The Annals of Statistics
1978, Vol. 6, No. 2, 422–433

PROPERTIES OF SEQUENCES OF PARTIAL SUMS OF POLYNOMIAL REGRESSION RESIDUALS WITH APPLICATIONS TO TESTS FOR CHANGE OF REGRESSION AT UNKNOWN TIMES[1]

BY IAN B. MACNEILL

University of Western Ontario

Limit processes are obtained for the sequences of partial sums of polynomial regression residuals. Properties of linear and quadratic functionals on the sequences are discussed. Distribution theory for Cramér–von Mises type functionals is obtained. An indication is given of the relevance of these results to the problem of testing for change of regression at unknown times.

1. Introduction and summary. The problem of testing for change of regression at unknown time was first considered by Quandt (1958, 1960) who proposed a test for no change versus one change based upon the likelihood ratio. Hinkley (1969) also discussed the likelihood ratio test and conjectured that the test statistic was approximately distributed as a χ_3^2 variable. A different and more easily applied approach was proposed by Farley, Hinich and McGuire (1970, 1975). Feder (1975), in dealing with the asymptotic distribution of the likelihood ratio statistics in regression models which have different forms in different regions in the domain of the independent variable, finds, in the case of unknown change point, the distribution to be complicated and to depend upon the configuration of the observation points of the independent variable. Quandt (1972) and Goldfield and Quandt (1973) attack a more complicated problem wherein each observation may be randomly chosen to come from one of two regression models.

Brown, Durbin and Evans (1975) propose tests based upon recursively generated residuals. Their test statistics utilize the sequence of partial sums of these residuals which, it turns out, are relatively easy to analyse since they are i.i.d. asymptotically. The widely used statistical computing packages have paid little attention to the computation of recursively generated residuals thus making it awkward to routinely apply the tests for change of regression suggested by Brown et al. However, these same packages nearly always make available printouts and plots of the raw regression residuals. In this paper we consider tests for change of polynomial regression at unknown times which are based on raw regression residuals. We first examine the large sample properties of the sequence of partial

Received March 1976; revised June 1977.

[1] This research was supported by the National Research Council of Canada and was completed while the author was visiting the Statistical Laboratory at the University of Cambridge.

AMS 1970 *subject classifications.* 60J65, 62J05.

Key words and phrases. Regression residuals, weak convergence, Brownian motion, Cramér–von Mises statistic, tests for change of parameters at unknown time.

sums of regression residuals. Limit processes are derived and used to obtain large sample distributions for certain functionals on the sequences of partial sums of residuals, including Cramér–von Mises type functionals. Related to these functionals are tests for change of regression at unknown time. Included are tests analogous to those proposed by Chernoff and Zacks (1964) and by Gardner (1969).

Assume $\{\varepsilon_j\}_{j=1}^{\infty}$ to be a sequence of independent and identically distributed random variables with zero means and variances $\sigma^2 < \infty$. Assume each component of the sequence to be defined on the same probability space, (Ω, A, P). Let $\{(t_{nj})_{j=1}^{n}\}_{n=1}^{\infty}$ be a triangular array of the nonstochastic independent variable and let $\{(Y_{nj})_{j=1}^{n}\}_{n=1}^{\infty}$ be a similar array of the dependent variable whose components are defined by

$$Y_{nj}(\omega) = \sum_{i=0}^{p} \beta_i t_{nj}^i + \varepsilon_j(\omega) \qquad \omega \in \Omega .$$

In the usual matrix formulation this becomes

$$\mathbf{Y}_n(\omega) = \mathbf{X}_n \boldsymbol{\beta}_p + \boldsymbol{\varepsilon}_n(\omega)$$

where the r, sth component of the design matrix is t_{nr}^s. The Gauss–Markov estimator for $\boldsymbol{\beta}_p$ is denoted by $\hat{\boldsymbol{\beta}}_{pn}(\omega)$ and is defined by

$$\hat{\boldsymbol{\beta}}_{pn}(\omega) = (\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{X}_n'\mathbf{Y}_n(\omega) ,$$

it being understood that the inverse exists. We suppress ω and the subscripts on the vectors and matrices where no confusion results. Our purpose is to discuss some large sample properties of

$$[\{\textstyle\sum_{i=1}^{j} (Y_{ni} - \hat{Y}_{ni})\}_{j=1}^{n}]_{n=1}^{\infty} \qquad \text{where} \quad \hat{Y}_{ni} = \hat{\boldsymbol{\beta}}'_{pn}\mathbf{t}_{ni},\ \mathbf{t}'_{ni} = (1, t_{ni}, t_{ni}^2, \cdots, t_{ni}^p) \quad \text{and} \quad t_{ni} = i/n .$$

2. Limit process for the sequence of partial sums of regression residuals. Let $S_j(\omega) = \sum_{i=1}^{j} \varepsilon_i(\omega)$ and define a sequence of stochastic processes $\{\theta_n(t), t \in [0, 1]\}_{n=1}^{\infty}$ possessing continuous sample paths by

$$\theta_n(t, \omega) = \frac{1}{\sigma n^{\frac{1}{2}}} S_{[nt]}(\omega) + (nt - [nt]) \frac{1}{\sigma n^{\frac{1}{2}}} \varepsilon_{[nt]+1}(\omega) ,$$

it being understood that $S_0(\omega) \equiv 0$. If $\mathbf{e}_{jn}$ is an $n \times 1$ vector whose first j components are 1 and the remainder zero then one can write

$$\textstyle\sum_{i=1}^{j} (Y_{ni} - \hat{Y}_{ni}) = \mathbf{e}_j'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\boldsymbol{\varepsilon}_n .$$

Whether one uses monic polynomials or orthogonal polynomials to define the components of $\mathbf{X}$ the projection $\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ assumes the same values. Hence we shall use orthogonal polynomials in the sequel since $\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ is then more easily evaluated.

