TESTS FOR CHANGE-POINTS BASED ON RECURSIVE U-STATISTICS

by

Pranab Kumar Sen

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1402

May 1982

Tests for Change-Points Based on Recursive U-Statistics*

PRANAB KUMAR SEN

University of North Carolina, Chapel Hill

Tests for change-points for location or regression models are often based on cumulative sums of recursive residuals. In the context of general estimable parameters (functionals of the underlying distributions), such recursive residuals may be defined in terms of recursive U-statistics. Along with some invariance principles for recursive U-statistics, asymptotic properties of some proposed tests for change-points are studied.

1. Introduction. Let $X_1,\ldots,X_n$ be independent random vectors (r.v.), taken at (ordered) time points $t_1,\ldots,t_n$, respectively, where X_i has a distribution function (d.f.) F_i, defined on the p ($\geq$ 1)-dimensional real space R^p, for $i\geq 1$. In the conventional models, one assumes that $F_1=\ldots=F_n=F$ (belonging to a family $\mathcal{F}$ of d.f.) and likes to draw statistical inference on a functional $\theta(F)$ of the d.f. F, defined on $\mathcal{F}$. There are, however, problems in which a change of the d.f. (and hence, the parameter $\theta(F)$) may occur at an unknown time point $\tau \in (t_1,t_n)$ [viz., Page (1957)]. As such, it may be of some interest to test for such a possible change at an unknown τ. Thus, one may desire to test the null hypothesis

(1.1) $\quad H_0\colon\; F_1=\ldots=F_n=F$ (unknown), $F \in \mathcal{F}$,

against the composite alternatives

(1.2) $\quad F_1=\ldots=F_q\neq F_{q+1}=\ldots=F_n$, for some q: $1 \leq q \leq n-1$,

where $F_1\in\mathcal{F}$, $F_n\in\mathcal{F}$ and $\mathcal{F}=\{F\colon\ \theta(F) \text{ exists}\}$.

AMS Subject ClassificationL 62E20, 60F17, 62L99.

Key Words & Phrases: Asymptotic power, contiguity, continuous mapping theorem, invariance principles, recursive residuals, tightness.

*Work supported by the National Heart, Lung and Blood Institute, Contract NIH-NHLBI-71-2243-L from the National Institutes of Health.

For linear and parametric models, tests for change-points have been considered by a host of workers; detailed accounts of these developments are given in Hinkley (1980) and Hackle(1980). *Recursive residuals* play a vital role in this context, and have been thoroughly studied by Brown, Durbin and Evans (1975) and others. Some further studies in this direction are due to Deshayes and Picard (1981) and Sen (1981a,b), among others. Such recursive residuals have not been used so extensively in the nonparametric case, though for the simple location model, *recursive ranking* [viz., Bhattacharya and Frierson (1981)] comes close to it.

The object of the present investigation is to formulate some change-point problems in terms of estimable parameters of the F_i, wherein *recursive U-statistics* are incorporated in providing analogues of recursive residuals. The proposed tests along with the recursive U-statistics are introduced in Section 2. Section 3 is devoted to the study of the distribution theory under H_0 in (1.1) and Section 4 deals with the asymptotic power properties of the proposed tests.

2. Recursive U-statistics and proposed tests.

For every $F \in \mathcal{F}$, consider an *estimable parameter* $\theta(F)$ of degree m (≥ 1), expressible as

$$(2.1)\qquad \theta(F) = \int_{R^{pm}} \cdots \int \phi(x_1,\ldots,x_m)\,dF(x_1)\ldots dF(x_m),\ F \in \mathcal{F},$$

where ϕ is a Borel measurable *kernel*, and, without any loss of generality, we may assume that $\phi(x_1,\ldots,x_m)$ is a symmetric function of its m arguments. Thus, if $X_1,\ldots,X_m$ are independent and identically distributed (i.i.d.)r.v. with d.f. F, then

$$(2.2)\qquad \theta(F) = E_F\phi(X_1,\ldots,X_m),\quad \forall\ F \in \mathcal{F}.$$

Hence, for $n \geq m$, under H_0 in (1.1), a symmetric, unbiased and optimal estimator of $\theta(F)$ is

$$(2.3)\qquad U_n = \binom{n}{m}^{-1} \sum_{1 \leq i_1 < \ldots < i_m \leq n} \phi(X_{i_1},\ldots,X_{i_m}) \ .$$

A very detailed study of the basic properties of U_n is due to Hoeffding (1948); some distributional results on U_n when the X_i are independent but not necessarily i.d. were also studied by him.

