The Annals of Statistics
1974, Vol. 2, No. 5, 950–962

TESTS FOR CHANGE OF PARAMETER AT UNKNOWN TIMES AND DISTRIBUTIONS OF SOME RELATED FUNCTIONALS ON BROWNIAN MOTION[1]

BY IAN B. MACNEILL

University of Western Ontario

Statistics are derived for testing a sequence of observations from an exponential-type distribution for no change in parameter against possible two-sided alternatives involving parameter changes at unknown points. The test statistic can be chosen to have high power against certain of a variety of alternatives. Conditions on functionals on $C[0, 1]$ are given under which one can assert that the large sample distribution of the test statistic under the null-hypothesis or an alternative from a range of interesting hypotheses is that of a functional on Brownian Motion. We compute and tabulate distributions for functionals defined by nonnegative weight functions of the form $\psi(s) = as^k$, $k > -2$. The functionals for $-1 \geqq k > -2$ are not continuous in the uniform topology on $C[0, 1]$.

1. Introduction and summary. This paper presents a derivation, using the method of Chernoff and Zacks (1964), of statistics useful in testing a sequence of observations for no change in parameter against possible two-sided alternatives involving parameter changes at unknown points. The observations are assumed to have distributions of the exponential type. The family of derived statistics is large enough that one can select from it a test statistic that has high power against certain of a variety of alternatives. Large sample distributions for the test statistic under the null hypothesis and under a broad range of alternatives are obtained using the theory of weak convergence. The distributions are related to those of certain functionals on Brownian Motion, but some of the functionals of prime interest are not continuous in the uniform topology on $C[0, 1]$. Conditions on functionals are given under which one can assert that the large sample distribution of the test statistic is that of a functional on Brownian Motion. We compute and tabulate the distribution for functionals defined by nonnegative weight functions of the form $\psi(s) = as^k$, $k > -2$.

2. Two-sided tests for change of parameter at unknown change points for distributions of the exponential type. We begin by posing the testing problem in a form related to that of Page (1957) and similar to that of Kander and Zachs (1966). We let $\{X_j\}_{j=1}^n$ be a sequence of independent random variables taken

Received May 1972; revised November 1973.

[1] This research was supported in part by the National Research Council of Canada.

AMS 1970 *subject classifications.* 60B10, 60E05, 60G50, 60H05, 60J65, 62E15, 62E20, 62F05.

Key words and phrases. Parameter changes at unknown times, functionals on Brownian motions, weak convergence, weight functions, exponential family, Karhünen–Loève expansions, stochastic integrals, Bessel functions.

from a one-parameter exponential density

$$f(X_i \mid \theta_i) = h(X_i)\exp\{\psi_1(\theta_i)U(X_i) + \psi_2(\theta_i)\} \tag{1}$$

with $\psi_1(\cdot)$ and $\psi_2(\cdot)$ having second order derivatives such that

$$\psi_1'(\theta) \neq 0\,, \tag{2}$$
$$0 < \{\psi_1'(\theta)\}^{-3}\{\psi_1''(\theta)\psi_2'(\theta) - \psi_2''(\theta)\psi_1'(\theta)\} < \infty\,.$$

We wish to test the hypothesis of no change in the parameter, i.e.,

$$H_0\colon \theta_i = \theta_0 \qquad\qquad i = 1, 2, \cdots, n$$

against certain alternatives involving changes in the parameter at unknown times. To specify alternatives we follow Gardner (1969) in letting $w_i = 1$ or 0 according to whether there is or is not a change in the parameter between X_i and X_{i+1}. Thus, a particular change sequence is specified by the vector $\mathbf{w}' = (w_1, \cdots, w_{n-1})$. If δ_i represents the change in the parameter after the ith observation (δ_i is allowed to be positive or negative) we find that

$$\begin{aligned} &f(X_k \mid \mathbf{w}, \theta_0, \delta_1, \cdots, \delta_{k-1}) \\ &\quad = h(X_k)\exp\{\psi_1(\theta_0 + \textstyle\sum_{i=1}^{k-1} w_i\delta_i)U(X_k) + \psi_2(\theta_0 + \sum_{i=1}^{k-1} w_i\delta_i)\} \\ &\quad \simeq h(X_k)\exp\{[\psi_1(\theta_0) + \psi_1'(\theta_0)\textstyle\sum_{i=1}^{k-1} w_i\delta_i]U(X_k) + \psi_2(\theta_0) + \psi_2'(\theta_0)\sum_{i=1}^{k-1} w_i\delta_i\}\,, \end{aligned}$$

the approximation being close if $\sum_{i=1}^{k-1} w_i\delta_i$ is small. To derive statistics for testing H_0 we use a Bayesian method introduced by Chernoff and Zacks (1964) that puts a priori distributions on the change points and on the nuisance parameters. Accordingly we assume that δ_i, $i = 1, \cdots, n-1$ are independent $N(0, \gamma^2)$ and obtain the joint density of $X_1, \cdots, X_n, \delta_1, \cdots, \delta_{n-1}$ given $\mathbf{w}$. Thus

$$\begin{aligned} &f(X_1, \cdots, X_n, \delta_1, \cdots, \delta_{n-1} \mid \mathbf{w}, \theta_0) \\ &\quad \simeq \{\textstyle\prod_{k=1}^{n} h(X_k)\}\exp\{n\psi_2(\theta_0) + \psi_1(\theta_0)\sum_{k=1}^{n} U(X_k)\}\{\prod_{k=2}^{n}\exp[\psi_2'(\theta_0)\sum_{i=1}^{k-1} w_i\delta_i]\} \\ &\qquad \times \{\textstyle\prod_{k=2}^{n}\exp[(\sum_{i=1}^{k-1} w_i\delta_i)\psi_1'(\theta_0)U(X_k)]\}\left\{\prod_{k=1}^{n-1}\frac{1}{(2\pi\gamma^2)^{\frac{1}{2}}}\exp\left[-\frac{1}{2\gamma^2}\delta_k^2\right]\right\}\,. \end{aligned}$$