For the case of observations taken at $j = 1, 2, \cdots, n$, Allan (1930), defining $\xi_j = (j - (n + 1)/2)$, shows that the orthogonal polynominal of degree m evaluated at j is

$$\phi_{n,m}(j) = \xi_j \sum_{q=0}^{[m/2]} \frac{(-1)^q \binom{m}{q,\,q,\,m-2q} \binom{\frac{1}{2}(n+m-1)}{q} \binom{\frac{1}{2}(n-m-1)+q}{q} (\xi_j + m/2 - 1 - q)!\,(q!)^2}{2^{2q}\binom{m-\frac{1}{2}}{q}(\xi_j - m/2 + q)!}$$

where

$$\binom{a}{b_1, b_2, \cdots, b_r} = \frac{a!}{b_1!\, b_2! \cdots b_r!}\,.$$

Since

$$\binom{\frac{1}{2}(n+m-1)}{q} = \frac{n^q}{2^q q!}\left\{1 + O\left(\frac{1}{n}\right)\right\},$$

$$\binom{\frac{1}{2}(n-m-1)+q}{q} = \frac{n^q}{2^q q!}\left\{1 + O\left(\frac{1}{n}\right)\right\}$$

and

$$\frac{\xi_j(\xi_j + m/2 - 1 - q)!}{(\xi - m/2 + q)!} = \xi_j^{m-2q}\left\{1 + O\left(\frac{1}{n^2}\right)\right\},$$

it follows that

$$(1) \qquad \phi_{n,m}(j) = n^m \sum_{q=0}^{[m/2]} \frac{(-1)^q \binom{m}{q,q,m-2q}}{2^{4q}\binom{m-\frac{1}{2}}{q}} \left(\frac{j}{n} - \frac{1}{2}\right)^{m-2q} (1 + K_{qn})$$

where K_{qn}, here and in the sequel, is $O(1/n)$. The mth component of $\mathbf{e}_r'\mathbf{X}$ is $\sum_{j=1}^r \phi_{nm}(j)$. Thus, if $t = r/n$, we have

$$(2) \qquad \sum_{j=1}^r \phi_{n,m}(j) = n^{m+1} \sum_{q=0}^{[m/2]} \frac{(-1)^q \binom{m}{q,q,m-2q}}{2^{4q}\binom{m-\frac{1}{2}}{q}} (1 + K_{qn}) \sum_{j=1}^r \frac{1}{n} (j/n - \tfrac{1}{2})^{m-2q}$$

$$= n^{m+1} \sum_{q=0}^{[m/2]} \frac{(-1)^q \binom{m}{q,q,m-2q}}{2^{4q}\binom{m-\frac{1}{2}}{q}} (1 + K_{qn}) \int_0^t (s - \tfrac{1}{2})^{m-2q}\, ds\,.$$

The mth diagonal component of $\mathbf{X}'\mathbf{X}$ is shown by Allen (1930) to be $d_{m,n}$ where

$$(3) \qquad d_{m,n} = \frac{(m!)^4}{(2m)!\,(2m+1)!} n(n^2-1)(n^2-4)\cdots(n^2-m^2)$$

$$= \frac{(m!)^4}{(2m)!\,(2m+1)!} n^{2m+1}\left\{1 + O\left(\frac{1}{n^2}\right)\right\}.$$

The mth component of $\mathbf{X}'\boldsymbol{\varepsilon}$ is $\sum_{j=1}^n \phi_{n,m}(j)\varepsilon_j$. It follows from (1) that

$$(4) \qquad \sum_{j=1}^n \phi_{n,m}(j)\varepsilon_j = n^m \sum_{q=0}^{[m/2]} \frac{(-1)^q \binom{m}{q,q,m-2q}}{2^{4q}\binom{m-\frac{1}{2}}{q}} (1 + K_{qn}) \sum_{j=1}^n (j/n - \tfrac{1}{2})^{m-2q}\varepsilon_j\,.$$

Now

$$\sum_{j=1}^n (j/n - \tfrac{1}{2})^{m-2q}\varepsilon_j = \sum_{k=0}^{m-2q} \binom{m-2q}{k}(-\tfrac{1}{2})^{m-2q-k} \sum_{j=1}^n (j/n)^k \varepsilon_j$$

and, for $k > 0$,

$$\sum_{j=1}^n (j/n)^k \varepsilon_j = \sum_{j=1}^n \left(\frac{1}{n}\right)^k \sum_{i=1}^{j^k} \varepsilon_j = \sum_{i=1}^{n^k} \left(\frac{1}{n}\right)^k \sum_{j=[(i-1)^{1/k}]+1}^n \varepsilon_j$$

$$= \sum_{i=1}^{n^k} \frac{1}{n^k} (S_n - S_{[(i-1)^{1/k}]})$$

$$= S_n - \sigma n^{\frac{1}{2}}(1 + K_{qn}) \int_0^1 \theta_n(s^{1/k})\, ds$$

$$= S_n - \sigma n^{\frac{1}{2}}(1 + K_{qn}) k \int_0^1 s^{k-1}\theta_n(s)\, ds\,.$$

