The Annals of Statistics
1982, Vol. 10, No. 1. 307–312

INVARIANCE PRINCIPLES FOR RECURSIVE RESIDUALS[1]

BY PRANAB KUMAR SEN

University of North Carolina, Chapel Hill

A general class of recursive residuals is defined by means of lower-triangular, orthonormal transformations. For these residuals, some weak invariance principles are established under appropriate regularity conditions. The theory is then incorporated in the study of robustness of some tests for change of parameters occurring at unknown time points.

1. Introduction. Various types of recursive residuals are commonly encountered in problems of inference about the change point of a sequence of random variables. In the context of testing for constancy of regression relationships over time, Brown, Durbin and Evans (1975) have considered some CUSUM tests based on suitably defined recursive residuals; the corresponding location problem treated earlier by several workers is included in their setup as a particular case. When the error components are assumed to be independent and identically distributed (i.i.d.) according to a normal distribution, these recursive residuals are mutually independent and distributed according to a common normal distribution, so the Brownian motion approximation can readily be incorporated for the study of the properties of some CUSUM (or Cramér-von Mises type) tests based on these residuals. The situation may differ considerably when the errors are not normally distributed: these residuals remain uncorrelated but not necessarily independent or normally distributed. Several discussants of the Brown et al. (1975) paper have raised the issues of incorporating more general forms of orthonormal transformations for generating recursive residuals and establishing weak invariance principles for the related CUSUM test statistics without imposing normality on the distribution of the errors.

The object of the present investigation is to study some weak invariance principles for recursive residuals generated by a class of orthonormal transformations when the errors are not necessarily normally distributed. Along with the preliminary notions, these orthonormal transformations are introduced in Section 2. Section 3 deals with the main theorems and their proofs. Section 4 is devoted to applications of the main theorems to some CUSUM procedures based on recursive residuals and relates to some of their asymptotic properties.

2. Preliminary notions. Consider the regression model

$$Y_t = \boldsymbol{\beta}_t'\mathbf{x}_t + e_t, \qquad t = 1, \cdots, n, \tag{2.1}$$

where, at time t, Y_t is the dependent variate, $\mathbf{x}_t = (x_{t1}, \cdots, x_{tk})'$ (for some $k \geq 1$) is the vector of regressors, $\boldsymbol{\beta}_t = (\beta_{t1}, \cdots, \beta_{tk})'$ is the vector of regression coefficients and the e_t are the error components. For testing the null hypothesis

$$H_0 : \boldsymbol{\beta}_t = \boldsymbol{\beta} \text{ (unknown) for all } t$$

along with the i.i.d. character of the e_t, i.e., the constancy of the regression relationships over time, Brown, Durbin and Evans (1975) considered the regression residuals

---

Received October 1980; revised May 1981.

[1] Work partially supported by the National Heart, Lung and Blood Institute, Contract No. NIH-NHLBI-71-2243-L from the National Institutes of Health.

AMS 1980 *subject classifications.* 60F17, 62F05.

Key words and phrases. Constancy of regression (location) relationship over time, CUSUM tests, invariance principles, orthonormal transformations, recursive residuals, tightness, unknown transition points, Wiener process.

(2.2) $$w_r = (Y_r - \mathbf{b}'_{r-1}\mathbf{x}_r)/\{1 + \mathbf{x}'_r(\mathbf{X}'_{r-1}\mathbf{X}_{r-1})^{-1}\mathbf{x}_r\}^{1/2}, \qquad k+1 \le r \le n,$$

where, for every $r(=k, \cdots, n)$,

(2.3) $$\mathbf{X}'_r = (\mathbf{x}_1, \cdots, \mathbf{x}_r), \quad \mathbf{Y}'_r = (Y_1, \cdots, Y_r) \quad \text{and} \quad \mathbf{b}_r = (\mathbf{X}'_r\mathbf{X}_r)^{-1}\mathbf{X}'_r\mathbf{Y}_r.$$

Their CUSUM test is then based on the statistics

(2.4) $$D_n^+ = \max_{k<r\le n} W_r/\{s_n(n-k)^{1/2}\} \quad \text{and} \quad D_n = \max_{k<r\le n} |W_r|/\{s_n(n-k)^{1/2}\},$$

where

(2.5) $$W_r = \sum_{i=k+1}^{r} w_i, \quad k+1 \le r \le n \quad \text{and} \quad s_n^2 = (n-k)^{-1}(\mathbf{Y}_n - \mathbf{X}_n\mathbf{b}_n)'(\mathbf{Y}_n - \mathbf{X}_n\mathbf{b}_n).$$