Note that for m=1 and $\phi(x)\equiv x$, $\theta(F)$ in (2.1) is the mean of the d.f. F and $U_n=n^{-1}\Sigma_{i=1}^n X_i=\bar{X}_n$ is the sample mean. In this special case, for normal F, tests have been proposed and studied by Chernoff and Zacks (1964) and others. These are based either on the residuals $\{\bar{X}_k-\bar{X}_n,\ 1\le k\le n\}$ or on the recursive residuals $\{X_{k+1}-\bar{X}_k,\ 1\le k\le n-1\}$. Note that if we let ${}_{n-k}\bar{X}=(n-k)^{-1}\Sigma_{i=k+1}^n X_i$, for $k=1,\ldots,n-1$, then

$$\bar{X}_k - \bar{X}_n = n^{-1}(n-k)\{\bar{X}_k - {}_{n-k}\bar{X}\},\quad 1\le k\le n-1 \quad . \tag{2.4}$$

Keeping this in mind, we may define as well

$${}_{n-k}U = \binom{n-k}{m}^{-1} \sum_{k+1\le i_1<\ldots<i_m\le n} \phi(X_{i_1},\ldots,X_{i_m})\ ,\quad m\le k\le n-m \quad , \tag{2.5}$$

and consider tests for change-points based on either of the following:

$$\{U_k - U_n;\ m\le k\le n\}\ , \tag{2.6}$$

$$\{U_k - {}_{n-k}U;\ m\le k\le n-m\}\ . \tag{2.7}$$

The identity in (2.4) holding for m=1, does not in general hold for $m\ge 2$. Nevertheless, the use of (2.6) and (2.7) generally leads to asymptotically equivalent results. To extend the recursive residuals to the general case, first, we rewrite

$$X_{k+1} - \bar{X}_k = (k+1)[\bar{X}_{k+1} - \bar{X}_k],\quad k\ge 1 \quad . \tag{2.8}$$

Thus, we may consider as well the following sequence

$$\{(k+1)[U_{k+1} - U_k],\ m\le k\le n-1\}\ . \tag{2.9}$$

To simplify (2.9) further, we introduce the *recursive U-statistics*

$$U_k^* = \binom{k-1}{m-1}^{-1} \sum_{1\le i_2<\ldots<i_m\le k-1} \phi(X_k, X_{i_2},\ldots,X_{i_m})\ ,\quad k\ge m\ . \tag{2.10}$$

Then, by (2.3) and (2.10), for every $k\ge m+1$,

$$\binom{k+1}{m}U_{k+1} = \binom{k}{m}U_k + \binom{k}{m-1}U_{k+1}^* \quad , \tag{2.11}$$

so that by (2.9) and (2.11),

$$(2.12) \qquad (k+1)(U_{k+1} - U_k) = m(U^*_{k+1} - U_k) \ \forall \ k \geq m \ .$$

Thus, we may consider the cumulative sums (CUSUM) of the residuals in (2.12) i.e.,

$$(2.13) \qquad W_k = \Sigma^k_{i=m+1}(U^*_i - U_{i-1}), \text{ for } k=m+1,\ldots,n; \ W_k=0, \ k\leq m \ ,$$

and the tests for the change-point may be based on $\{W_k, \ m<k\leq n\}$.

Before we proceed to construct the test statistics, we introduce the following notations. Following Hoeffding (1948), we let

$$(2.14) \qquad \phi_c(x_1,\ldots,x_c) = \int_{R^{p(m-c)}}\cdots\int \phi(x_1,\ldots,x_c,x_{c+1},\ldots,x_m)dF(x_{c+1})\ldots dF(x_m),$$

for $0\leq c\leq m$, $\psi_c(x_1,\ldots,x_c) = \phi_c(x_1,\ldots,x_c) - \theta(F)$ and

$$(2.15) \qquad \zeta_c = E_F\psi^2_c(X_1,\ldots,X_c) \ , \ 0\leq c\leq m \ .$$

Then $\zeta_0=0$, and for every $n\geq m$, whenever $\zeta_m<\infty$, under H_0,

$$(2.16) \qquad \mathrm{Var}(U_n-\theta(F)) = \binom{n}{m}^{-1} \Sigma^m_{c=1}\binom{m}{c}\binom{n-m}{m-c}\zeta_c$$

$$= m^2n^{-1}\zeta_1 + O(n^{-2}) \ .$$

Let then for every i (=1,...,n),

$$(2.17) \qquad U_{ni} = \binom{n-1}{m-1}^{-1} \Sigma^*_{n,i}\phi(X_i,X_{i_2},\ldots,X_{i_m}) \ ,$$

where the summation $\Sigma^*_{n,i}$ extends over all $1\leq i_2<\ldots<i_m\leq n$ with $i_j\neq i$, for $j=2,\ldots,m$. Also, let

$$(2.18) \qquad s^2_n = (n-1)^{-1}\Sigma^n_{i=1}[U_{ni} - U_n]^2 \ .$$

Then, it follows from Sen (1960) that

$$(2.19) \qquad s^2_n \xrightarrow{p} \zeta_1 = m^{-2} \lim_{n\to\infty} n \ \mathrm{Var}\{U_n - \theta(F)\} \ ;$$

in fact, it follows from Sproule (1974) and Sen (1977) that $s^2_n\to\zeta_1$ almost surely, as $n\to\infty$.