Integrating with respect to δ_k, $k = 1, 2, \cdots, n-1$ to obtain the marginal density we see that

$$\begin{aligned} &f(X_1, X_2, \cdots, X_m \mid \mathbf{w}, \theta_0) \\ &\qquad \simeq \{\textstyle\prod_{k=1}^{n} h(X_k)\exp[\psi_1(\theta_0)U(X_k) + \psi_2(\theta_0)]\} \\ &\qquad\qquad \times \exp[\gamma^2/2\textstyle\sum_{k=1}^{n-1} w_k(\sum_{j=k+1}^{n}\{\psi_2'(\theta_0) + \psi_1'(\theta_0)U(X_j)\})^2]\,. \end{aligned}$$

Upon letting γ^2 become arbitrarily small the approximation becomes close, and dropping γ^2 we see that a likelihood ratio statistic for testing H_0 against the change sequence $\mathbf{w}$ is

$$\Lambda_{\mathbf{w}}(X_1, X_2, \cdots, X_n) = \sum_{k=1}^{n-1} w_i\left[\sum_{j=k+1}^{n}\left(\frac{\psi_2'(\theta_0) + \psi_1'(\theta_0)\mu_0}{\psi_1'(\theta_0)\tau_0} + \frac{U(X_j) - \mu_0}{\tau_0}\right)\right]^2,$$

where $\mu_0 = E_{\theta_0}[U(X_i)]$ and $\tau_0^2 = \mathrm{Var}_{\theta_0}[U(X_i)]$. With the assumptions (2) made on the functions $\psi_1(\cdot)$ and $\psi_2(\cdot)$ one is able to pass derivatives through the integral in $\int f(x \mid \theta)\,dx$ (see Lehmann (1959), page 52) which enables one to derive the following:

$$\psi_2'(\theta_0) + \psi_1'(\theta_0)E_{\theta_0}[U(X_j)] = 0\,, \tag{3}$$
$$\mathrm{Var}_{\theta_0}[U(X_j)] = (\psi_1'(\theta_0))^{-3}[\psi_1''(\theta_0)\psi_2'(\theta_0) - \psi_2''(\theta_0)\psi_1'(\theta_0)]\,.$$

Consequently, we may write

$$\Lambda_{\mathbf{w}}(X_1, \cdots, X_n) = \sum_{k=1}^{n-1} w_k \left[\sum_{j=k+1}^{n} \frac{\{\psi_1'(\theta_0)U(X_j) + \psi_2'(\theta_0)\}\{\psi_1'(\theta_0)\}^{\frac{1}{2}}}{\{\psi_1''(\theta_0)\psi_2'(\theta_0) - \psi_2'(\theta_0)\psi_1'(\theta_0)\}} \right]^2 .$$

If we assign to the collection of possible change sequences, $\{\mathbf{w}\}$, a prior distribution, $P(\mathbf{w})$, we obtain the test statistic

$$\begin{aligned} T_n(P) &= \textstyle\sum_{\{\mathbf{w}\}} P(\mathbf{w})\Lambda_{\mathbf{w}}(X_1, \cdots, X_n) \\ &= \sum_{k=1}^{n-1} P_k \left[\sum_{j=k+1}^{n} \frac{\{\psi_1'(\theta_0)U(X_j) + \psi_2'(\theta_0)\}\{\psi_1'(\theta_0)\}^{\frac{1}{2}}}{\{\psi_1''(\theta_0)\psi_2'(\theta_0) - \psi_2''(\theta_0)\psi_1'(\theta_0)\}} \right]^2 , \end{aligned} \tag{4}$$

where $\{P_k\}_{k=1}^{n-1}$ is a nonnegative weight sequence whose components are given by

$$P_k = \textstyle\sum_{\{\mathbf{w}\}} P(\mathbf{w})w_k\,, \qquad k = 1, 2, \cdots, n-1\,. \tag{5}$$

3. Large sample distribution theory for $T_n(P)$. We first establish some notation and recall some well-known results for sums of independent and identically distributed random variables. Let $\{X_j\}_{j=1}^{\infty}$ be such a sequence, each component having zero mean and positive variance $\sigma^2 < \infty$ and each defined on the same probability space $(\Omega, \mathscr{A}, P)$. Let $\{S_k\}_{k=0}^{\infty}$ with $S_k = \sum_{j=1}^{k} X_j$, $S_0 \equiv 0$ be the corresponding sequence of partial sums. If we pair S_k with the point k/n then $[\{(\sigma n^{\frac{1}{2}})^{-1}S_k\}_{k=0}^{n}]_{n=1}^{\infty}$ can be used to define a sequence of stochastic processes $\{Y_n(t, \omega), t \in [0, 1], \omega \in \Omega\}_{n=1}^{\infty}$ possessing continuous sample paths by the relation

$$Y_n(t, \omega) \equiv (\sigma n^{\frac{1}{2}})^{-1}S_{[nt]}(\omega) + (nt - [nt])(\sigma n^{\frac{1}{2}})^{-1}X_{[nt]+1}(\omega)\,,$$

where $[x]$ indicates the largest integer in x.

Let C denote the space of continuous functions on the unit interval with the uniform topology and let $\mathscr{F}$ denote the Borel subsets of C. Then Y_n induces a measure, P_n, on $(C, \mathscr{F})$ and Theorem 10.1 Billingsley (1968) assures us that P_n converges weakly to Wiener measure W, which is the probability measure on $(C, \mathscr{F})$ corresponding to that of the standard Brownian Motion process having continuous sample paths. This process, which is a measurable map from some probability space to $(C, \mathscr{F})$ and which we denote by $\{B(t), t \in [0, 1]\}$, is Gaussian with zero mean and covariance kernel $E[B(s)B(t)] = \min(s, t)$. The sequence $\{Y_n\}_{n=1}^{\infty}$ is said to converge weakly to Brownian Motion. The paths of these processes can be conditioned to be zero at the end points by considering the sequence $[\{(\sigma n^{\frac{1}{2}})^{-1}(S_k - kn^{-1}S_n)\}_{k=0}^{n}]_{n=1}^{\infty}$. The limit process is now the Brownian Bridge process, a Gaussian process denoted by $\{B_0(t), t \in [0, 1]\}$ with paths zero at $t = 0$ and $t = 1$, zero mean and covariance kernel $\{\min(s, t) - st\}$. Similar weak convergence results hold for the sequence of "reversed" partial sums. That is, suppose $X_{jn}^* = X_{n-j+1}$, $j = 1, 2, \cdots, n$ and $S_{kn}^* = \sum_{j=1}^k X_{jn}^*$, $S_{0n}^* = 0$. Then $[\{(\sigma n^{\frac{1}{2}})^{-1} S_{kn}^*\}_{k=0}^n]_{n=1}^\infty$ and $[\{(\sigma n^{\frac{1}{2}})^{-1}(S_{kn}^* - kn^{-1}S_{nn}^*)\}_{k=0}^n]_{n=1}^\infty$ define sequences of stochastic processes that converge weakly to Brownian Motion and Brownian Bridge processes respectively. We have use for these results in studying the large sample distribution of $T_n(P)$, the statistic derived in Section 2 for detecting changes in parameters at unknown change points. In this connection Page (1957) noted that it is convenient to view the sequence of partial sums from the vantage point of the last element.