When the e_t are i.i.d. r.v.'s (random variables) with a normal distribution, the w_i are also i.i.d. r.v.'s with the same normal distribution, so that the Brownian motion approximation for the W_r can readily be incorporated in the study of the distribution of D_n^+ or D_n. However, when the e_t are not normally distributed, the w_r are not necessarily independent nor normally distributed, and hence, invariance principles for these recursive residuals remain to be explored. Note that by (2.1) and (2.2), for each $r(> k)$, w_r can be expressed as a linear combination of $e_i (i \le r)$ plus a nonstochastic component which vanishes under H_0. Keeping this in mind, we conceive of a sequence $\{U_i, i \ge 1\}$ of i.i.d. r.v.'s, assume that

(2.6) $$EU_i = 0 \quad \text{and} \quad 0 < \sigma^2 = EU_i^2 < \infty,$$

let $\mathbf{U}_n = (U_1, \cdots, U_n)'$, $n \ge 1$, and define

(2.7) $$\mathbf{V}_n = (V_{m+1}, \cdots, V_n)' = \mathbf{A}_n\mathbf{U}_n, \quad n \ge m+1,$$

where m is a nonnegative integer, $\mathbf{A}_n$ is an $(n-m) \times n$ matrix with row vectors $\mathbf{a}_{nj} = (\mathbf{a}_j, \mathbf{O}_{n-j})$, $m+1 \le j \le n$ and the $\mathbf{a}_{nj}$ satisfy the following orthonormality condition:

(2.8) $$\mathbf{A}_n\mathbf{A}'_n = \mathbf{I}_{n-m}, \quad \text{i.e. } \mathbf{a}_j\mathbf{a}'_j = 1 \quad \text{for} \quad j = m+1, \cdots, n.$$

The class of residuals $\mathbf{V}_n$, generated by the class of lower triangular or trapizoidal matrices $\mathbf{A}_n$ in (2.8), contains the w_r as a special case, where $m = k$. In view of the assumed normality of the e_t in (2.1), Brown et al. (1975) did not require any further condition on the $\mathbf{x}_t$. However, for the general case to be treated here, we may need some of the following conditions:

(2.9) $$\lim_{n\to\infty} \max_{m<k\le n}\{n^{-1}(\sum_{j=k}^{n} |a_{jk}|)^2\} = 0,$$

(2.10) $$\sup_{n>m}\{\max_{m<k\le n}\max_{1\le j\ge k}(k\,|a_{kj} - \delta_{kj}|)\} \le c_1 < \infty,$$

(2.11) $$\sup_{n>m}\{\max_{1\le i\le n-1} i\,|\sum_{k=(m\vee i)+2}^{n}(a_{ki} - a_{ki+1})|\} \le c_2 < \infty,$$

where δ_{rs} stands for the Kronecker delta and c_1, c_2 are finite positive numbers. It may be remarked that (2.10) implies (2.9), but not conversely. Also, for recursive residuals generally, the a_{kk} are close to 1, $a_{kj}, j > k$ are all equal to 0 and the $a_{kj}, j < k$ are close to 0. These conditions are reflected in (2.9) through (2.11). The main results are presented in the next section.

3. The main results. For every $n > m$, we introduce a stochastic process $W_n = \{W_n(t), t \in [0, 1]\}$ by letting $V_i = 0$, $i \le m$ and

(3.1) $$W_n(t) = \sum_{i\le k_n(t)} V_i/\{\sigma(n-m)^{1/2}\}, \quad k_n(t) = m + [(n-m)t], \qquad 0 \le t \le 1,$$

where σ and the V_i are defined as in (2.6), (2.7) and (2.8). Then, W_n belongs to the $D[0, 1]$ space endowed with the Skorokhod J_1-topology. Also, let $W = \{W(t), t \in [0, 1]\}$ be a standard Wiener process on [0, 1]. We are primarily interested in the following weak convergence result:

(3.2) $$W_n \to_{\mathscr{D}} W, \quad \text{in the } J_1\text{-topology on } D[0, 1].$$

Towards this, we consider the following two theorems under different sets of regularity conditions.