When $k = 0$ the sum reduces to S_n. This implies that

$$(5)\qquad \sum_{j=1}^{n} (j/n - \tfrac{1}{2})^{m-2q}\varepsilon_j = S_n(\tfrac{1}{2})^{m-2q} - \sigma n^{\frac{1}{2}}(1 + K_{qn})(m - 2q) \int_0^1 (s - \tfrac{1}{2})^{m-2q-1}\theta_n(s)\, ds\,.$$

Substitution of (5) into (4) gives

$$(6)\qquad \begin{aligned} \frac{1}{\sigma n^{\frac{1}{2}}} \sum_{j=1}^{n} \phi_{n,m}(j)\varepsilon_j &= n^m \sum_{q=0}^{[m/2]} \frac{(-1)^q\binom{m}{q,q,m-2q}}{2^{4q}\binom{m-\frac{1}{2}}{q}} (1 + K_{qn}) \\ &\quad \times \{(\tfrac{1}{2})^{m-2q}\theta_n(1) - (m - 2q)(1 + K_{qn}) \int_0^1 (s - \tfrac{1}{2})^{m-2q-1}\theta_n(s)\, ds\}\,. \end{aligned}$$

Equations (2), (3) and (6) yield

$$(7)\qquad \begin{aligned} &\frac{1}{\sigma n^{\frac{1}{2}}} \mathbf{e}_r'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\boldsymbol{\varepsilon}_n(\omega) \\ &= \theta_n(r/n, \omega) - \frac{1}{\sigma n^{\frac{1}{2}}} \sum_{m=0}^{p} d_{m,n}^{-1} \sum_{j=1}^{r} \phi_{n,m}(j) \sum_{j'=1}^{n} \phi_{n,m}(j')\varepsilon_{j'}(\omega) \\ &= \theta_n(r/n, \omega) - \sum_{m=0}^{p} (2m + 1) \\ &\quad \times \left\{\sum_{q=0}^{[m/2]} \frac{(-1)^q\binom{2m}{m,q,q,m-2q}}{2^{4q}\binom{m-\frac{1}{2}}{q}} (1 + K_{qn}) \int_0^{r/n} (s - \tfrac{1}{2})^{m-2q}\, ds\right\} \\ &\quad \times \left\{\sum_{q=0}^{[m/2]} \frac{(-1)^q\binom{2m}{m,q,q,m-2q}}{2^{4q}\binom{m-\frac{1}{2}}{q}} (1 + K'_{qn}) \right. \\ &\quad \left. \times [(\tfrac{1}{2})^{m-2q}\theta_n(1, \omega) - (m - 2q) \int_0^1 (s - \tfrac{1}{2})^{m-2q-1}\theta_n(s, \omega)\, ds]\right\}\,. \end{aligned}$$

For the case $p = 0$ and $t = r/n$ one obtains

$$\frac{1}{\sigma n^{\frac{1}{2}}} \sum_{i=1}^{r} \{Y_{ni}(\omega) - \hat{Y}_{ni}(\omega)\} = \theta_n(t, \omega) - t\theta_n(1, \omega)\,.$$

For the case $p = 1$ one obtains

$$\begin{aligned} \frac{1}{\sigma n^{\frac{1}{2}}} \sum_{i=1}^{r} (Y_{ni}(\omega) - \hat{Y}_{ni}(\omega)) &= \theta_n(t, \omega) - t\theta_n(1, \omega) + 3t(1 - t) \frac{n^2}{n^2 - 1} \theta_n(1, \omega) \\ &\quad - 6t(1 - t) \frac{n^2}{n^2 - 1} \int_0^1 \theta_n(s, \omega)\, ds \end{aligned}$$

thus indicating the precise nature of K_{qn} for $p = 1$. By letting $W_{pn}(t, \omega)$ equal the right-hand side of (7) with r/n replaced by the continuous variable t one sees that $W_{pn}(\cdot)$ is a function of $\theta_n(\cdot)$. In this way, for each $p = 0, 1, 2, \cdots$, the relation $W_{pn}(t, \omega) = h_{pn}\{\theta_n(t, \omega)\}$ defines a sequence of stochastic processes $\{W_{pn}(t, \omega), t \in [0, 1], \omega \in \Omega\}_{n=1}^{\infty}$ with continuous simple paths and a sequence of functions $\{h_{pn}(\cdot)\}_{n=1}^{\infty}$ from $C[0, 1]$ into $C[0, 1]$.

Let P_n be the distribution of $\theta_n(\cdot)$ on $C[0, 1]$ and let P_{pn} be the probability measure generated in $C[0,1]$ by $W_{pn}(\cdot)$. Now denote by $\{B(t), t \in [0, 1]\}$ the standard Brownian motion process with continuous sample paths. The process is a measurable map from some probability space to $C[0, 1]$. Such a process is Gaussian with zero mean and has $B(0) = 0$ and $E[B(t)B(s)] = \min(s, t)$. The measure generated in $C[0, 1]$ by Brownian motion is Wiener measure denoted by W. Next define another process $\{B_p(t), t \in [0, 1]\}$ called a generalized Brownian bridge and a function $h_p(\cdot)$ from $C[0, 1]$ into $C[0, 1]$ by

$$\begin{aligned} B_p(t) &= h_p\{B(t)\} \\ &\equiv B(t) - \sum_{m=0}^{p} (2m+1) \left\{ \sum_{q=0}^{[m/2]} \frac{(-1)^q \binom{2m}{m,q,q,m-2q}}{2^{4q}\binom{m-\frac{1}{2}}{q}} \int_0^t (s - \tfrac{1}{2})^{m-2q}\, ds \right\} \\ &\quad \times \left\{ \sum_{q=0}^{[m/2]} \frac{(-1)^q \binom{2m}{m,q,q,m-2q}}{2^{4q}\binom{m-\frac{1}{2}}{q}} \right. \\ &\quad \left. \times \left[(\tfrac{1}{2})^{m-2q} B(1) - (m-2q) \int_0^1 (s - \tfrac{1}{2})^{m-2q-1} B(s)\, ds \right] \right\}. \end{aligned} \tag{8}$$

