–487.
[23] Pettitt, A. N. (1979). A non-parar
135.
[24] Pettitt, A. N. (1980). A simple cum
observations. *Biometrika* **67**, 79–
[25] Pragman. K. (1988). Bahadur eff
198–217.
[26] Sen, P. K. (1977). Tied-down Wi
some applications. *Ann. Statist.* **5**
[27] Sen, P. K. (1978). Invariance pri
215–236.
[28] Sen, P. K. (1980). Asymptotic th
occurring at an unknown time pc
[29] Sen, P. K. (1981). *Sequential Non*



r the abrupt changepoint model (2.1) at
;, $n = 50$, and $\theta = 0.5$ or 1.0; in Panel C,
Panel E, $n = 200$, and $\theta = 0.1$ or 0.3. The
veland's (1981) LOWESS method, with
·smoother for each statistic corresponds

profitably be utilized for selection of a test statistic sensitive to that alternative. A further advantage of M_n^+ and P_n^+ is that they provide natural estimates and confidence sets for the changepoint location. See Worsley (1986), Siegmund (1986, 1988), and James, James, and Siegmund (1988) for further discussion.

References

[1] Bahadur R. R. (1960). Stochastic comparison of tests. *Ann. Math. Statist.* **31**, 276–295.

[2] Bahadur R. R. (1967). Rates of convergence of estimates and test statistics. *Ann. Math. Statist.* **38**, 303–324.

[3] Billingsley, P. (1968). *Convergence of Probability Measures.* John Wiley, New York.

[4] Cleveland, W. S. (1981). LOWESS: A program for smoothing scatterplots by robust locally weighted regression. *American Statistician* **35**, 54.

[5] Csörgő, M. and Horváth, L. (1987). Nonparametric tests for the changepoint problem. *J. Statist. Plann. Inf.* **17**, 1–9.

[6] Csörgő, M. and Horváth, L. (1988). Nonparametric methods for changepoint problems. In: PR Krishnaiah and CR Rao, eds. *Handbook of Statistics*, North-Holland, Amsterdam, **7**, 403–425.

[7] Erman, M., Hajdukovic, R., Cohen, R., Pasche, B. and Mitler, M. (1990a). A comparison of subjective and objective sleep variables. *Sleep Res.* **19**, 220.

[8] Erman, M., Hajdukovic, R., Cohen, R., Pasche, B., Rossel, C. and Mitler, M. (1990b). Effectiveness of LEET in the treatment of insomnia. *Sleep Res.* **19**, 221.

[9] Hájek, J. and Šidák, Z. (1967). *Theory of Rank Tests.* Academic Press, New York.

[10] Hušková, M. (1970). Asymptotic distribution of simple linear rank statistics for testing symmetry. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **14**, 308–322.

[11] Hušková, M. (1988). Adaptive procedures for detection of change. *Statistics and Decisions* **6**, 137–148.

[12] Hušková, M. (1989). Nonparametric tests for shift and change in regression at an unknown time point. In: P Hackl, ed. *Analysis and Forecasting of Economic Structural Change.* Springer-Verlag, Berlin, 71–85.

[13] James, B. James, K. L. and Siegmund, D. (1987). Tests for a change-point. *Biometrika* **74**, 71–83.

[14] James, B. James, K. L. and Siegmund, D. (1988). Conditional boundary crossing probabilities with applications to changepoint problems. *Ann. Prob.* **16**, 825–839.

[15] Jandhyala, V. K. and MacNeill, I. B. (1991). Tests for parameter changes at unknown times in linear regression models. *J. Statist. Plann. Inf.* **27**, 291–316.

[16] Joe, H., Koziol, J. A. and Petkau, A. J. (1981). Comparison of procedures for testing the equality of survival distributions. *Biometrics* **37**, 327–340.

[17] Kim, H. J. and Siegmund, D. (1989). The likelihood ratio test for a change-point in simple linear regression. *Biometrika* **76**, 409–423.

[18] Koziol, J. A. (1987). A note on non-parametric tests for the changepoint problem. *Biom. J.* **29**, 791–794.

[19] Koziol, J. A. (1991). A note on signed rank tests for the changepoint problem. Unpublished technical report.

[20] Koziol, J. A., Erman, M., Pasche, B., Hajdukovic, R. and Mitler, M. E. (1993). Assessing a changepoint in a sequence of repeated measurements with application to a low-energy emission therapy sleep study. *J. Appl. Statist.* **20**, 393–400.