We first prove a result concerning $[\{S_k\}_{k=1}^{n-1}]_{n=1}^\infty$ and the Brownian Motion process.

THEOREM 1. *Let $\psi(\cdot)$ be a nonnegative measurable function defined on the unit interval such that $\int_0^1 t\psi(t)\,dt < \infty$. If*

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

$$Z_n = \sum_{j=1}^{n-1} P_n(j/n)((\sigma n^{\frac{1}{2}})^{-1} S_j)^2 \qquad \text{and} \qquad Z = \int_0^1 \psi(t) B^2(t)\,dt,$$

then

$$\lim_{n\to\infty} P[Z_n \leqq \alpha] = P[Z \leqq \alpha]. \tag{6}$$

PROOF. Since the functional $F_\psi(\cdot)$ defined by $F_\psi(B) = \int_0^1 \psi(t)B^2(t)\,dt$ is not necessarily continuous in the uniform topology on C we cannot verify (6) by appealing directly to the weak convergence of the sequence of partial sums to Brownian Motion and to the Donsker result. However, we are able to apply a result due to Wichura (1971), page 1769, to establish (6).

We first let $Z_n{}^* = \int_0^1 \psi(t) Y_n{}^2(t)\,dt$, and define a sequence of measurable functions $\{A_n\}_{n=1}^\infty$ on the functionals by

$$A_n Z = \sum_{j=1}^{n-1} P_n(j/n) B^2(j/n)$$

and

$$A_n Z_m{}^* = \sum_{j=1}^{n-1} P_n(j/n) Y_m{}^2(j/n).$$

For $m > n$ and $j/n \leqq t \leqq (j+1)/n$

$$E|Y_m{}^2(t) - Y_m{}^2(j/n)| \leqq [E\{Y_m(t) - Y_m(j/n)\}^2 E\{Y_m(t) + Y_m(j/n)\}^2]^{\frac{1}{2}}$$
$$\leqq \left\{\left(\frac{1}{n} + \frac{2}{m}\right)\left(\frac{j+1}{n} + \frac{3}{m}\right)\right\}^{\frac{1}{2}}.$$

Consequently

$$E|Z_m{}^* - A_n Z_m{}^*|$$
$$\leqq E|\int_0^{1/n} \psi(t) Y_m{}^2(t)\,dt| + E|\sum_{j=1}^{n-1} \int_{j/n}^{(j+1)/n} \psi(t)\{Y_m{}^2(t) - Y_m{}^2(j/n)\}\,dt|$$
$$\leqq \int_0^{1/n} t\psi(t)\,dt + \left(\frac{1}{n} + \frac{2}{m}\right)^{\frac{1}{2}} \sum_{j=1}^{n-1} \int_{j/n}^{(j+1)/n} \left(\frac{j+1}{n} + \frac{3}{m}\right)^{\frac{1}{2}} \psi(t)\,dt.$$

Since, for $\varepsilon > 0$,

$$P[|Z_m - A_n Z_m{}^*| \geqq \varepsilon] \leqq \varepsilon^{-1} E|Z_m - A_n Z_m{}^*|$$

and since $\int_0^1 t\psi(t)\,dt < \infty$ implies that $\lim_{n\to\infty} n^{-\frac{1}{2}} \int_{1/n}^1 t^{\frac{1}{2}}\psi(t)\,dt = 0$ we have that

$$\lim_{n\to\infty} \limsup_{m\to\infty} P[|Z_m - A_n Z_m{}^*| \geqq \varepsilon] = 0. \tag{7}$$

By a similar argument we can show that

$$\lim_{n\to\infty} P[|Z - A_n Z| \geqq \varepsilon] = 0 \tag{8}$$

since

$$E|Z - A_n Z| \leqq \int_0^{1/n} t\psi(t)\,dt + 3/n^{\frac{1}{2}} \int_{1/n}^1 t^{\frac{1}{2}}\psi(t)\,dt\,.$$

Finally, due to the weak convergence of Y_n to Brownian Motion and the consequent convergence of the finite dimensional distributions, it follows that as $m \to \infty$, $A_n Z_m{}^*$ converges in distribution to $A_n Z$. Combining this with (7) and (8) enables us to invoke Wichura's result to prove that $Z_n{}^*$ converges in distribution to Z. We verify (6) by noting that

$$E|Z_n - Z_n{}^*| \leqq \int_0^{1/n} t\psi(t)\,dt + \sum_{j=1}^{n-1} n^{-\frac{1}{2}} \int_{j/n}^{(j+1)/n} \psi(t)\{(4j+1)/n\}^{\frac{1}{2}}\,dt\,,$$

which implies that $Z_n - Z_n{}^*$ converges in distribution to zero.

The terms in the inner sum of (4) are equal to

$$\left\{\sum_{j=k+1}^{n} \frac{U(X_j) - \mu_0}{\tau_0}\right\}_{k=1}^{n-1}$$

and form what we referred to at the beginning of this section as a sequence of "reversed" partial sums. With some simple modifications Theorem 1 can be applied to these "reversed" partial sums. In fact, if $\psi(\cdot)$ satisfies the conditions of Theorem 1 and if $\{P_k\}_{k=1}^{n-1}$ is defined by

$$P_{n-k} = \int_{k/n}^{(k+1)/n} \psi(t)\,dt\,, \tag{9}$$

then $n^{-1}T_n(P)$ converges in distribution to $\int_0^1 \psi(t)B^2(t)\,dt$.