The "slack" in the generalized Brownian bridge becomes less as p increases. For $p = 0$, (8) yields the Brownian bridge

$$B_0(t) = h_0\{B(t)\} \equiv B(t) - tB(1)$$

and for $p = 1$, (8) yields

$$B_1(t) = h_1\{B(t)\} \equiv B(t) - tB(1) + 6t(1-t)\{\tfrac{1}{2}B(1) - \textstyle\int_0^1 B(s)\, ds\}.$$

The generalized Brownian bridge is Gaussian and has $B_p(0) = B_p(1) = E\{B_p(t)\} = 0$. The covariance kernel is given by

$$\begin{aligned} K_p(s, t) &= E\{B_p(s)B_p(t)\} \\ &= \min(s, t) - \textstyle\sum_{m=0}^{p} (2m+1) g_m(t) g_m(s) \end{aligned}$$

where

$$g_m(t) = \sum_{q=0}^{[m/2]} \frac{(-1)^q \binom{2m}{m,q,q,m-2q}}{2^{4q}\binom{m-\frac{1}{2}}{q}} \int_0^t (t' - \tfrac{1}{2})^{m-2q}\, dt'. \tag{9}$$

The first six values of $g_m(t)$ are given in Table 1.

TABLE 1
Selected values for $g_m(t)$ which defines the m*th term of the covariance kernel of $B_p(t)$*

m	$g_m(t)$
0	t
1	$-t(1-t)$
2	$t(1-t)(1-2t)$
3	$-t(1-t)(1-5t+5t^2)$
4	$t(1-t)(1-2t)(1-7t+7t^2)$
5	$-t(1-t)(1-14t+56t^2-84t^3+42t^4)$

If the probability measure generated in $C[0, 1]$ by $B_p(\cdot)$ is denoted by W_p then we have:

THEOREM 1. *P_{pn} converges weakly to W_p.*

PROOF. Note that for each p (i) $P_{pn} = P_n h_{ph}^{-1}$ and (ii) $W_p = Wh_p^{-1}$. Now, from Theorem 10.1 (Billingsley (1968)), we have that P_n converges weakly to W. Since $h_{pn}(\cdot)$ and $h_p(\cdot)$ are continuous in the uniform topology on $C[0, 1]$, and $h_{pn}(\cdot)$ converges to $h_p(\cdot)$ (in the sense that if $\{x_n\}_{n=1}^{\infty}$ and x are elements of $C[0, 1]$ and x_n converges uniformly to x, then $h_{pn}(x_n)$ converges uniformly to $h_p(x)$), it follows from Theorem 5.5 (Billingsley (1968)) that P_{pn} converges weakly to W_p.

3. Linear and quadratic functionals of sequences of partial sums of regression residuals. We determine the conditions under which certain functionals on the sequences of partial sums converge in distribution to the same functionals on the generalized Brownian bridge. We first consider functionals of a quadratic or Cramér–von Mises type.

THEOREM 2. *Assume $\psi(\cdot)$ to be a nonnegative weight function such that $\int_0^1 t(1-t)\psi(t)\,dt < \infty$. If $F_\psi(f) = \int_0^1 \psi(t) f^2(t)\,dt$ then*

$$\lim_{n\to\infty} P[F_\psi(W_{pn}) \leqq \alpha] = P[F_\psi(B_p) \leqq \alpha] \tag{10}$$

uniformly in α.

PROOF. In the event that the functional $F_\psi(\cdot)$ was continuous in the uniform topology on $C[0, 1]$, (10) could be justified by first appealing to Theorem 1 and secondly to the Donsker result as embodied in Theorem 5.2 of Billingsley (1968). However, certain of the functionals satisfying the hypotheses of the theorem are discontinuous, e.g., $\psi(t) = [t(1-t)]^{-1}$. The proof of the result for the more general case requires weight functions $\psi_\eta(\cdot)$ defined as follows. Let $0 < \eta < \frac{1}{2}$. Then define

$$\begin{aligned} \psi_\eta(t) &= \psi(t) \quad && \eta \leqq t \leqq 1-\eta \\ &= 0 && \text{otherwise.} \end{aligned}$$

$F_{\psi_\eta}(\cdot)$ is continuous. A consequence of Theorem 1 and the aforementioned Donsker result is that for $\varepsilon > 0$ one can choose $N_{\varepsilon\eta}$ such that if $n > N_{\varepsilon\eta}$ then, uniformly in α,

$$|\Pr\{F_{\psi_\eta}(W_{pn}) \leqq \alpha\} - \Pr\{F_{\psi_\eta}(B_p) \leqq \alpha\}| < \varepsilon\,. \tag{11}$$