[21] Lombard, F. (1987). Rank tests for changepoint problems. *Biometrika* **74**, 615–624.

[22] Pasche, B., Erman, M. and Mitler, M. (1990). Diagnosis and management of insomnia. *N. Engl. J. Med.* **323**, 486–487.

[23] Pettitt, A. N. (1979). A non-parametric approach to the change-point problem. *Appl. Statist.* **28**, 126–135.

[24] Pettitt, A. N. (1980). A simple cumulative sum type statistic for the change-point problem with zero-one observations. *Biometrika* **67**, 79–84.

[25] Pragman. K. (1988). Bahadur efficiency of rank tests for the change-point problem. *Ann. Statist.* **16**, 198–217.

[26] Sen, P. K. (1977). Tied-down Wiener process approximations for aligned rank order processes and some applications. *Ann. Statist.* **5**, 1107–1123.

[27] Sen, P. K. (1978). Invariance principles for linear rank statistics revistics revisited. *Sankhya. A.* **40**, 215–236.

[28] Sen, P. K. (1980). Asymptotic theory of some tests for a possible change in the regression slope occurring at an unknown time point. *Z. Wahrschein-lichkeitstheorie verw. Geb.* **52**, 203–218.

[29] Sen, P. K. (1981). *Sequential Nonparametrics.* John Wiley, New York.

[30] Shaban, S. A. (1980). Changepoint problem and two-phase regression: An annotated bibliography. *Int. Statist. Rev.* **48**, 83–93.

[31] Siegmund, D. (1985). *Sequential Analysis: Tests and Confidence Intervals.* Springer-Verlag, New York.

[32] Siegmund, D. (1986). Boundary crossing probabilities and statistical applications. *Ann. Statist.* **14**, 361–404.

[33] Siegmund, D. (1988). Confidence sets in changepoint problems. *Int. Statist. Rev.* **56**, 31–48.

[34] Tang, S. M. and MacNeill, I. B. (1993). The effect of serial correlation on tests for parameter change at unknown time. *Ann. Statist.* **21**, 552–575.

[35] van Eeden, C. (1963). The relation between Pitman's asymptotic relative efficiency of two tests and the correlation coefficient between their test statistics. *Ann. Math. Statist.* **34**, 1442–1451.

[36] Wells, M. T. (1990). The relative efficiency of goodness-of-fit statistics in the simple and composite hypothesis-testing problem. *J. Am. Statist. Assoc.* **85**, 459–463.

[37] Wieand, H. S. (1976). A condition under which the Pitman and Bahadur approaches to efficiency coincide. *Ann. Statist.* **4**, 1003–1011.

[38] Wolfe, D. A. and Schechtman, E. (1984). Nonparametric statistical procedures for the changepoint problem. *J. Statist. Plann. Inf.* **9**, 389–396.

[39] Worsley, K. J. (1986). Confidence regions and tests for a changepoint in a sequence of exponential family variables. *Biometrika* **73**, 91–104.

[40] Zacks, S. (1983). Survey of classical and Bayesian approaches to the changepoint problem: Fixed sample and sequential procedures of testing and estimation. In: MH Rizvi, JS Rustagi and DO Siegmund, eds. *Recent Advances in Statistics*, Academic Press, New York, 245–269.

James A. Koziol, Ph. D.
Department of Molecular and
Experimental Medicine
The Scripps Research Institute
10666 North Torrey Pines Road
La Jolla, CA 92037

Statistics **27** (1996) 339–343
Reprints available directly from the p
Photocopying permitted by license or

TESTING FOR A

Bul

(Received 19

Summary. A test based on Spearmar
rankings are chosen randomly and un
is that rankings are biased due to ord
for complete rankings.

AMS subject classification: 62G10

Key words: Partial rankings, probabil

1. INTRODUCTION

The statistic D, introduced by
group of judges shows favorab
ranks in analyzing data permi
Suppose that a judge ranks
permutation of the integers
permutation group S_n.
Suppose than that there are
ing ranking $\pi_j = \langle \pi_j(1), \ldots, \pi_j(n$
and uniformly distributed ran
or not the rankings have mov
Verducci (1988) test the null h
does not depend on the order
tend to give worse scores to th

where $v_0 = \langle n, \ldots, 1 \rangle$ and the c
The test rejects the null hyp
for small values of D.
An incomplete analysis arise
number of objects. There are
testants are admitted to the ne
bias in ordering of these conte