What remains to complete the asymptotic distribution theory for $T_n(P)$ is to compute the probability distributions for $\int_0^1 \psi(t)B^2(t)\,dt$ for a wide range of weight functions $\psi(\cdot)$. In Section 4 we partially solve this problem by obtaining the distributions for the family of weight functions defined by $\psi(t) = at^k$, $k > -2$. Note that for $-2 < k \leqq -1$ the corresponding functionals are not continuous in the uniform topology. However, this is no hardship since $\psi(\cdot)$ satisfies the condition of Theorem 1. The weight sequence $\{P_k\}_{k=1}^{n-1}$ may emphasize changes in certain regions and the functions $\psi(t) = at^k$ can be used to model certain of these sequences. A uniform weight sequence, which corresponds to the case of at most one change (AMOC) at time m having a uniform prior, calls for the weight function with $k = 0$. A weight sequence that puts large weight on early changes could be approximated with k large, whereas a sequence that puts large weight on late changes could be approximated with k small (subject to $k > -2$). Using $k = -1$ one weights the change times proportional to the reciprocal of the variance of the partial sums.

If the initial level θ_0 is unknown then μ_0 will be unknown. Under the conditions assumed μ_0 may be consistently estimated by $\bar{U}_n = n^{-1} \sum_{j=1}^n U(X_j)$. Replacing μ_0 with $\bar{U}_n$ results in a substantial change in the distribution in that the limit process is the Brownian Bridge rather than Brownian Motion. Thus to obtain large sample distributions for the modified test statistic $T_n^0(P)$ defined by

$$T_n^0(P) = \sum_{k=1}^{n-1} P_k \left[\sum_{j=k+1}^{n} \frac{U(X_j) - \bar{U}_n}{\tau_0} \right]^2$$

one must first examine the conditions under which $n^{-1}T_n^0(P)$ converges in distribution to $\int_0^1 \psi(t)B^2(t)\,dt$.

The folllowing result, analogous to Theorem 1, pertains to $[\{(\sigma n^{\frac{1}{2}})^{-1}(S_k - (k/n)S_n)\}_{k=1}^{n-1}]_{n=1}^{\infty}$ and to the Brownian Bridge process having continuous sample paths.

THEOREM 2. *Let $\psi(t)$ be a nonnegative weight function defined on the unit interval such that $\int_0^1 t(1-t)\psi(t)\,dt < \infty$. If $P_n(j/n) = \int_{(2j-1)/2n}^{(2j+1)/2n} \psi(t)\,dt$, $j = 1, 2, \cdots, n-1$,*

$$Z_n^0 = \sum_{j=1}^{n-1} P_j[(\sigma n^{\frac{1}{2}})^{-1}(S_j - (j/n)S_n)]^2 \qquad \text{and} \qquad Z^0 = \int_0^1 \psi(t)B_0^2(t)\,dt\,,$$

then

$$\lim_{n\to\infty} P[Z_n^0 \leqq \alpha] = P[Z^0 \leqq \alpha]\,.$$

The proof is similar to that of Theorem 1. To apply this theorem to the "reversed" partial sums of $n^{-1}T_n^0(P)$ one need only define $\{P_k\}_{k=1}^{n-1}$ by

$$P_{n-k} = \int_{(2k-1)/2n}^{(2k+1)/2n} \psi(t)\,dt\,. \tag{10}$$

To complete the asymptotic distribution theory for $n^{-1}T_n^0(P)$ one requires the probability distributions for $\int_0^1 \psi(t)B_0^2(t)\,dt$ for a variety of weight functions. The collection of such distributions presently available is not very large. However, for $\psi(t) \equiv 1$ the distribution is tabulated by Anderson and Darling (1952), and for $\psi(t) = (t(1-t))^{-1}$ several important percentage points are given by Anderson and Darling (1954). The latter function, which puts large weight on changes near the end points, defines a functional that is not continuous in the uniform topology on C, as is claimed by Anderson and Darling (1952) page 197. However, the proof that $n^{-1}T_n^0(P)$ converges in distribution to $\int_0^1 \{t(1-t)\}^{-1}B_0^2(t)\,dt$ follows from Theorem 2.

We now examine the large sample distribution of the statistic $T_n(P)$ as defined by (4) under alternatives to H_0. For fixed alternatives the statistic always rejects H_0 if the sample size is large enough. However, if the alternative approaches the null hypothesis at a rate that is neither too fast nor too slow, more interesting large sample results are available. Since the sequence of independent random variables $\{X_k\}_{k=1}^{n}$ with densities $f(X_k \mid \theta_k)$ given by (1) have means and variances given by (3) that are both functions of θ we will denote them by $\mu(\theta) = E_\theta[U(X)]$ and $\tau^2(\theta) = \mathrm{Var}_\theta\,[U(X)]$. Then the following theorem may be proved.

THEOREM 3.

(i) *Assume that $\mu(\cdot)$ and $\tau^2(\cdot)$ satisfy assumptions* (2) *and that the first two derivatives of $\mu(\cdot)$ and $\tau^2(\cdot)$ exist and are continuous in a neighborhood of θ_0.*

(ii) *Let $h(\cdot)$ be a function bounded and Riemann-integrable on* $[0, 1]$ *and let $H(t) = \int_0^t \{h(1-s) - \mu_0\}\,ds$.*

(iii) *Let* $\{g_n(t), t \in [0, 1]\}_{n=1}^{\infty}$ *be a sequence of functions such that*

$$g_n(t) - \theta_0 = \frac{(H(t) - \mu_0)\tau_0}{n^{\frac{1}{2}}\mu'(\theta_0)} + O\left(\frac{1}{n^{\frac{1}{2}+\delta}}\right). \qquad \delta > 0,$$

uniformly in t.