$\{N_{\varepsilon\eta}, 0 < \eta < \frac{1}{2}\}$ can be chosen to be nondecreasing in the sense that if $\eta_1 < \eta_2$ then $N_{\varepsilon\eta_1} \geqq N_{\varepsilon\eta_2}$. It can next be observed that

$$\begin{aligned} E|F_{\psi_\eta}(B_p) - F_\psi(B_p)| &= \textstyle\int_0^\eta + \int_{1-\eta}^1 \psi(t) E\{B_p^2(t)\}\,dt \\ &= \textstyle\int_0^\eta + \int_{1-\eta}^1 \psi(t) t(1-t) G_p(t)\,dt \end{aligned}$$

where $G_p(t)$ is a polynomial of degree $2p$ since $t = 0$ and $t = 1$ are zeros of $K_p(t, t)$. Hence, for $\varepsilon > 0$ one can select $\eta_{1\varepsilon} > 0$ sufficiently small that if $0 < \eta < \eta_{1\varepsilon}$ then

$$|\Pr\{F_{\psi_\eta}(B_p) \leqq \alpha\} - \Pr\{F_\psi(B_p) \leqq \alpha\}| < \varepsilon \tag{12}$$

uniformly in α. It can also shown that there exists $K > 0$, independent of n such that

$$E|F_\psi(W_{pn}) - F_{\psi_\eta}(W_{pn})| = \int_0^\eta + \int_{1-\eta}^1 \psi(t)E\{W_{pn}^2(t)\}\,dt$$
$$\leqq K \int_0^\eta + \int_{1-\eta}^1 \psi(t)K_p(t, t)\,dt\,.$$

Consequently, for $\varepsilon > 0$, one can select $\eta_{2\varepsilon} > 0$ sufficiently small that if $0 < \eta < \eta_{2\varepsilon}$ then

$$|\Pr\{F_\psi(W_{pn}) \leqq \alpha\} - \Pr\{F_{\psi_\eta}(W_{pn}) \leqq \alpha\}| < \varepsilon \tag{13}$$

uniformly in n and α. The proof of the theorem is completed by combining (11), (12) and (13). For $\varepsilon > 0$, first choose $\eta < \min(\eta_{1\varepsilon}, \eta_{2\varepsilon})$ and then select $N_{\varepsilon\eta}$ such that if $n > N_{\varepsilon\eta}$

$$|\Pr\{F_\psi(W_{pn}) \leqq \alpha\} - \Pr\{F_\psi(B_p) \leqq \alpha\}| < 3\varepsilon$$

uniformly in α.

We can now exploit this result to demonstrate the relationship between the functional on sequences of partial sums and on the generalized Brownian bridge.

COROLLARY 3. *Assume*

$$R_n(j/n) = \int_{(2j-1)/2n}^{(2j+1)/2n} \psi(t)\,dt \qquad j = 1, 2, \cdots, n-1$$

where $\psi(\cdot)$ *satisfies the conditions of Theorem* 2. *Then*

$$\sum_{j=1}^{n-1} R_n(j/n) \left\{\frac{1}{\sigma n^{\frac{1}{2}}} \sum_{i=1}^{j} (Y_{ni} - \hat{Y}_{ni})\right\}^2$$

converges in distribution to $\int_0^1 \psi(t)B_p^2(t)\,dt$.

PROOF. Since $(\sigma n^{\frac{1}{2}})^{-1} \sum_{i=1}^{j} (Y_{ni} - \hat{Y}_{ni}) \equiv W_{pn}(j/n)$, we have

$$\int_0^1 \psi(t)W_{pn}^2(t)\,dt - \sum_{j=1}^{n-1} R_n(j/n) \left\{\frac{1}{\sigma n^{\frac{1}{2}}} \sum_{i=1}^{j} (Y_{ni} - \hat{Y}_{ni})\right\}^2$$
$$= \sum_{j=0}^{n-1} \left\{\int_{j/n}^{j/n+\frac{1}{2}n} \psi(t)(W_{pn}^2(t) - W_{pn}^2(j/n))\,dt \right.$$
$$\left. + \int_{j/n+\frac{1}{2}n}^{(j+1)/n} \psi(t)\left(W_{pn}^2(t) - W_{pn}^2\left(\frac{j+1}{n}\right)\right)dt\right\}.$$

We note that

$$E|W_{pn}^2(t) - W_{pn}^2(j/n)| \leqq [E\{W_{pn}(t) + W_{pn}(j/n)\}^2 E\{W_{pn}(t) - W_{pn}(j/n)\}^2]^{\frac{1}{2}}\,.$$

The dominant term in $W_{pn}(t) - W_{pn}(j/n)$ is $\{\theta_n(t) - t\theta_n(1)\} - \{\theta_n(j/n) - j/n\theta_n(1)\}$. Using arguments similar to those of MacNeill (1974), page 973, with $m = n$ one can show that, for $\varepsilon > 0$,

$$\lim_{n\to\infty} P[\int_0^1 \psi(t)[\{\theta_n(t) - t\theta_n(1)\} - \{\theta_n(j/n) - j/n\theta_n(1)\}]^2 \geqq \varepsilon] = 0\,.$$

[Note the misprint in this article; an asterisk should appear on Z_m in equation (7) and in the equation above it.] The proof is completed by applying similar arguments to the remaining terms in $\{W_{pn}(t) - W_{pn}(j/n)\}^2$ and then applying Theorem 2.

A result, analogous to Theorem 2 and its corollary, for certain linear functionals is given in Theorem 4. The proof of this result is similar to the preceding proofs.