Consider then the studentized statistics

$$(2.20) \qquad D^+_{n1} = (n-m)^{-\frac{1}{2}}s^{-1}_n\{\max_{m<k\leq n} W_k\} \ ,$$

$$(2.21)\qquad D_{n1} = (n-m)^{-\frac{1}{2}} s_n^{-1} \{ \max_{m<k\leq n} |W_k| \} ,$$

which we like to propose at the test statistic for one and two-sided changes, respectively. Consider also the statistics

$$(2.22)\qquad D_{n2}^{+} = n^{-\frac{1}{2}} m^{-1} s_n^{-1} \{ \max_{m\leq k\leq n} k(U_k - U_n) \} ,$$

$$(2.23)\qquad D_{n2} = n^{-\frac{1}{2}} m^{-1} s_n^{-1} \{ \max_{m\leq k\leq n} k|U_k - U_n| \} ,$$

$$(2.24)\qquad D_{n3}^{+} = n^{-\frac{1}{2}} m^{-1} s_n^{-1} \{ \max_{m\leq k\leq n-m} \frac{(n-k)k}{n} (U_k - {}_{n-k}U) \} ,$$

$$(2.25)\qquad D_{n4}^{+} = n^{-\frac{1}{2}} m^{-1} s_n^{-1} \{ \max_{m\leq k\leq n-m} \frac{k(n-k)}{n} |U_k - {}_{n-k}U| \} .$$

Although, we shall mainly concentrate on D_{n1}^{+} and D_{n1}, we will present, side by side, the picture with the other statistics as well. In each case, the hypothesis H_0 is rejected if the statistic is adequately large. For the study of their critical values, we proceed in Section 3, to study some invariance principles related to these U-statistics, when H_0 holds.

3. Distribution theory under H_0.

Let us introduce some stochastic processes. Let $Z_n^{(1)} = \{Z_n^{(1)}(t),\ 0\leq t\leq 1\}$ be defined by letting

$$(3.1)\qquad Z_n^{(1)}(k/n) = \begin{cases} 0, & k\leq m-1 , \\ k(U_k - U_n)/(\sqrt{n}\ m\zeta_1^{\frac{1}{2}}), & k=m,\dots,n; \end{cases}$$

$$(3.2)\qquad Z_n^{(1)}(t) = Z_n^{(1)}(k/n),\ \text{for } k/n \leq t < (k+1)/n,\ 0\leq k\leq n-1 .$$

Thus, $Z_n^{(1)}$ belongs to the space $D[0,1]$, endowed with the Skorokhod J_1-topology. Also, let $Z_n^{(2)} = \{Z_n^{(2)}(t),\ 0\leq t\leq 1\}$ be defined by letting

$$(3.3)\qquad Z_n^{(2)}(k/n) = \begin{cases} 0 , & 0\leq k<m,\ n-m+1 \leq k \leq n , \\ \frac{(n-k)k}{n}(U_k - {}_{n-k}U)/(\sqrt{n}\ m\zeta_1^{\frac{1}{2}}), & m\leq k\leq n-m ; \end{cases}$$

$$(3.4)\qquad Z_n^{(2)}(t) = Z_n^{(2)}(k/n),\ \text{for } k/n \leq t < (k+1)/n,\ 0\leq k\leq n-1 .$$

(3.13) $\quad Z_n \underset{\mathcal{D}}{\to} Z$, in the J_1-topology on $D[0,1]$,

where Z is a standard Wiener process on $[0,1]$. Since $Z_n(0)=0$, with probability 1, (3.13) ensures that

(3.14) $\quad \sup_{0\le t\le 1} |Z_n(t)| = O_p(1)$,

and for every $\varepsilon>0$ and $\eta>0$, there exist a δ: $0<\delta<1$ and a sample size n_0, such that

(3.15) $\quad P\{\omega_\delta(Z_n)>\varepsilon\} < \eta$, $\quad \forall \; n\ge n_0$,

where the modulus of continuity $\omega_\delta(x)$ is defined by

(3.16) $\quad \sup\{|x(t)-x(s)|: 0\le s<t\le s+\delta\le 1\}$, $0<\delta<1$.