(iv) *Let* $\psi(\cdot)$ *be as in Theorem* 1 *and let* $\{P_k\}_{k=1}^{n-1}$ *be defined by* (9). *Then, under the hypothesis that* $\theta_k = g(k/n)$, $k = 1, 2, \cdots, n$, $n^{-1}T_n(P)$ *converges in distribution to* $\int_0^1 \psi(t)(H(t) + B(t))^2\, dt$ *where* $\{B(t), t \in [0, 1]\}$ *is standard Brownian Motion.*

PROOF. The test statistic may be expressed as

$$n^{-1}T_n(P) = \sum_{k=1}^{n-1} \frac{P_k}{\tau_0^2}\Big[n^{\frac{1}{2}} \sum_{j=k+1}^{n} \frac{1}{n}\{\mu(g_n(j/n)) - \mu(\theta_0)\}$$
$$+ \frac{1}{n^{\frac{1}{2}}} \sum_{j=k+1}^{n} \{U(X_j) - \mu(g_n(j/n))\}\Big]^2.$$

By expanding $\mu(\cdot)$ about θ_0 and using the special nature of $\{g_n(t)\}_{n=1}^{\infty}$ it can be shown that, for $k = [nt]$, $[n^{\frac{1}{2}} \sum_{j=k+1}^{n} n^{-1}\{\mu(g_n(j/n)) - \mu_0\}]_{k=0}^{n-1}$ converges uniformly to $H(t)$ as $n \to \infty$. Since $\{[U(X_j) - (g_n(j/n))]/\tau_0\}_{j=1}^{n}$ are zero mean variables that are independent but not identically distributed the weak convergence result appealed to earlier does not apply. However, we have that

$$\tau^2[g_n(j/n)] - \tau_0^2 = \frac{(h(j/n) - \mu_0)\tau_0^{2\prime}(\theta_0)}{n^{\frac{1}{2}}\mu'(\theta_0)} + O\left(\frac{1}{n}\right), \tag{11}$$

which implies that $\{\tau^2(g_n(j/n))\}_{j=1}^{n}$ converges to the constant function $\tau^2(\theta) \equiv \tau_0^2$ at a rate equal to $(1/n^{\frac{1}{2}})$. This enables us to apply the lemma to Theorem 10.4 of Billingsley (1968), page 69, to demonstrate the tightness of the probability measures generated in C by the sequences of processes defined by the "reversed" partial sums $[n^{-\frac{1}{2}} \sum_{j=k+1}^{n} \{(U(X_j) - \mu(g_n(j/n))\}]_{k=0}^{n-1}$. Since (11) implies that $\{U(X_j) - \mu(g_n(j/n))\}_{j=1}^{n}$ have variances that are almost equal for large n we can verify the conditions of the Lindeberg theorem to show that $n^{-\frac{1}{2}} \sum_{j=[tn]}^{n} \{[U(X_j) - \mu(g_n(j/n))]/\tau_0\}$, $t \in (0, 1)$ converges in distribution to a normal variable with zero mean and variance $(1 - t)$. The finite dimensional distributions of these processes can be shown to be asymptotically those of standard Brownian Motion by applying a technique due to Cramér and Wold that reduces the multidimensional problem to the one-dimensional problem just discussed. Consequently $[n^{-\frac{1}{2}} \sum_{j=k+1}^{n} \{U(X_j) - \mu(g_n(j/n))\}]_{k=0}^{n-1}$ forms a sequence of "reversed" partial sums that converges weakly to the Brownian Motion process with parameter $\sigma^2 = \tau_0^2$. Since $h(\cdot)$ is bounded, $H(t) = 0(t)$, which along with (iv) implies that $\int_0^1 \psi(t)H(t)\, dt < \infty$. The method of proof for Theorem 1 can now be used to show that

$$\lim_{n\to\infty} P[n^{-1}T_n(P) \leqq \alpha] = P[\textstyle\int_0^1 \psi(t)(H(t) + B(t))^2\, dt \leqq \alpha], \qquad \alpha > 0,$$

where $B(\cdot)$ is standard Brownian motion.

Section 4 contains a discussion of the distribution theory of $\int_0^1 \psi(t)(H(t) + B(t))^2\,dt$ for $H(\cdot)$ satisfying the conditions of Theorem 3 and for $\psi(t) = at^k$, $k > -2$.

We now consider an interesting special case. Assuming that observations are normally distributed with variance σ^2 and initial mean μ_0 the test statistic for testing no change in mean reduces to

$$T_n = \sigma^{-2} \sum_{k=1}^{n-1} P_k\{\sum_{j=k+1}^{m} (X_j - \mu_0)\}^2$$

which is what one would obtain using the method of Gardner (1969). Gardner goes further and considers the above problem with the initial level not specified and derives the statistic

$$Q_n = \sigma^{-2} \sum_{\{\mathbf{w}\}} P(\mathbf{w}) \sum_{k=1}^{n-1} w_k[\sum_{j=k+1}^{n} (X_j - \bar{X})]^2 = \sigma^{-2} \sum_{k=1}^{n-1} P_k[\sum_{j=k+1}^{n} (X_j - \bar{X})]^2,$$

with $\{P_k\}$ given by (5) and (10). As noted above, the limit process is the Brownian Bridge. Consequently, if $\psi(\cdot)$ satisfies the condition of Theorem 2 the asymptotic distribution of $n^{-1}Q_n$ is that of $\int_0^1 \psi(t)B_0^2(t)\,dt$. Viewed in this framework it is natural to expect Gardner's result which states that for at most one change (AMOC) at time m having a uniform prior the asymptotic distribution of the quadratic form in normal variables is the same as that of the Cramér–von Mises goodness-of-fit test statistic. It can also be noted that the same asymptotic results hold for variables from any distribution (exponential family or otherwise) having finite variance.

Sen (1971) has considered the statistic Q_n with normal observations in the AMOC situation (uniform prior) in a multivariate framework and has also derived small sample distributions for the univariate case, which indicates that the sample size does not need to be very large for the asymptotic theory to provide a good approximation, a fact that is also noted by Gardner.

The difference between the statistic $T_n(P)$ given by (4) and the statistic derived by Kander and Zacks (1966) stems mainly from the fact that they choose to restrict parameter changes to be of one sign. Under these circumstances, in the AMOC case, they obtain, for testing H_0 against a possible change at time k having a prior distribution $\{P_k\}_{k=1}^{n-1}$, the statistic

$$T_n'(P) = \sum_{k=1}^{n-1} P_k \sum_{j=k+1}^{n} \frac{U(X_j) - \mu_0}{\tau_0}.$$

In the following theorem the statistic is viewed as a weighted average of "reversed" partial sums. We may then use the theory of weak convergence of probability measures to demonstrate conditions under which interesting large sample distributions for this statistic exist.