THEOREM 4. *Let $\psi(\cdot)$ be a nonnegative weight function such that $\int_0^1 \{t(1-t)\}^{\frac{1}{2}}\psi(t)\,dt < \infty$. If $R_n(j/n), j = 1, 2, \cdots, n-1$ are defined as in Corollary 3, then $\sum_{j=1}^{n-1} R_n(j/n)\{(\sigma n^{\frac{1}{2}})^{-1} \sum_{i=1}^{j} (Y_{ni} - \hat{Y}_{ni})\}$ converges in distribution to $\int_0^1 \psi(t)B_p(t)\,dt$ which is distributed as a normal variable with zero mean and variance τ_p^2 where*

$$\tau_p^2 = \int_0^1 \int_0^1 \psi(s)\psi(t) \min(s, t)\,ds\,dt - \sum_{m=0}^{p} (2m+1)\{\int_0^1 g_m(s)\psi(s)\,ds\}^2$$

with $g_m(s)$ defined by (9).

4. Distributions for Cramér–von Mises type statistics. The distributions for the stochastic integral, $\int_0^1 \psi(t)B_p^2(t)\,dt$, can, in theory, be calculated by applying the method that Anderson and Darling (1952) applied to the Brownian bridge. Consider the process $\{(\psi(t))^{\frac{1}{2}}B_p(t),\ t \in (0, 1)\}$. The method consists of expanding this process by computing a set of orthonormal functions $\{\phi_{pn}(\cdot)\}_{n=1}^{\infty}$ and a set of zero mean, uncorrelated, normal random variables $\{b_{pn}\}_{n=1}^{\infty}$ with $\mathrm{Var}(b_{pn}) = \lambda_{pn}$ such that

$$\lim_{n\to\infty} E\{(\psi(t))^{\frac{1}{2}}B_p(t) - \sum_{k=1}^{n} b_{pk}\phi_{pk}(t)\}^2 = 0\,.$$

If $\psi(\cdot)$ satisfies the condition of Theorem 2 and those stated by Anderson and Darling (1952), page 199, then the characteristic function for $\int_0^1 \psi(t)B_p^2(t)\,dt$ is

$$\Phi_{p\psi}(s) = \prod_{n=1}^{\infty} (1 - 2\,is\lambda_{pn})^{-\frac{1}{2}}\,.$$

We now consider the special case of $\psi(t) \equiv 1$.

THEOREM 5. *The characteristic function of $\int_0^1 B_p^2(t)\,dt$ is*

$$\Phi_p(s) = \left\{\frac{4\Gamma(p+\frac{1}{2})\Gamma(p+\frac{3}{2})}{\pi(\frac{1}{2}(is/2)^{\frac{1}{2}})^{2p-1}}\, j_{p-1}\left(\frac{is}{2}\right)^{\frac{1}{2}} j_p\left(\frac{is}{2}\right)^{\frac{1}{2}}\right\}^{-\frac{1}{2}} \tag{14}$$

where $j_p(\cdot)$ is the pth order spherical Bessel function of the first kind.

PROOF. If the representation and the orthogonality conditions are to hold simultaneously then $\{\phi_{pn}(t), \lambda_{pn}\}_{n=1}^{\infty}$ must satisfy the Fredholm equation

$$\int_0^1 K_p(s, t)\phi_p(s)\,ds = \lambda_p\,\phi_p(t) \tag{15}$$

where $K_p(s, t)$ is given by (9). The eigenvalues satisfying (15) are found to be

$$\lambda_{p,2n-1} = \frac{1}{4Z_{p-1,n}^2}$$

$$\lambda_{p,2n} = \frac{1}{4Z_{p,n}^2} \qquad n = 1, 2, \cdots$$

where $Z_{p,n}$ is the nth positive zero of the pth order spherical Bessel function of the first kind. These eigenvalues are found by differentiating (15) $p+2$ times thus obtaining the differential equation,

$$\lambda\phi_p^{(p+2)}(t) + \phi_p^{(p)}(t) = 0\,, \tag{16}$$

together with $p+2$ conditions that the solution must satisfy. The conditions, along with the orthonormality requirement for the eigenfunctions, enable one to determine the constants in the general solution and to obtain the eigenvalues. For example, consider the case $p = 1$. The solution to (16) is

$$\phi_1(t) = A + Be^{-it/\lambda^{\frac{1}{2}}} + Ce^{2t/\lambda^{\frac{1}{2}}}$$

subject to $\phi_1(0) = \phi_1(1) = 0$ and

$$6 \int_0^1 s(1-s)\phi_1(s)\,ds = \lambda\phi_1''(1)\,. \tag{17}$$

$\phi_1(0) = 0$ implies

$$\phi_1(t) = \{Ce^{it/(2\lambda^{\frac{1}{2}})} - Be^{-it/(2\lambda^{\frac{1}{2}})}\}2i\sin\{t/(2\lambda^{\frac{1}{2}})\}$$

and $\phi_1(1) = 0$ implies that either $\sin\{1/(2\lambda^{\frac{1}{2}})\} = 0$, and hence

$$\lambda_{1,2n-1} = (4n^2\pi^2)^{-1} \qquad n = 1, 2, \cdots,$$

or $C = Be^{-i/\lambda^{\frac{1}{2}}}$, and hence

$$g(t) = B\{e^{i(s-1)/\lambda^{\frac{1}{2}}} + e^{-is/\lambda^{\frac{1}{2}}} - e^{-i/\lambda^{\frac{1}{2}}} - 1\}\,.$$

Substituting $g(t)$ into (17) yields

$$\lambda_{1,2n} = (4Z_{1,n}^2)^{-1} \qquad n = 1, 2, \cdots$$

where $Z_{1,n}$ is the nth positive zero of the first order spherical Bessel function of the first kind.