Actually $\{U_k-\theta(F),\ k\ge m\}$ is a reverse martingale, and hence, by (3.11) and the Hájek-Rényi-Chow inequality for reverse (sub-) martingales, we have for every q: $q\le n$,

(3.17) $\quad P\{\max_{k\le q}|Z_n(\frac{k}{n})| \ge K(q/n)^{\frac{1}{2}}\} \le 2K^{-2}$, $\forall\ K>0$.

Now, by (3.10) and (3.11), for every $k\ge m+1$, we have

(3.18) $\quad (n\zeta_1)^{-\frac{1}{2}}W_k = Z_n(\frac{k}{n}) - \Sigma_{i=m}^{k-1}\frac{1}{i}Z_n(\frac{i}{n}) - Z_n(\frac{m}{n})$,

where $|Z_n(\frac{m}{n})| \overset{p}{\to} 0$, as $n\to\infty$. Note that if we let

(3.19) $\quad Z_n^*(t) = Z_n(t) - \int_{0+}^{t-} s^{-1}Z_n(s)ds$, $0<t\le 1$,

then, by (3.12), (3.14) and (3.17) (where we let $q\to\infty$ but $n^{-1}q\to 0$),

(3.20) $$\max_{k\le n}|\Sigma_{i=m}^{k-1}\frac{i}{m}Z_n(\frac{i}{n}) - \int_{0+}^{k/n-} s^{-1}Z_n(s)ds|$$

$$= \max_{k\le n}|\frac{1}{n}\Sigma_{i=m}^{k-1}Z_n(\frac{i}{n})/(\frac{i}{n}) - \Sigma_{i=m}^{k-1}Z_n(s)\log\frac{i+1}{i}|$$

$$\le \max_{k\le n}\{\Sigma_{i=m}^{k-1}|Z_n(\frac{i}{n})|i^{-2}\} \le \Sigma_{k=m}^{n-1}|k^{-2}Z_n(\frac{k}{n})|$$

$$= \Sigma_{k=m}^{q-1}k^{-2}|Z_n(\frac{k}{n})| + \Sigma_{k=q}^{n-1}k^{-2}|Z_n(\frac{k}{n})|$$

$$\le \max_{k\le q-1}|Z_n(\frac{k}{n})| \cdot \frac{2}{m} + \max_{k\le n}|Z_n(\frac{k}{n})| \cdot 2q^{-1}$$

$$= O_p((\tfrac{q}{n})^{\frac{1}{2}}) + O_p(1).\ 2q^{-1} \overset{p}{\to} 0 \quad \text{as } q\to\infty \text{ but } q/n\to 0.$$

Therefore, we may use the continuous mapping theorem, and hence, for the finite dimensional distributions of $Z_n^{(3)}$, it suffices to use (3.13) along with the continuous mapping: $Z^*(t) = Z(t) - \int_0^t Z(s)s^{-1}ds$, $0<t<1$; $Z^*(0)=0$. It is easy to verify that $Z^* = \{Z^*(t);\ 0\le t\le 1\}$ is also a standard Wiener process. We proceed next to show that Z_n^*, defined by (3.19) is tight. Towards this, note that by (3.15), it suffices to show only that $\{\int_{0+}^{t-} s^{-1}Z_n(s)ds,\ 0<t\le 1\}$ is tight. For this, note that for every $\eta>0$,

$$
\begin{aligned}
(3.21)\qquad & \sup_{0<t<\eta}\left|\int_{0+}^{t-} s^{-1}Z_n(s)ds\right| \\
& \le \max_{k\le n\eta}\Sigma_{i=m}^{k}\ i^{-1}|Z_n(\tfrac{i}{n})| \\
& \le \Sigma_{i=m}^{[n\eta]} i^{-1}(Z_n^2(\tfrac{i}{n}))^{\frac{1}{2}},
\end{aligned}
$$

where the right hand side has the expectation $\le \Sigma_{i=m}^{[n\eta]} i^{-1}(i/n)^{\frac{1}{2}} \sim 2\sqrt{\eta}$, and hence, choosing η (> 0), sufficiently small, the left hand side can be made arbitrarily small, in probability. On the other hand, for $s''>s'\ge\eta>0$,

$$
\begin{aligned}
(3.22)\qquad & \sup_{s''\ge t\ge s'}\left|\int_{s'}^{t} s^{-1}Z_n(s)ds\right| \le \sup_{s'\le s\le s''}|Z_n(s)|(s''-s')/s' \\
& \le \left(\sup_{0<t\le 1}|Z_n(t)|\right)\eta^{-1}(s''-s'),\ \forall\ 0<\eta\le s'\le s''\le 1.
\end{aligned}
$$

Hence, by (3.14) and (3.22), we conclude that for $s''-s'<\delta$, (3.22) can be bounded (uniformly in s') by $\delta\eta^{-1}(\sup_{0\le t\le 1}|Z_n(t)|)$ and this can be made to converge to 0, in probability, by letting $\delta\downarrow 0$. This establishes the desired tightness property, and hence, we arrive at the following theorem (where note that $Z^* \underset{\mathcal{D}}{=} Z$).