THEOREM 4. *Let $\psi(t)$ be a nonnegative measurable function defined on the unit interval such that $\int_0^1 t^{\frac{1}{2}}\psi(t)\,dt < \infty$ and assume that* (i), (ii) *and* (iii) *of Theorem* 3 *hold. Then with $\{P_k\}_{k=1}^{n-1}$ given by* (9), *$W(s) = \int_s^1 \psi(t)\,dt$ and under the hypothesis that $\theta_k = g(k/n)$, $n^{-\frac{1}{2}}T_n'(P)$ converges in distribution to the Gaussian variable $\int_0^1 \psi(t)(H(t) + B(t))\,dt$ having mean $\int_0^1 \psi(t)H(t)\,dt$ and variance $\int_0^1 W^2(s)\,ds < \infty$.*

We omit the details of the proof since they are similar to those of Theorems 1 and 3.

As with $T_n(P)$, if $\bar{U}_n$ replaces μ_0, Brownian Motion must be replaced by the Brownian Bridge to obtain the appropriate asymptotic distributions.

Page (1955) originally formulated the problem and proposed a range-type statistic for testing a one-sided alternative. Later (1957) he proposed the use of a one-sided Smirnov statistic.

Among a host of other test statistics that might be considered in this context are the range as discussed by Feller (1951), and a statistic discussed by Watson (1961) and Stephens (1963), (1964). A tabulation of the range, and a discussion of the Stephens-Watson statistic when the initial level is known, are given by MacNeill (1971). For a discussion of the distribution theory when the initial level is assumed unknown one can refer to Feller and Watson. Nadler and Robbins (1971) consider a two-sided version of Page's original statistic and show that it is equivalent to the range.

4. Distributions of certain functionals on Brownian Motion. We now concern ourselves with the distributions of the stochastic integrals of Section 3. Our approach is to apply to Brownian Motion the methods that Anderson and Darling (1952) applied to Brownian Bridge processes. We consider the process $\{(\psi(t))^{\frac{1}{2}}B(t),\, t \in [0, 1]\}$ and obtain its Karhünen-Loève expansion. This expansion requires that we compute a set of functions $\{\phi_n(\cdot)\}_{n=1}^{\infty}$ defined on the unit interval and a set of zero mean random variables $\{b_n\}_{n=1}^{\infty}$ such that

$$E((\psi(t))^{\frac{1}{2}}B(t) - \textstyle\sum_{n=1}^{\infty} b_n \phi_n(t))^2 = 0\,, \qquad t \in [0, 1]\,,$$

with

$$\int_0^1 \phi_n(t)\phi_m(t)\,dt = 1 \quad m = n\,, \qquad = 0 \quad m \neq n\,,$$

and

$$E(b_n b_m) = \lambda_n \quad m = n\,, \qquad = 0 \quad m \neq n\,.$$

Because Brownian Motion is Gaussian, $b_n \sim N(0, \lambda_n)$. The characteristic functions for the stochastic integrals of Theorems 1 and 3 are given in the following result.

THEOREM 5. *Assume* (i) $\psi(\cdot)$ *and* $H(\cdot)$ *satisfy the assumptions of Theorem* 3, *and* (ii) *the kernel* $k(s, t) = \min(s, t)(\psi(t))^{\frac{1}{2}}(\psi(s))^{\frac{1}{2}}$ *is continuous on the closed unit square or continuous except at* $(0, 0)$ *with* $\partial k(s, t)/\partial s$ *continuous for* $0 < s,\, t \leqq 1$, $s \neq t$ *and bounded for* $\varepsilon \leqq s \leqq 1$ *for* $\varepsilon > 0$. *Then, if* $\{\lambda_k, \phi_k(\cdot)\}_{k=1}^{\infty}$ *are the eigenvalues and eigenfunctions of the Karhünen-Loève expansion for the process* $\{(\psi(t))^{\frac{1}{2}}B(t),\, t \in [0, 1]\}$, *and* $\nu_k = \int_0^1 (\psi(t))^{\frac{1}{2}}H(t)\phi_k(t)\,dt$, $k = 1, 2, \cdots$, *the characteristic function of* $\int_0^1 \psi(t)(H(t) + B(t))^2\,dt$ *is*

$$\Phi_{\psi,H}(t) = \{\textstyle\prod_{k=1}^{\infty} (1 - 2it\lambda_k)^{-\frac{1}{2}}\} \exp[it \textstyle\sum_{k=1}^{\infty} \nu_k^2(1 - 2it\lambda_k)^{-1}]\,.$$

PROOF. If the representation and the orthogonality conditions of the expansion are to hold simultaneously then $\phi_n(t)$, λ_n, $n = 1, 2, \cdots$ must satisfy the Fredholm equation

$$\int_0^1 \min(s, t)(\psi(t))^{\frac{1}{2}}(\psi(s))^{\frac{1}{2}}\phi(s)\,ds = \lambda\phi(t)\,. \tag{12}$$

Condition (ii) has been chosen analogous to the conditions of Anderson and Darling (1952) so that we can obtain, by solving (12), the spectral representation of the self-adjoint operator induced on $L^2[0, 1]$ by $k(s, t)$ and hence obtain the required sequences $\{\lambda_n, \phi_n(\cdot)\}_{n=1}^{\infty}$. By Parseval's Theorem $\sum_{k=1}^{\infty} \nu_k^2 = \int_0^1 \psi(t)H^2(t)\,dt$ and (i) assures us that the latter is finite. Hence, by replacing $B(t)$ and $H(t)$ with their expansions we see that $\int_0^1 \psi(t)(H(t) + B(t))^2\,dt$ is distributed as $\sum_{k=1}^{\infty}(b_k + \nu_k)^2$. Since $(b_k + \nu_k)^2/\lambda_k$ is a noncentral χ_1^2 variable, $\sum_{k=1}^{\infty}(b_k + \nu_k)^2$ is a weighted sum of such variables having the characteristic function $\Phi_{\psi,H}(t)$.