THEOREM 1. *Under* (2.6), (2.7), (2.8), (2.9) *and* $\nu_4 = EU_i^4 < \infty$, (3.2) *holds.*

THEOREM 2. *Under* (2.7), (2.8), (2.10) *and* (2.11), (2.6) *insures* (3.2).

Note that the 4th moment condition in Theorem 1 has been counter-balanced in Theorem 2 by the extra conditions (2.10) and (2.11).

PROOF OF THEOREM 1. To establish (3.2), we need to show that the finite-dimensional distributions (f.d.d.) of $\{W_n\}$ converge to those of W and that $\{W_n\}$ is *tight*. For arbitrary $r(\geq 1)$, $0 \leq t_1 < \cdots < t_r \leq 1$ and $\lambda = (\lambda_1, \cdots, \lambda_r)' \neq 0$. On letting $k_n(t_j) = k_j$, $1 \leq j \leq r$, we have

$$\text{(3.3)} \qquad \begin{aligned} \textstyle\sum_{j=1}^r \lambda_j W_n(t_j) &= \{\sigma(n-m)^{1/2}\}^{-1} \textstyle\sum_{j=1}^r \lambda_j \sum_{i \leq k_j} V_i \\ &= \{\sigma(n-m)^{1/2}\}^{-1} \textstyle\sum_{i=1}^n (\sum_{j=1}^r \lambda_j \sum_{s=i}^{k_j} a_{si}) U_i = \sum_{i=1}^n c_{ni} U_i/\sigma, \end{aligned}$$

say, where, conventionally, $\sum_k^r a_{sk} = 0$ for $k > r$. Now, by (2.6) and (2.8), (3.3) has zero mean and variance

$$\textstyle\sum_{i=1}^n c_{ni}^2 \to \sum_{j=1}^r \sum_{\ell=1}^r \lambda_j \lambda_\ell (t_j \wedge t_\ell) = E\{\sum_{j=1}^r \lambda_j W(t_j)\}^2 > 0.$$

Hence, to establish the convergence of f.d.d.'s of $\{W_n\}$ to those of W, it suffices to show that the right hand side of (3.3) has asymptotically a normal distribution. For this, we appeal to a special central limit theorem in Hájek and Šidák (1967, page 153), and thereby, we require only to show that

$$\text{(3.4)} \qquad \lim_{n\to\infty} \max_{1 \leq i \leq n} c_{ni}^2 = 0.$$

Since $\lambda'\lambda < \infty$, we obtain from (3.3) that for every i, $1 \leq i \leq n$,

$$\text{(3.5)} \qquad c_{ni}^2 \leq r(\lambda'\lambda)(\textstyle\sum_{s=i}^{k_j} a_{si})^2/(n-m) \leq \{rn/(n-m)\}(\lambda'\lambda)\{n^{-1}(\sum_{s=i}^n |a_{si}|)^2\},$$

so that (2.9) insures (3.4). To establish the tightness of $\{W_n\}$, note that for every $(m \leq) k_1 < k_2 < k_3\ (\leq n)$, by virtue of (2.7) and (2.8), we may write

$$\text{(3.6)} \qquad \textstyle\sum_{i=k_j+1}^{k_{j+1}} V_i = \sum_{i=k_j+1}^{k_{j+1}} \mathbf{a}_i \mathbf{U}_n = \mathbf{b}_{nj}\mathbf{U}_n, \quad \text{say } (j = 1, 2)$$

where

$$\text{(3.7)} \qquad \mathbf{b}_{nj}\mathbf{b}_{nj}' = (k_{j+1} - k_j), j = 1, 2 \quad \text{and} \quad \mathbf{b}_{n1}\mathbf{b}_{n2}' = 0, \text{ by (2.8)}.$$

Further, note that $E(U_{i_1}U_{i_2}U_{i_3}U_{i_4}) = 0$ whenever at least one of the indices i_1, i_2, i_3, i_4 occurs with multiplicity 1. Hence, we obtain by some routine steps that

$$\text{(3.8)} \qquad \begin{aligned} E\{(\textstyle\sum_{i=k_1+1}^{k_2} V_i)^2(\sum_{i=k_2+1}^{k_3} V_i)^2\}/(\sigma^4(n-m)^2) &\leq (\nu_4/\sigma^4)(n-m)^{-2}(\textstyle\sum_{i=1}^n b_{1i}^2)(\sum_{j=1}^n b_{2j}^2) \\ &\quad + 3(n-m)^{-2}\{(\textstyle\sum_{i=1}^n b_{1i}^2)(\sum_{j=1}^n b_{2j}^2)\} \\ &= \{(\nu_4/\sigma^4) + 3\}(k_3 - k_2)(k_2 - k_1)/(n-m)^2. \end{aligned}$$