Theorem 3.1. Under H_0 in (1.1), *for* $0<\zeta_1<\zeta_m<\infty$,

$$
(3.23)\qquad Z_n^{(3)} \underset{\mathcal{D}}{\to} Z,\ \textit{in the } J_1\textit{-topology on } D[0,1].
$$

Note that for every $\lambda\ge 0$,

$$
(3.24)\qquad P\{\sup_{0\le t\le 1} Z(t)\ge\lambda\} = 2P\{Z(1)\ge\lambda\} = \sqrt{\frac{2}{\pi}}\int_{\lambda}^{\infty} e^{-\frac{1}{2}t^2}dt,
$$

$$(3.25)\qquad P\{\sup_{0\le t\le 1}|Z(t)|\ge\lambda\} = 1 - \sum_{k=-\infty}^{\infty}(-1)^k[\Phi((2k+1)\lambda)-\Phi((2k-1)\lambda)] ,$$

where Φ is the standrad normal d.f., and

$$(3.26)\qquad P\{\sup_{0\le t\le 1} Z^o(t)\ge\lambda\} = e^{-2\lambda^2} ,$$

$$(3.27)\qquad P\{\sup_{0\le t\le 1}|Z^o(t)|\ge\lambda\} = 2\sum_{m=1}^{\infty}(-1)^{m+1}e^{-2m^2\lambda^2} .$$

Let λ_α^+, λ_α, $\lambda_{0\alpha}^+$ and λ be the values of λ for which the right hand sides of (3.24), (3.25), (3.26) and (3.27) are all equal to α $(0<\alpha<1)$. Then, by virtue of (2.19)-(2.25), (3.1)-(3.6), (3.7), (3.9), (3.23), and (3.24)-(3.27), we conclude that the asymptotic critical values for D_{n1}^+, D_{n1}, D_{n2}^+ (or D_{n3}^+) and D_{n2} (or D_{n3}) are λ_α^+, λ_α, $\lambda_{0\alpha}^+$ and $\lambda_{0\alpha}$, respectively, where α $(0<\alpha<1)$ corresponds to the desired significance level. Note that

$$(3.28)\qquad \lambda_\alpha^+ = \tau_{\alpha/2},\ \Phi(\tau_\varepsilon) = 1-\varepsilon;\ \ \lambda_{0\alpha}^+ = \sqrt{-\tfrac{1}{2}\log\alpha} ,$$

while for α close to 0, $\lambda_\alpha \sim \lambda_{\alpha/2}^+$ and $\lambda_{0\alpha} \sim \sqrt{-\frac{1}{2}\log\alpha/2}$. Instead of the unweighted W_k in (2.20)-(2.21), we could have also taken the standardized version $(k-m)^{-\frac{1}{2}}W_k$. In that case, if in the range $m\le k\le n$, we replace the lower limit m by ηn for some $\eta>0$, then we are also able to use appropriate Wiener process approximation with square root boundaries, and critical values for such a case, for various $\eta>0$, have been tabulated by DeLong (1981). A similar case holds when we use the standardized forms in (2.22)-(2.25), provided we replace m by $[n\eta]$, for some $\eta>0$. A choice of $\eta=.05$, .01 or .1 may work out quite well.

4. Asymptotic power properties. The asymptotic distribution theory of U-statistics, when the F_i are not necessarily identical, has also been studied by Hoeffding (1948; Sec. 8). Let us introduce the following notations. Let

$$(4.1)\qquad \bar{F}_k(x) = k^{-1}\Sigma_{i=1}^k F_i(x),\ x\in E^p,\ k\ge 1.$$

Define $\theta(\bar{F}_k)$ as in (2.1) with F being replaced by $\bar{F}_k$, $k\ge 1$. Also, define $\zeta_1(\bar{F}_k)$ as in (2.14)-(2.15) with F being replaced by $\bar{F}_k$, $k\ge 1$. Further, define

$\phi_1^{(F)}(x_1)$ as in (2.14), where the superscript F signifies the role of F in the integration performed. Also, let

$$\gamma(F_i;\, F_j) = \int\phi_1^{(F_j)}(x)dF_i(x) - \theta(F_j)\ , \tag{4.2}$$

$$\Delta_k^2 = k^{-1}\Sigma_{i=1}^k\gamma^2(F_i,\overline{F}_k),\ k\geq 1. \tag{4.3}$$