The nth cumulant, K_n, of the stochastic integral is

$$K_n = 2^{n-1}(n-1)!\sum_{k=1}^{\infty}(\lambda_k^n + n\lambda_k^{n-1}\nu_k)\,. \tag{13}$$

We now complete the analysis for the family of weight functions $\psi(s) = as^k$, $k > -2$, and for $H(t) = 0$, which is the null hypothesis. Letting $\phi(s) = (\psi(s))^{\frac{1}{2}}h(s)$ we see that (12) is equivalent to

$$h''(s) + \frac{1}{\lambda}\psi(x)h(s) = 0\,,$$

subject to $h(0) = 0$ and $h'(1) = 0$. It can be shown that

$$h_n(s) = As^{\frac{1}{2}}J((k+2)^{-1}; 2(a/\lambda_n)^{\frac{1}{2}}s^{(k+2)/2}(k+2)^{-1})\,,$$

where $J(\nu; x)$ is a Bessel function of the first kind, A is a normalizing constant and

$$\lambda_n = 4a/[(k+2)^2\{j(-(k+1)/(k+2); n)\}^2]$$

with $j(\nu; n)$ being the nth zero of $J(\nu; x)$. Thus, we have the following:

COROLLARY 6. *The characteristic function for* $T_k = \int_0^1 at^kB^2(t)\,dt$, $k > -2$, *is*

$$\Phi_{T_k}(s) = (\Gamma(k+2)^{-1}((2ais)^{\frac{1}{2}}/(k+2))^{(k+1)/(k+2)} \times J(-(k+1)/(k+2); 2(k+2)^{-1}(2ais)^{\frac{1}{2}}))^{-\frac{1}{2}}\,.$$

Letting $D(2is) = (\Phi_{T_k}(s))^2$ and $\lambda = 2is$ one can derive the following expression for the probability distribution function:

$$\Omega_\psi(\alpha) = 1 - \frac{1}{\pi}\sum_{n=1}^{\infty}(-1)^{n-1}\int_{1/\lambda_{2n-1}}^{1/\lambda_{2n}} \frac{e^{-(\lambda/2)\alpha}}{\lambda(-D(\lambda))^{\frac{1}{2}}}\,d\lambda \tag{14}$$

where $\psi(s) = as^k$, $k > -2$.

A distribution such as (14) was first (essentially) derived by Smirnov (1936) and, as has been noted by Darling (1957), it is a far from ideal computing formula due to the singularities at the end points of the intervals over which the integrations are performed. However, all the zeros of the Bessel functions are real, simple and well-spaced. Consequently, by some simple substitutions, each integral can be put in the form

$$I = \int_{-1}^{1} \frac{g(x)}{(1 - x^2)^{\frac{1}{2}}}\, dx\,, \tag{15}$$

where $g(x)$ is well behaved over the interval $[-1, 1]$ and then (15) can be effectively evaluated by Gaussian quadrature of the Chebychev form. This is one of the techniques used by Grad and Solomon (1955) for the evaluation of the probability distributions for positive definite quadratic forms. We have evaluated (6) with $\psi(s) = as^k$ for $k = -\frac{3}{4}, -\frac{1}{2}, -\frac{1}{4}, 0, 1, \frac{3}{2}, 2, 3, 4$ with $a = k + 1$ and for $k = -1, -\frac{5}{4}, -\frac{3}{2}, -\frac{7}{4}$ with $a = k + 2$. For $k > -1$, this choice of normalization makes $\psi(\cdot)$ a probability density (and hence useable as a proper prior measure) and for $-2 < k \leqq -1$, since $\psi(\cdot)$ is not integrable, a is chosen arbitrarily to yield $E[T_k] = 1$. The results of the above computations appear in Table 1.

TABLE 1
Selected percentage points for $\Omega_{\psi_k}(\alpha) = P[\int_0^1 as^k B^2(s)\, ds \leqq \alpha]$

$\Omega_{\psi_k}(x)$	$k =$												
	-1.75	-1.50	-1.25	-1.00	-0.75	-0.50	-0.25	0.0	1.00	1.50	2.00	3.00	4.00
0.01	0.2634	0.1781	0.1378	0.1135	0.0194	0.0284	0.0326	0.0345	0.0337	0.0321	0.0303	0.0271	0.0244
0.025	0.3079	0.2151	0.1696	0.1417	0.0245	0.0361	0.0418	0.0444	0.0443	0.0425	0.0404	0.0366	0.0332
0.05	0.3546	0.2557	0.2055	0.1740	0.0304	0.0453	0.0528	0.0565	0.0576	0.0556	0.0533	0.0488	0.0448
0.10	0.4209	0.3160	0.2605	0.2248	0.0399	0.0601	0.0708	0.0765	0.0804	0.0786	0.0762	0.0711	0.0664
0.50	0.8295	0.7495	0.6982	0.6617	0.1268	0.2042	0.2551	0.2905	0.3621	0.3794	0.3915	0.4069	0.4164
0.90	1.7883	1.9951	2.1200	2.2066	0.4542	0.7734	1.0113	1.1958	1.6509	1.7883	1.8938	2.0453	2.1490
0.95	2.2288	2.6060	2.8337	2.9889	0.6204	1.0628	1.3956	1.6557	2.3041	2.5017	2.6539	2.8735	3.0244
0.975	2.6864	3.2453	3.5795	3.8052	0.7937	1.3643	1.7959	2.1347	2.9842	3.2443	3.4453	3.7357	3.9357
0.99	3.3121	4.1194	4.5975	4.9183	1.0300	1.7752	2.3415	2.7875	3.9108	4.2563	4.5236	4.9105	5.1773
Mean	1	1	1	1	$\frac{1}{5}$	$\frac{1}{3}$	$\frac{3}{7}$	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{5}{7}$	$\frac{3}{4}$	$\frac{4}{5}$	$\frac{5}{6}$

If $k = 0$, we obtain a special case that has been partially tabulated by MacNeill (1971) and verified by Rothman and Woodroofe (1972). The computational technique used in those instances is more convenient, especially for small values of the argument, and was used as a check on the accuracy of the results. Agreement to the number of places used in the table was obtained.

With the normalizing constant as chosen above it can be seen that T_k converges in distribution to a χ_1^2 variable as $k \to \infty$, to the constant 1 as $k \downarrow -2$ and to the constant 0 as $k \downarrow -1$. As long as one keeps in mind the discontinuity at $k = -1$ caused by the different normalizations one can use Table 1 and the above limiting cases to obtain, by interpolation, approximations to the percentage points for any $k > -2$. For example, linear interpolation yields 0.3768 and cubic interpolation yields 0.3803 as the median for $k = 1.5$ whereas the exact result as given in the table is 0.3794.