Since $W_n(0) = 0$, with probability 1, (3.1), (3.8) and Theorem 15.6 of Billingsley (1968) insure the tightness of W_n. □

PROOF OF THEOREM 2. The convergence of the f.d.d.'s of $\{W_n\}$ to those of W follows as in Theorem 1, since (2.10) implies (2.9), and hence we need to establish the tightness part only. For this, note that by (3.1),

$$\text{(3.9)} \qquad \begin{aligned} W_n(t) &= \{\sigma(n-m)^{1/2}\}^{-1} \textstyle\sum_{i \leq k_n(t)} a_{ii} U_i + \{\sigma(n-m)^{1/2}\}^{-1} \sum_{i \leq k_n(t)} \sum_{j<i} a_{ij} U_j \\ &= W_{n1}(t) + W_{n2}(t), \quad t \in [0, 1], \quad \text{say}. \end{aligned}$$

Since the $a_{ii}U_i$ are independent r.v.'s and (2.10) insures that $n^{-1} \sum_{i=1}^n a_{ii}^2 \to 1$ and

$n^{-1}\{\max_{1\le i\le n} a_{ii}^2\} \to 0$, as $n \to \infty$, under (2.6) and (2.10), $\{W_{n1}\}$ converges weakly to W; see Problem 1 on page 67 of Billingsley (1968) in this respect. This, in turn insures that $\{W_{n1}\}$ is tight. Thus, it remains to establish the tightness of $\{W_{n2}\}$.

Let $S_k = U_1 + \cdots + U_k$, $k \ge 1$ and $S_0 = 0$. Then, by the Hájek-Rényi inequality, for every $c > 0$, $0 \le a < ½$ and $q \ge 1$,

$$P\{\max_{1\le k\le q}(k/n)^a n^{-1/2}|S_k| > \sigma c\} \le c^{-2}\{n^{-1}\textstyle\sum_{k=1}^{q}(k/n)^{-2a}\} \le c^{-2}\int_0^{q/n} t^{-2a}\,dt = (1-2a)^{-1}c^{-2}(q/n)^{1-2a}. \tag{3.10}$$

Also, $U_i = S_i - S_{i-1}$, $i \ge 1$, so that for every q, $m + 1 < q \le n$,

$$\{\sigma(n-m)^{1/2}\}^{-1}(\textstyle\sum_{i=m+1}^{q}\sum_{j=1}^{i} a_{ij}U_j) = \{(n-m)/n\}^{-1/2}[\sum_{j=1}^{m}\{a_{m+1j} + \sum_{i=m+2}^{q}(a_{ij} - a_{ij+1})\}(S_j/\sigma\sqrt{n}) + \sum_{j=m+1}^{q-1}\{a_{j+1j} + \sum_{i=j+2}^{q}(a_{ij} - a_{ij+1})\}(S_j/\sigma\sqrt{n})],$$

where, by (2.10), $|ia_{ij}| \le c_1$, $\forall\ 1 \le j \le i-1$, $i \ge m+1$, while, by (2.11),

$$|j\textstyle\sum_{i=(j\vee m)+2}^{q}(a_{ij} - a_{ij+1})| \le c_2,\ \forall\ i \ge 1,\ q \ge m+1.$$

Hence, from (3.10) and the above, we have for $q = m + [(n-m)\delta] + 1$, $0 < \delta < 1$,

$$\begin{aligned}\max_{m<k\le q}\{\sigma(n-m)^{1/2}\}^{-1}|\textstyle\sum_{i=m+1}^{k}\sum_{j=1}^{i-1} a_{ij}U_j| &\le c(c_1+c_2)\{(n-m)/n\}^{-1/2}\textstyle\sum_{j=1}^{q} j^{-1}(j/n)^a \\ &\le c(c_1+c_2)\{(n-m)/n\}^{-1/2}\int_0^{q/n} t^{-1+a}\,dt \\ &= c(c_1+c_2)\{(n-m)/n\}^{-1/2}a^{-1}(q/n)^a,\end{aligned} \tag{3.11}$$