Then, as in Sen (1969), we may state the basic result of Hoeffding (1948, Theorem 8.1) in a simplified version as follows: for every t: $0<t\leq 1$, as $n\to\infty$,

$$[nt]^{\frac{1}{2}}[U_{[nt]}-\theta(\overline{F}_{[nt]})] \sim N(0,\ m^2\{\zeta_1([nt])-\Delta_{[nt]}^2\})\ , \tag{4.4}$$

where the regularity conditions of Hoeffding (1948) are presumed. As such, if we let

$$\gamma_n^+ = \max_{m<k\leq n}\{\frac{1}{n-m}\ \Sigma_{i=m+1}^k\gamma(F_i;\ \overline{F}_{i-1})\} \tag{4.5}$$

and if $\underline{\mathit{lim}}\ \gamma_n^+ > 0$, then the consistency of D_{n1}^+ can be established by some standard steps. Similarly, if $\gamma_n = \max_{m<k\leq n}\{\frac{1}{n-m}|\ \Sigma_{i=m+1}^k\gamma(F_i;\ \overline{F}_{i-1})|\}$ and $\underline{\mathit{lim}}\ \gamma_n > 0$, then the test based on D_{n1} is consistent. A similar picture holds for the other tests too.

Note that under (1.2), if we let F_1=F and F_n=G, then

$$\gamma(F_i,\overline{F}_{i-1}) = \begin{cases} 0, & i\leq q \\ \frac{q}{i-1}\int\phi_1^{(\overline{F}_{i-1})}(x)d[G(x)-F(x)], & q+1\leq i\leq n, \end{cases} \tag{4.6}$$

where $\overline{F}_{i-1} = (i-1)^{-1}\{qF + (i-1-q)G\}$, $q+1\leq i\leq n$. Consequently, when q/n is away from 0 or 1 and the ϕ_1 functions behave regularly, then for fixed $F\neq G$, $\underline{\mathit{lim}}\ \gamma_n^+$ (or $\underline{\mathit{lim}}\ \gamma_n$) will be positive, and hence, the tests will be consistent. As such, for the study of the asymptotic power properties, we will confine ourselves to some local alternatives for which the asymptotic power function exists and is different from 1 (or 0).

Recall that $\underset{\sim}{X}_n = (X_1,\ldots,X_n)$ belongs to a measure space $(\mathcal{X}_n,\ \mathcal{A}_n,\ \mu_n)$ where, under (1.1), the d.f. F generates a sequence $\{P_n\}$ of probability measures, assumed to be absolutely continuous (with respect to $\{\mu_n\}$); under

(1.2) or more generally, for $(F_1,\ldots,F_n)$, let $\{Q_n\}$ be the sequence of probability measures, also assumed to be absolutely continuous (μ_n). We may then formally define the (product) densities $p_n = dP_n/d\mu_n$ and $q_n = dQ_n/d\mu_n$, $n\geq 1$. We assume that

(4.7) $\{q_n\}$ are *contiguous* to $\{p_n\}$;

for some good discussion on the contiguity of probability measures, we may refer to Hájek and Šidák (1967, Ch. VI). To fit this contiguous model with our (1.2), we need to elaborate a bit more. First, we assume that the change point τ $(= \tau_n)$ belongs to (t_{q_n}, t_{q_n+1}), where q_n is $\nearrow$ in n with

(4.8) $$\lim_{n\to\infty}(n^{-1}q_n) = \pi:\ 0<\pi<1.$$

Secondly, (1.2) holds with (4.8) and with $F_1=F$ and F_n satisfying

(4.9) $$\sqrt{n}(F_n(x)-F(x)) = H(x),\ x\in E^p,$$

and the choice of H satisfies the contiguity in (4.7). Note that for this model, the log-likelihood ratio criterion is

(4.10) $$\Sigma_{i=1}^n \log\{dF_i(X_i)/dF(X_i)\}$$

$$= \Sigma_{i=q_n+1}^n \log\{1 + n^{-\frac{1}{2}}dH(X_i)/dF(X_i)\}\ ,$$

and hence, the contiguity may be verified easily be invoking the asymptotic normality of (4.10) (under (4.8) and H_0 in (1.1)). Let us denote the sequence of alternatives in (4.7)-(4.9) by $\{K_n\}$. Note that if we let

(4.10) $$\gamma = \int\phi_1^{(F)}(x)dH(x),$$

$$= \int\phi_1^{(F)}(x)[dH(x)/dF(x)]dF(x)\ ,$$

then γ exists whenever $\int[dH(x)/dF(x)]^2dF(x)$ is finite [as is needed to ensure the normality in (4.10)]. It follows from (4.6), (4.9) and (4.10) that whenever $\frac{i}{n}\to t$: $0\leq t\leq 1$,

(4.11) $$\delta(t) = \lim_{n\to\infty}\{\sqrt{n}\gamma(F_i,\overline{F}_{i-1})|K_n\} = \begin{cases}0, & t\leq\pi\\ \pi\gamma/t, & \pi<t\leq 1\end{cases}\ .$$

[Note that this implies that under $\{K_n\}$,

(4.12) $\lim_{n\to\infty} \sqrt{n-m}\ \gamma_n^+ = (-\pi\log\pi)\gamma,\quad \lim_{n\to\infty} \sqrt{n-m}\ \gamma_n = (-\pi\log\pi)|\gamma|,$

and the limits vanishes whenever $\pi\to0$ or 1.]