Replacing $B_p(\cdot)$ with the expansion in the stochastic integral we see that $\int_0^1 B_p^2(t)\,dt$ is distributed as $\sum_{n=1}^\infty b_{pn}^2$ and hence has characteristic function

$$\Phi_p(s) = \prod_{n=1}^\infty (1 - 2is\lambda_{pn})^{-\frac{1}{2}} = \prod_{n=1}^\infty \left(1 - \frac{is}{2Z_{p-1,n}^2}\right)^{-\frac{1}{2}} \prod_{n=1}^\infty \left(1 - \frac{is}{2Z_{p,n}^2}\right)^{-\frac{1}{2}}.$$

We obtain the result by noting that the infinite product representation of $j_p(t)$ is

$$j_p(t) = \frac{\pi^{\frac{1}{2}}(\frac{1}{2}t)^p}{2\Gamma(p+\frac{3}{2})} \prod_{n=1}^\infty \left(1 - \frac{t^2}{Z_{p,n}^2}\right).$$

The first few Bessel functions are

$$j_{-1}(t) = \frac{\cos t}{t}\,,$$

$$j_0(t) = \frac{\sin t}{t}\,,$$

$$j_1(t) = \frac{\sin t}{t^2} - \frac{\cos t}{t}\,,$$

$$j_2(t) = \left(\frac{3}{t^3} - \frac{1}{t}\right)\sin t - \frac{3}{t^2}\cos t\,,$$

and others may be obtained from the formula

$$j_{p+1}(t) = \frac{p}{t} j_p(t) - j_p'(t)\,.$$

If one lets $p = 0$ in (14) one obtains the result given by Anderson and Darling (1952):

$$\Phi_0(s) = \left(\frac{(2is)^{\frac{1}{2}}}{\sin (2is)^{\frac{1}{2}}}\right)^{\frac{1}{2}}.$$

If one lets $p = 1$ one obtains

$$\Phi_1(s) = \frac{is}{2\cdot 3^{\frac{1}{2}}}\left(\sin^2\left(\frac{is}{2}\right)^{\frac{1}{2}} - \left(\frac{is}{2}\right)^{\frac{1}{2}}\sin\left(\frac{is}{2}\right)^{\frac{1}{2}}\cos\left(\frac{is}{2}\right)^{\frac{1}{2}}\right)^{-\frac{1}{2}}.$$

If $D_p(2is) = (\Phi_p(s))^{-2}$ and $\lambda = 2is$ then $\Omega_p(\alpha) = P[\int_0^1 B_p^2(t)\,dt \leqq \alpha]$ is given by

$$\Omega_p(\alpha) = 1 - \frac{1}{\pi}\sum_{n=1}^{\infty}(-1)^{n-1}\int_{1/\lambda_{2n-1}}^{1/\lambda_{2n}} \frac{e^{-(\lambda/2)\alpha}}{\lambda(-D_p(\lambda))^{\frac{1}{2}}}\,d\lambda.$$

TABLE 2

Selected quantiles for $\Omega_p(\alpha) = P[\int_0^1 B_p^2(t)\,dt \leqq \alpha]$

Probabilities	p						Brownian motion
	0	1	2	3	4	5	
0.01	0.024798	0.017269	0.013799	0.011652	0.010156	0.009040	0.0345
0.025	0.030351	0.020260	0.015867	0.013224	0.011416	0.010087	0.0444
0.05	0.036562	0.023409	0.017986	0.014807	0.012671	0.011118	0.0565
0.10	0.046015	0.027886	0.020911	0.016954	0.014349	0.012485	0.0765
0.50	0.118880	0.055548	0.037513	0.028527	0.023079	0.019404	0.2905
0.90	0.347305	0.119220	0.071460	0.050559	0.038875	0.031446	1.1958
0.95	0.461361	0.147891	0.085955	0.059658	0.045243	0.036208	1.6557
0.975	0.580614	0.177468	0.100670	0.068799	0.051590	0.040925	2.1347
0.99	0.743458	0.217746	0.120482	0.081002	0.060010	0.047150	2.7875
Mean	$\frac{1}{6}$	$\frac{1}{15}$	$\frac{3}{70}$	$\frac{2}{63}$	$\frac{5}{198}$	$\frac{3}{143}$	$\frac{1}{2}$

Selected quantiles for this distribution for $p = 0, 1, 2, 3, 4, 5$ are given in Table 2. The case $p = 0$ was (essentially) obtained by Smirnov (1936).

The nth cumulant of the stochastic integral $\int_0^1 B_p^2(t)\,dt$ is given by

$$K_{pn} = 2^{n-1}(n-1)!\sum_{k=1}^{\infty}\lambda_{pk}^{n}. \tag{18}$$

To facilitate computation of the cumulants we note the following formulae which are derived from formulae given by Watson (1944), page 502:

$$\begin{aligned}\sum_{m=1}^{\infty}(Z_{p,m})^{-2n} &= \{2(2p+3)\}^{-1} && n = 1\\ &= \{2(2p+3)^2(2p+5)\}^{-1} && n = 2\\ &= \{(2p+3)^3(2p+5)(2p+7)\}^{-1} && n = 3.\end{aligned}$$

Hence the first three cumulants are:

$$K_{p1} = \frac{p+1}{2(2p+1)(2p+3)},$$

$$K_{p2} = \frac{2p^2+5p+4}{4(2p+1)^2(2p+3)^2(2p+5)},$$

$$K_{p3} = \frac{4p^3+16p^2+27p+16}{2(2p+1)^3(2p+3)^3(2p+5)(2p+7)}.$$

Good approximations to higher order cumulants may be obtained by using the first few terms of (18).