with probability greater than

$$1 - (q/n)^{1-2a}c^{-2}(1-2a)^{-1}. \tag{3.12}$$

Thus, for every $\varepsilon > 0$ and $\eta > 0$, there exist a δ, $0 < \delta < 1$, and a sample size $n_0 = n_0(\varepsilon, \eta)$, such that for $q = m + [(n-m)\delta] + 1$, the right hand side of (3.11) is bounded from above by ε, while (3.12) is bounded from below by $1 - \eta$, for every $n \ge n_0$, *i.e.*,

$$P\{\sup_{0\le t\le\delta}|W_{n2}(t)| > \varepsilon\} < \eta, \qquad \forall\ n \ge n_0. \tag{3.13}$$

On the other hand, defining $k_n(s) = k$ and $k_n(t) = q$, for $\delta \le s < t \le 1$, we have by (2.10),

$$\begin{aligned}E\{W_{n2}(t) - W_{n2}(s)\}^2 &= \{\sigma^2(n-m)\}^{-1}E[\{\textstyle\sum_{i=1}^{q-1}(\sum_{j=k\vee i+1}^{q} a_{ji})U_i\}^2] \\ &= (n-m)^{-1}\textstyle\sum_{i=1}^{q-1}(\sum_{j=k\vee i+1}^{q} a_{ji})^2 \\ &\le (n-m)\textstyle\sum_{i=1}^{q-1}(q - k\vee i)\sum_{j=ki+1}^{q} a_{ji}^2 \\ &\le c_1^2(n-m)^{-1}\{k(q-k)\textstyle\sum_{j=k+1}^{q} j^{-2} + \sum_{i=k+1}^{q}(q-i)\sum_{j=i+1}^{q} j^{-2}\} \\ &\le c_1^2\delta^{-1}\{(q-k)/(n-m)\}^2.\end{aligned} \tag{3.14}$$

Hence, using Theorem 12.3 of Billingsley (1968, page 95) along with (3.14), we conclude that for every $\varepsilon > 0$ and $\eta > 0$, and $\delta > 0$ defined by (3.13), there exist a $\rho : 0 < \rho < 1$ and a sample size n_0', such that for every $n \ge n_0'$,

$$P\{\sup_{\delta\le s<t\le s+\rho\le 1}|W_{n2}(t) - W_{n2}(s)| > \varepsilon\} < \eta. \tag{3.15}$$

Then, (3.13) and (3.17) insure the tightness of $\{W_{n2}\}$. □

4. Some applications to CUSUM tests. We shall discuss the role of Theorems 1 and 2 in the context of some CUSUM tests considered in the literature.

Consider first the simple location model, where, in (2.1), $k = 1$, and $x_t = 1$, $\forall\ t$. In this case, $b_r = \bar{Y}_r = r^{-1} \sum_{i=1}^{r} Y_i$, $r \geq 1$ and

$$w_r = (Y_r - \bar{Y}_{r-1})/\{1 + (r-1)^{-1}\}^{1/2} = \{(r-1)Y_r - \textstyle\sum_{i=1}^{r-1} Y_i\}/\{r(r-1)\}^{1/2}, \quad r \geq 2. \tag{4.1}$$

Thus, we have here $m = 1$ and for every $j \geq 2$,

$$a_{jj} = (1 - j^{-1})^{1/2}, \quad a_{ji} = -\{j(j-1)\}^{-1/2}, \quad 1 \leq i < j - 1. \tag{4.2}$$

Now (4.2) insures (2.10), while (2.11) holds with $c_2 = 0$. Thus, by our Theorem 2, we conclude that (3.2) holds under (2.6). Also, $s_n \to \sigma$ in probability under (2.6); see Sen and Puri (1970). Hence for the CUSUM test for a shift in location, under H_0 and (2.6),

$$D_n^+ \to_{\mathscr{D}} \sup_{0\leq t\leq 1} W(t) \quad \text{and} \quad D_n \to_{\mathscr{D}} \sup_{0\leq t\leq 1} |W(t)|. \tag{4.3}$$

A similar result holds for the Cramér-von Mises type statistic based on the recursive residuals in (4.1). This explains the robustness of the CUSUM tests against non-normality of the distribution of the errors; finiteness of the second moment of the e_t suffices.