The main advantage of invoking contiguity in this context is to insure that the tightness property of the $Z_n^{(j)}$, $1\le j\le 3$, establishes in Section 3, under H_0 in (1.1) (i.e., under $\{P_n\}$ measure) remains in tact under $\{K_n\}$ (i.e., under $\{Q_n\}$ measure) as well. Hence, to study the weak convergence of the $Z_n^{(j)}$ $(1\le j\le 3)$ or Z_n under $\{K_n\}$, it suffices to establish the convergence of their finite-dimensional distributions to those of appropriate Gaussian functions (with some drifts); see Theorem 4.3.4 of Sen (1981) in this context. For Z_n, defined in (3.11)-(3.12), the weak convergence (under H_0) was studied by Miller and Sen (1972) by writing

(4.13) $Z_n = \Sigma_{h=1}^{m}\binom{m}{h}Z_{n,h}$,

where $Z_{n,1}$ involves independent summands and showing that

(4.14) $\max_{2\le h\le m}\ \sup_{0\le t\le 1}|Z_{n,h}(t)| \overset{p}{\to} 0$, under H_0.

Note that by (4.13)-(4.14),

(4.15) $||Z_n - Z_{n,1}|| \overset{p}{\to} 0$, under H_0,

where by the contiguity and (4.14)-(4.15),

(4.15) $||Z_n - Z_{n,1}|| \overset{p}{\to} 0$, under $\{K_n\}$ as well

Hence, it suffices to consider the finite dimensional distributions of $\{Z_{n,1}\}$, where

(4.16) $Z_{n,1}(t) = [n\zeta_1(F)]^{-\frac{1}{2}}\Sigma_{i=1}^{k}\{\phi_1^{(F)}(X_i) - \theta(F)\}$, $k=[nt]$,

for every $0<t\le1$. Again, $Z_{n,1}(t)$ is asymptotically normal under H_0 and it involves a triangular array of independent summands. Hence, by an appeal to a result of Behnen and Neuhaus (1975), we conclude that $Z_{n,1}(t)$ is asymptotically normal (under $\{K_n\}$) with mean $(t-\pi)\gamma/[\zeta_1(F)]^{\frac{1}{2}}$ (for $t\ge\pi$ and otherwise 0) and unit variance. The case of more than one t (i.e., $Z_{n,1}(t_1),\dots,Z_{n,1}(t_r)$)

follows on parallel lines. A very similar case holds for $Z_n^{(1)}$ and $Z_n^{(2)}$. Further, the assumed contiguity also ensures (2.19) under $\{K_n\}$ as well. Thus, under $\{K_n\}$, we have

$$(4.16)\qquad Z_n^{(3)} \xrightarrow[\mathcal{D}]{} Z + \mu, \quad Z_n^{(1)} \xrightarrow[\mathcal{D}]{} Z^o + \mu^o, \quad Z_n^{(2)} \xrightarrow[\mathcal{D}]{} Z^o + \mu^o ,$$

where Z and Z^o are Brownian motion and bridge, respectively and the drift functions $\mu = \{\mu(t),\ 0\le t\le 1\}$ and $\mu^o = \{\mu^o(t)3\ 0\le t\le 1\}$ are specified by

$$(4.17)\qquad \mu(t) = \begin{cases} 0, & 0<t\le\pi \\ \pi\{\log(t/\pi)\}\gamma/\{\zeta_1(F)\}^{\frac{1}{2}}, & \pi\le t\le 1 \ ; \end{cases}$$

$$(4.18)\qquad \mu^o(t) = \begin{cases} -t(1-\pi)\gamma/\{\gamma_1(F)\}^{\frac{1}{2}}, & 0\le t\le\pi \ , \\ -\pi(1-t)\gamma/\{\zeta_1(F)\}^{\frac{1}{2}}, & \pi\le t\le 1. \end{cases}$$