The kernel with weight function $\psi(s) = s^{-1}$ has a particular appeal since it is the reciprocal of the variance of Brownian Motion just as $\psi(s) = [s(1 - s)]^{-1}$ is asymptotically the reciprocal of the variance of the process for the Anderson–Darling statistic, some approximate small sample results for which are given by Lewis (1961).

If the characteristic function of Theorem 5 cannot be expressed in closed form, inversion using (14) may be difficult. However, one may obtain approximations to the distribution function using the first few cumulants given by (13) and a Gram–Charlier-type expansion using Laguerre instead of Hermite polynomials. For $x > 0$

$$1 - \Omega_{\psi,H}(x) = e^{-x} \sum_{n=0}^{\infty} c_n P_n(x) ,$$

where

$$P_n(x) = \sum_{m=0}^{n} (-1)^m \binom{n}{n-m} \sum_{k=0}^{m} \frac{x^{m-k}}{m-k!}$$

and c_n is a constant obtained by replacing the kth power of the variable in the nth Laguerre polynomial with the kth moment of $\Omega_{\psi,H}(\cdot)$. To facilitate computation of the cumulants when the weight function is $\psi(t) = at^k$, $k > -2$, we note the following formulas:

$$\begin{aligned}\sum_{m=1}^{\infty} [j\{-(k+1)/(k+2); m\}]^{-2n} &= 2^{-2}(k+2) & n &= 1 ,\\ &= 2^{-4}(k+2)^3(k+3)^{-1} & n &= 2 ,\\ &= 2^{-5}(k+2)^5(k+3)^{-1}(2k+5)^{-1} & n &= 3 .\end{aligned}$$

Exact values of the cumulants for $n > 3$ can be obtained by computing the moments for T_k, but good approximations result by using only the first term in the above sum. To obtain an indication of the accuracy of this approximation technique we used the last four quantiles of Table 1 for $k = -1, 0$ as arguments in the Laguerre expansion truncated at $n = 6$. The results are compared with the exact probabilities in Table 2.

TABLE 2

Laguerre-type approximations to $\Omega_{\psi_k}(\cdot)$ using quantiles of Table 1

Exact Probabilities	$k = -1$	$k = 0$
0.90	0.9135	0.9092
0.95	0.9517	0.9535
0.975	0.9711	0.9738
0.99	0.9878	0.9876

For an example under an alternative hypothesis we let $\psi(t) \equiv 1$ and $H(t) = t$. Then

$$\lambda_k = 4\{(2k+1)\pi\}^{-2} , \qquad g_k(t) = 2^{\frac{1}{2}} \cos\{(2k+1)\pi/2(t-1)\} ,$$

$$\nu_k = 2^{\frac{1}{2}}\lambda_k , \qquad \prod_{k=1}^{\infty} (1 - 2it\lambda_k)^{-\frac{1}{2}} = (\cos(2it)^{\frac{1}{2}})^{-\frac{1}{2}} \qquad \text{and}$$

$$K_n = 2^{3n-2}(2^{2n} - 1)(1 + n2^{\frac{1}{2}})(n-1)!\,|B_{2n}|/(2n)!$$

where B_j is the jth Bernoulli number. Using the Laguerre expansion truncated at $n = 6$ we obtain as approximations to the 0.90, 0.95, 0.975 and 0.99 quantiles the values 2.68, 3.55, 4.37 and 5.32 respectively.

Acknowledgments. The author wishes to thank a referee for drawing attention to the need for care in dealing with the functionals defined by $\psi(s) = as^k$, $-2 < k \leqq -1$ and for indicating the class of weight functions under which Theorems 1–4 are proved.

REFERENCES

[1] 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.

[2] ANDERSON, T. W. and DARLING, D. A. (1954). A test of goodness of fit. *J. Amer. Statist. Assoc.* **49** 765–769.

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

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

[5] DARLING, D. A. (1957). The Kolmogorov-Smirnov Cramér-von Mises tests. *Ann. Math. Statist.* **28** 823–838.

[6] FELLER, W. (1951). The asymptotic distribution of the range of sums of independent variables. *Ann. Math. Statist.* **22** 427–432.

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

[8] GRAD, A. and SOLOMON, H. (1955). Distribution of some quadratic forms and some applications. *Ann. Math. Statist.* **26** 464–477.

[9] KANDER, Z. and ZACKS, S. (1966). Test procedures for possible changes in parameters of statistical distributions occurring at unknown time points. *Ann. Math. Statist.* **37** 1196–1210.

[10] LEHMANN, E. L. (1959). *Testing Statistical Hypotheses.* Wiley, New York.

[11] LEWIS, P. A. W. (1961). Distribution of the Anderson-Darling statistic. *Ann. Math. Statist.* **32** 1118–1123.

[12] MACNEILL, I. B. (1971). Limit processes for co-spectral and quadrature spectral distribution functions. *Ann. Math. Statist.* **42** 81–96.

[13] NADLER, J. and ROBBINS, N. B. (1971). Some characteristics of Page's two-sided procedure for detecting a change in a location parameter. *Ann. Math. Statist.* **42** 538–551.

[14] PAGE, E. S. (1955). A test for a change of parameters occurring at an unknown point. *Biometrika* **42** 523–527.

[15] PAGE, E. S. (1957). On problems in which a change in a parameter occurs at an unknown point. *Biometrika* **44** 248–252.

[16] ROTHMAN, E. D. and WOODROOFE, M. (1972). A Cramér-von Mises type statistic for testing symmetry. *Ann. Math. Statist.* **43** 2035–2038.

[17] SEN, A. (1971). Tests for change in mean and a sequential ranking procedure. Unpublished Ph. D. dissertation. Univ. of Toronto.

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

[19] STEPHENS, M. A. (1963). The distribution of the goodness-of-fit statistic, U_N^2. I. *Biometrika* **50** 303–313.

[20] STEPHENS, M. A. (1964). The distribution of the goodness-of-fit statistic, U_N^2. II. *Biometrika* **51** 393–397.

[21] WATSON, G. S. (1961). Goodness-of-fit tests on a circle. *Biometrika* **48** 109–115.

DEPARTMENT OF APPLIED MATHEMATICS
UNIVERSITY OF WESTERN ONTARIO
LONDON, ONTARIO, CANADA