Consider next the general regression model in (2.1). First, assume that for some $\lambda > ½$,

$$\max_{1\leq k\leq n} \mathbf{x}_k'(\mathbf{X}_{n-1}'\mathbf{X}_{n-1})^{-1}\mathbf{x}_k = O(n^{-\lambda}), \quad \forall\ n \geq m + 1. \tag{4.4}$$

Then, by (2.2) and (2.3), for every $j \geq m + 1$,

$$a_{jj} = \{1 + \mathbf{x}_j'(\mathbf{X}_{j-1}'\mathbf{X}_{j-1})^{-1}\mathbf{x}_j\}^{-1/2}, \tag{4.5}$$

$$a_{ji} = \mathbf{x}_j'(\mathbf{X}_{j-1}'\mathbf{X}_{j-1}')^{-1}\mathbf{x}_i/\{1 + \mathbf{x}_j'(\mathbf{X}_{j-1}'\mathbf{X}_{j-1})^{-1}\mathbf{x}_j\}^{1/2}, \quad i = 1, \cdots, j-1, \tag{4.6}$$

so that

$$a_{ji}^2 \leq \{\mathbf{x}_j'(\mathbf{X}_{j-1}'\mathbf{X}_{j-1})^{-1}\mathbf{x}_i\}^2 \leq \mathbf{x}_j'(\mathbf{X}_{j-1}'\mathbf{X}_{j-1})^{-1}\mathbf{x}_j\mathbf{x}_i'(\mathbf{X}_{j-1}'\mathbf{X}_{j-1})^{-1}\mathbf{x}_i = O(j^{-2\lambda}), \tag{4.7}$$

for every $i: 1 \leq i \leq j - 1$ and $j \geq m + 1$. Further, by (2.9) and (4.7),

$$\begin{aligned} \max_{m\leq k\leq n}\{n^{-1}(\textstyle\sum_{j=k}^{n} |a_{jk}|)^2\} &\leq \max_{m<k\leq n}\{n^{-1}(n-k) \textstyle\sum_{j=k}^{n} a_{jk}^2\} \\ &\leq \max_{m<k\leq n}\{n^{-1}(n-k)(a_{jj}^2 + \textstyle\sum_{j=k+1}^{n} a_{jk}^2)\} \\ &\leq \max_{m<k\leq n}\{1 + O(k^{-2\lambda+1})\} = O(1), \end{aligned} \tag{4.8}$$

as $\lambda > ½$. Hence, (2.9) holds, so that Theorem 1 holds under (4.4) and $\nu_4 < \infty$.

Next, we proceed to relax the condition that $\nu_4 < \infty$. Note that

$$\begin{aligned} |a_{ji} - a_{ji+1}| &= |\{\mathbf{x}_j'(\mathbf{X}_{j-1}'\mathbf{X}_{j-1})^{-1}(\mathbf{x}_i - \mathbf{x}_{i+1})\}/\{1 + \mathbf{x}_j'(\mathbf{X}_{j-1}'\mathbf{X}_{j-1})^{-1}\mathbf{x}_j\}^{1/2}| \\ &\leq |\mathbf{x}_j'(\mathbf{X}_{j-1}'\mathbf{X}_{j-1})^{-1}(\mathbf{x}_i - \mathbf{x}_{i+1})|, \quad \forall\ 1 \leq i \leq j-2, \quad j \geq m+1. \end{aligned} \tag{4.9}$$

Thus, if, for every $j \geq m + 1$ and $1 \leq i \leq j - 2$,

$$|(\mathbf{x}_i - \mathbf{x}_{i+1})'(\mathbf{X}_{j-1}'\mathbf{X}_{j-1})^{-1}\mathbf{x}_j| = O(j^{-2}), \tag{4.10}$$

then, (2.11) holds, while (2.10) holds when (4.4) holds with $\lambda = 1$. As a result, Theorem 2 holds when (4.4) holds with $\lambda = 1$ and (4.10) holds. Under these extra conditions on the a_{ij}, we do not need that $\nu_4 < \infty$. As a simple example, consider the classical polynomial regression model, where $\mathbf{x}_t' = (1, t, \cdots, t^p)$, for some $p (\geq 0)$, $t = 1, \cdots, n$. In this case, (4.4) holds with $\lambda = 1$ and (4.10) holds, so that Theorem 2 applies under (2.6). For this simple model, some alternative tests are due to MacNeill (1978). However, his test statistic is based on the residuals $Y_i - \mathbf{b}_n'\mathbf{x}_i$, $i = 1, \cdots, n$ (which are not the recursive ones) and the invariance principles (considered by him) follow more easily by an appeal to the existing results in Billingsley (1968). By contrast, our results do not follow directly from such theorems. Thus, the current Theorem 2 provides the robustness picture of CUSUM and related tests for polynomial regression models based on the recursive residuals.