Consequently, the asymptotic power function (under $\{K_n\}$) of the tests based on D_{n1}^+ and D_{n1} are given by

$$(4.19)\qquad \lim_{n\to\infty} P\{D_{n1}^+ \ge \lambda_\alpha^+ | K_n\}$$

$$= P\{Z(t) + \mu(t) \ge \lambda_\alpha^+ \text{ for some } t\colon 0\le t\le 1\} \ ;$$

$$(4.20)\qquad \lim_{n\to\infty} P\{D_{n1} \ge \lambda_\alpha | K_n\}$$

$$= P\{|Z(t) + \mu(t)| \ge \lambda_\alpha \text{ for some } t\colon 0\le t\le 1\} \ ,$$

where $\mu(t)$ is specified by (4.17). Similarly, the asymptotic power functions for D_{n2}^+ (or D_{n3}^+) and D_{n2} (or D_{n3}) are given by the boundary crossing probabilities for the drifted Brownian bridge $Z^o + \mu^o$, for the one and two sided cases, respectively.

The case of more than one change-points can also be accommodated in the above setup, where in (4.9), we need to replace H(x) by

$$(4.21)\qquad \sqrt{n}(F_{ni}(x) - F(x)) = H_i(x), \quad 1\le i\le n,$$

where the H_i can be one of r ($\ge$ 1) different $H^{(1)},\ldots,H^{(r)}$, depending on where i/n belongs to (one of r nonoverlapping intervals on (0,1)). As such, (4.11), (4.17) and (4.18) will have to be replaced by more complicated segmented curves.

We conclude this section by the remark that the asymptotic power functions of D_{n1}^{+} (or D_{n1}) and D_{n2}^{+} (or D_{n2}) do not correspond to forms wherein the classical Pitman-efficiency measure can be adapted to study the asymptotic relative efficiency of one with respect to the other. In this respect, for various π $(0<\pi<1)$ and $\gamma/\{\zeta_1(F)\}^{\frac{1}{2}}$, numerical comparisons of the boundary crossing probabilities should provide some broad ideas.

REFERENCES

BHATTACHARYA, P.K. and FRIERSON, D. (1981). A nonparametric control chart for detecting small disorders. *Ann. Statist. 9*, 544-554.

BROWN, R.L., DURBIN, J. and EVANS, J.M. (1975). Techniques for testing constancy of regression relationships over time (with discussions). *J. Roy. Statist. Soc. Ser. B 37*, 149-192.

CHERNOFF, H. and ZACKS, S. (1964). Estimating the current mean of a normal distribution which is subject change in time. *Ann. Math. Statist. 35*, 999-1018,

DELONG, D. (1981). Crossing probabilities for a square root boundary by a Bessel process. *Comm. Statist. Ser. A 10*, 2197-2213.

DESHAYES, J. and PICARD, D. (1981). Tests de rupture de regression: Comparison asymptotic. *Teor. Vero. Prem.*

HACKLE, P. (1980). *Testing the Constancy of Regression Relationships Over Time*. Gottingen: Vandenhoeck & Ruprecht.

HÁJEK, J. and ŠIDÁK, Z. (1967). *Theory of Rank Tests*. New York: Academic Press.

HINKLEY, D.V. (1980). Bibliography on articles on change-points. *Tech. Rept.*, Univ. of Minnesota.

HOEFFDING, W. (1948). A class of statistics with asymptotically normal distribution. *19*, 293-325.

HOEFFDING, W. (1961). The strong law of large numbers for U-statistics. *Inst. Statist. Univ. North Carolina*, Mimeo Rep. No. 302.

MILLER, R.G. Jr. and SEN, P.K. (1972). Weak convergence of U-statistics and Von Mises' differentiable functions. *Ann. Math. Statist. 43*, 31-41.

NEUHAUS, G. and BEHNEN, K. (1975). A central limit theorem under contiguous alternatives. *Ann. Statist. 3* , 1349-1354.

PAGE, E.S. (1957). On problems in which a change of parameters occurs at an unknown time point. *Biometrika 44*, 248-252.

SEN, P.K. (1960). On some convergence properties of U-statistics. *Calcutta Statist. Asso. Bull. 10*, 1-18.

SEN, P.K. (1969). On a robustness property of a class of nonparametric tests based on U-statistics. *Calcutta Statist. Asso. Bull. 18*, 51-60.

SEN, P.K. (1977). Some invariance principles relating to jack-knifing and their role in sequential analysis. *Ann. Statist. 5*, 315-329.

SEN, P.K. (1981). *Sequential Nonparametrics: Invariance Principles and Statistical Inference*. New York : John Wiley.

SEN, P.K. (1982a). Invariance principles for recursive residuals. *Ann. Statist. 10*, 307-312.

SEN, P.K. (1982b). Asymptotic theory of some tests for constancy of regression relationships over time. *Math. Operat. Statist., Ser. Statist. 13*.

SPROULE, R.N. (1979). Asymptotic properties of U-statistics. *Trans. Amer. Math. Soc. 199*, 55-64.