5. Tests for change of regression at unknown times. Let $(t_{ni}, Y_{ni})_{i=1}^n$ be defined and related as in Section 1 with p fixed and known. It is desired to test

$$H_0: E(Y_{ni}) = \boldsymbol{\beta}'_{p0}\mathbf{t}_{ni} \qquad i = 1, 2, \cdots, n$$

versus

$$H_A: E(Y_{ni}) = \boldsymbol{\beta}'_{p0}\mathbf{t}_{ni} \qquad i = 1, 2, \cdots, k$$
$$E(Y_{ni}) = \boldsymbol{\beta}'_{p1}\mathbf{t}_{ni} \qquad i = k+1, \cdots, n$$

with $\boldsymbol{\beta}_{p0} \neq \boldsymbol{\beta}_{p1}$ and $\boldsymbol{\beta}_{p0}, \boldsymbol{\beta}_{p1}, k$ unknown. For the case $p = 0$ Gardner (1969) has proposed as test statistic the quadratic form

$$Q_n = \sigma^{-2} \sum_{j=1}^{n-1} R_n\left(\frac{n-j}{n}\right)\{\sum_{i=j+1}^n (Y_{ni} - \bar{Y}_n)\}^2 .$$

Analogous to this we propose the following statistic:

$$Q_{pn} = \sigma^{-2} \sum_{j=1}^{n-1} R_n\left(\frac{n-j}{n}\right)\{\sum_{i=j+1}^n (Y_{ni} - \hat{Y}_{ni})\}^2 .$$

From Corollary 3 it follows that the large sample distribution of $n^{-1}Q_{pn}$ is that of $\int_0^1 \psi(t)B_p^2(t)\,dt$. If σ^2 is unknown then, without altering the asymptotic distribution theory, σ^2 may be replaced with a consistent estimator, such as the usual variance estimator based on the sum of squares of residuals.

One-sided tests analogous to those proposed by Chernoff and Zacks (1964) may be based upon Theorem 4.

In the event that the regression parameters are assumed known then, under H_0, the sequence of partial sums

$$\left[\left\{\frac{1}{\sigma n^{\frac{1}{2}}} \sum_{i=1}^j (Y_{ni} - \boldsymbol{\beta}'_{p0}\mathbf{t}_{ni})\right\}_{j=1}^n\right]_{n=1}^\infty$$

converges weakly to the standard Brownian motion process $\{B(t), t \in [0, 1]\}$. Cramér–von Mises type functionals of the form $\int_0^1 \psi(t)B^2(t)\,dt$ are considered in MacNeill (1974) where selected quantiles of the distribution are tabulated for $\psi(t) = at^k$ for a range of values of $k > -2$.

Acknowledgments. The author wishes to thank a referee for helpful comments on an earlier version of this paper.

REFERENCES

[1] ALLAN, F. E. (1930). The general form of the orthogonal polynomials for simple series, with proofs of their simple properties. *Proc. Roy. Soc. Edinburgh* **50** 310–320.

[2] ANDERSON, T. W. and DARLING, D. A. (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *Ann. Math. Statist.* **23** 193–212.

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

[4] BROWN, R. L., DURBIN, J. and Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *J. Roy. Statist. Soc. B* **37** 149–192.

[5] CHERNOFF, H. and ZACKS, S. (1964). Estimating the current mean of a normal distribution which is subjected to changes in time. *Ann. Math. Statist.* **35** 999–1018.

[6] FARLEY, J. U. and HINICH, M. J. (1970). A test for a shifting slope coefficient in a linear model. *J. Amer. Statist. Assoc.* **65** 1320–1329.

[7] FARLEY, J. U., HINICH, M., and MCGUIRE, T. W. (1971). Testing for a shift in the slopes of a multivariate linear time series model. Technical Report WP77–70–1, Graduate School of Industrial Administration, Carnegie-Mellon Univ.

[8] FEDER, PAUL I. (1975). The log likelihood ratio in segmented regression. *Ann. Statist.* **3** 84–97.

[9] GARDNER, L. A. (1969). On detecting changes in the mean of normal variates. *Ann. Math. Statist.* **40** 116–126.

[10] GOLDFIELD, S. M. and QUANDT, R. E. (1973). A Markov model for switching regressions. *J. Econometrics* **1** 3–16.

[11] HINKLEY, D. V. (1969). Inference about the intersection in two-phase regression. *Biometrika* **56** 495–504.

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

[13] QUANDT, R. E. (1958). The estimation of the parameters of a linear regression system obeying two separate regimes. *J. Amer. Statist. Assoc.* **53** 873–880.

[14] QUANDT, R. E. (1960). Tests of the hypothesis that a linear regression system obeys two separate regimes. *J. Amer. Statist. Assoc.* **55** 324–330.

[15] QUANDT, R. E. (1972). A new approach to estimating switching regressions. *J. Amer. Statist. Assoc.* **67** 306–310.

[16] SMIRNOV, N. V. (1936). Sur la distribution de ω^2. *C.R. Acad. Sci. Paris* **202** 449–452.

[17] WATSON, G. N. (1944). *Theory of Bessel Function*, 2nd ed. Cambridge Univ. Press.

DEPARTMENT OF APPLIED MATHEMATICS
UNIVERSITY OF WESTERN ONTARIO
LONDON N6A 5B9
ONTARIO, CANADA