We conclude this section with some discussion on some related tests and their robustness properties. For the normal theory location-change model, some CUSUM tests (not based on recursive residuals) are discussed in Sen and Srivastava (1975), while MacNeill (1974) has considered some CUSUM tests for some exponential distributions. Invariance principles for these test statistics follow by more direct applications of the existing results in Billingsley (1968). The same conclusion applies to the CUSUM types tests considered by Schweder (1976): tests based on recursive residuals can be constructed for the shift model and for the asymptotic distribution theory, one need not confine to normal distribution of the errors. For some nonparametric procedures, see Bhattacharya and Frierson (1981), A. K. Sen and Srivastava (1975) and P. K. Sen (1977, 1978).

We have considered the invariance principles without any specific rates of convergence for finite or moderately large sample sizes. Such a rate of convergence depends on the particular design matrix $\mathbf{A}_n$ as well as the underlying distribution of the errors. Any uniformity result over a class of $\mathbf{A}_n$ and/or a class of error distributions may be quite difficult to obtain.

So far, we have considered the case where the null hypothesis H_0 holds, i.e., the U_t all have expectation 0. If $U_t = \xi_t + e_t$, where the e_t are i.i.d r.v.'s and we define $\mu_i = \sum_{j \le i} \xi_j a_{ij}$, $i \ge m + 1$, then, for the $V_i - \mu_i$, $i \ge m + 1$, Theorems 1 and 2 apply. Hence, whenever $\{(\sigma(n-m)^{1/2})^{-1} \mu_{k_n(t)},\, t \in [0, 1]\}$ converges to a smooth function $\gamma = \{\gamma(t), t \in [0, 1]\}$, the asymptotic power of the CUSUM tests can be expressed in terms of the boundary crossing probability of a drifted Wiener process $W + \gamma$. A smooth γ arises typically in the context of Pitman type local alternatives. In general, γ is not so simple as to allow an algebraic expression for this probability. However, the recent results obtained by DeLong (1980) (in a different context) present excellent prospects for an adequate simulation study.

5. Acknowledgement. The author is grateful to Professor D. V. Hinkley for providing him a copy of a bibliography of articles on change points, prepared by him, which had led to some deeper discussion in Section 4.

REFERENCES

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

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

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.

DELONG, D. (1980). Some asymptotic properties of a progessively censored nonparametric test for multiple regression. *J. Multivarate Anal.* **10,** 363–369.

DESHAYES, J. and PICARD, D. (1980). Tests de rapture de regression: comparison asymptotique. Tech. Report, Univ. Paris-Sud.

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

MACNEILL, I. B. (1974). Tests for change of parameter at unknown time and distributions of some related functionals of Brownian motion. *Ann. Statist.* **2** 950–962.

MACNEILL, I. B. (1978). Properties of sequence of partial sums of polynomial regression residuals with applications to tests for change of regression at unknown times. *Ann. Statist.* **6** 422–433.

SCHWEDER, T. (1976). Some optimal methods to detect structural shift or outliers in regression. *J. Amer. Statist. Assoc.* **71** 491–501.

SEN, A. K. and SRIVASTAVA, M. S. (1975). On tests for detecting changes in mean. *Ann. Statist.* **3** 98–108.

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

SEN, P. K. (1978). Invariance principles for linear rank statistics revisited. *Sankhya, Ser. A.* **40** 215–236.

SEN, P. K. and PURI, M. L. (1970). Asymptotic theory of likelihood ratio and rank order tests in some multivariate linear models. *Ann. Math. Statist.* **41** 87–100

SCHOOL OF PUBLIC HEALTH
DEPARTMENT OF BIOSTATISTICS
THE UNIVERSITY OF NORTH CAROLINA
ROSENAN HALL 201 H
CHAPEL HILL, NC 27514