The Annals of Mathematical Statistics
1968, Vol. 39, No. 3, 840–854

INFERENCE PROBLEMS ABOUT PARAMETERS WHICH ARE SUBJECTED TO CHANGES OVER TIME[1]

BY CHANDAN K. MUSTAFI[2]

Columbia University

Introduction. Consider a situation where a random variable is observed sequentially over time and the distribution of this random variable is subject to a possible change at every point in the sequence. We discuss some problems connected with this situation.

In the first two sections, we assume that the change is random in nature and affects the mean of the distribution. The study of this problem is centered about a model introduced by Chernoff and Zacks [1]. We first consider the problem of estimating the current value of the mean on the basis of a set of observations taken up to present. The problem has been treated in some detail before [1] but it has been assumed throughout implicitly that certain parameters occurring in the model are known. In Section one, we derive a procedure for estimating the current value of the mean on the basis of a set of observations taken at successive time points when nothing is known about the other parameters occurring in the model.

Section two considers another important aspect of the problem, namely, to estimate the various points of change. We handle the problem in the framework of empirical Bayes procedure and use an idea similar to that of Tainiter [6] to derive a sequence of tests to be applied at each stage. This sequence of tests will be shown to be "asymptotically reasonable" in a certain sense.

In Section three, we consider n independent observations of a random variable taken at successive time points. It is further assumed that the distribution of the random variable belongs to the one parameter exponential family. We examine the problem of testing the equality of these n parameters against the alternative that the parameter has changed r-times at some unknown points where r is some finite positive integer less than n. A test procedure is obtained by generalizing the techniques used by Kander and Zacks [2] who studied the case $r = 1$. Under quite general conditions, the distribution of the test statistic is shown to be asymptotically normal both under the null and the alternative hypotheses. This kind of problem has also been studied by Page [5].

1. Estimation of the present value of the mean. Let $\mathbf{x}' = (x_1, x_2, \cdots, x_n)$ be the observation vector where the observations are taken at n successive time points. Then following Chernoff and Zacks [1], the co-ordinates of $\mathbf{x}$ satisfy the

Received 19 April 1967; revised 26 October 1967.

[1] This is a part of the author's doctoral dissertation written at Columbia University. The research was sponsored by grant no. GK-695 of the National Science Foundation.

[2] Now at Rutgers—The State University.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.* www.jstor.org

following relations:

$$x_i = \mu_n + \epsilon_i + \textstyle\sum_{k=i}^{n-1} J_k Z_k \qquad (i = 1, 2, \cdots, n-1) \tag{1}$$
$$= \mu_n + \epsilon_n \qquad (i = n),$$

where μ_n is the present value of the mean. We assume that the variables ϵ_i $(i = 1, 2, \cdots, n)$, J_i $(i = 1, 2, \cdots, n-1)$, Z_i $(i = 1, 2, \cdots, n-1)$ are all independent with

$$E(\epsilon_i) = 0; \quad V(\epsilon_i) = \lambda^2 \quad (i = 1, 2, \cdots, n)$$
$$E(Z_i) = 0; \quad V(Z_i) = \sigma^2 \quad (i = 1, 2, \cdots, n-1) \tag{2}$$
$$P(J_i = 1) = P = 1 - P(J_i = 0) \qquad (i = 1, 2, \cdots, n-1).$$

We, however, do not make any assumption about the actual form of the distribution of the random variables. The random variable J_i assumes the value 1 if there is a change between the time points i and $(i + 1)$ while it assumes the value 0 in case there is no change between these points. The random variable Z_i represents the amount of change when a change is present between the points i and $(i + 1)$. The dispersion matrix $\mathbf{V}$ of $\mathbf{x}$ is [1]:

$$\mathbf{V} = \lambda^2 \mathbf{I}_n + \sigma^2 P \textstyle\sum_{k=1}^{n-1} \mathbf{W}_n^{(k)} \tag{3}$$

where $\mathbf{I}_n$ is the $n \times n$ identity matrix; $\mathbf{W}_n^{(k)}$ is an $n \times n$ matrix whose upper left $k \times k$ submatrix consists of elements equal to 1 and all of whose other elements are zero. The BLUE (best linear unbiased estimate) of μ_n on the basis of n observations is given by [1]

$$\hat{\mu}_n = [\textstyle\sum_{i=1}^{n-1} \xi_i^{(n-1)} + 1]^{-1} [\sum_{i=1}^{n-1} \xi_i^{(n-1)} x_i + x_n] \tag{4}$$

with

$$V(\hat{\mu}_n) = \lambda^2 [\textstyle\sum_{i=1}^{n-1} \xi_i^{(n-1)} + 1]^{-1} \tag{5}$$

where

$$\xi_1^{(n-1)} = [v_1 v_2 \cdots v_{n-2}(v_{n-1} - 1)]^{-1}, \tag{6}$$
$$\xi_i^{(n-1)} = (v_{i-1} - 1)[v_{i-1} v_i \cdots v_{n-2}(v_{n-1} - 1)]^{-1} \quad (i = 2, 3, \cdots, n-1)$$

with

$$v_1 = 2 + c, \qquad v_i = 2 + c - v_{i-1}^{-1} \qquad (i = 2, 3, \cdots, n-1) \tag{7}$$

where $c = \sigma^2 P \lambda^{-2} > 0$.

Suppose, r is some positive integer less than n. Then from (3), $\mathbf{V}$ can be rewritten as

$$\mathbf{V} = \begin{pmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \lambda^2 \mathbf{I}_r + \sigma^2 P \sum_{k=1}^{r-1} \mathbf{W}_r^{(k)} \end{pmatrix} \tag{8}$$

where $\mathbf{V}_{11}$, $\mathbf{V}_{12}$ and $\mathbf{V}_{21}$ are matrices of orders $n - r \times n - r$, $n - r \times r$ and $r \times n - r$ respectively.

Equation (8) shows that the dispersion matrix of the last r observations can be obtained replacing n by r in equation (3). Consequently, the BLUE $\hat{\mu}_{n,r}$ of μ_n and its variance $V(\hat{\mu}_{n,r})$ based on the last r observations can be obtained from equations (4), (5) and (6) replacing n by r. We now prove some results concerning the limiting behavior of the sequences $\{v_k\}$ and $\{\sum_{i=1}^{r} \xi_i^{(r)}\}$.

LEMMA 1.1. $\lim_{k\to\infty} v_k = A$, *where*

$$A = 1 + \tfrac{1}{2}c[1 + (1 + 4/c)^{\frac{1}{2}}]. \tag{9}$$

PROOF. We assert that the sequence $\{v_k\}$ is monotone decreasing and bounded below by 1. Suppose, $v_k > 1$; then $-v_k^{-1} > -1$.

$$v_{k+1} = 2 + c - v_k^{-1} > 1 + c > 1.$$

Since $v_1 > 1$, the second part of the assertion follows by induction. Again, let us suppose that $v_{k-1} > v_k$; then $-v_{k-1}^{-1} > -v_k^{-1}$. Hence $v_{k+1} = 2 + c - v_k^{-1} < 2 + c - v_{k-1}^{-1} = v_k$.

Since, $v_1 > v_2$, the first part of the assertion also follows by induction. Since $\{v_k\}$ is monotone decreasing and bounded below it must converge to a limit. Let L be this limit. Then from (7), we observe that L must satisfy the condition

$$L^2 - (2 + c)L + 1 = 0. \tag{10}$$

It is easy to see [4] that the two roots of equation (10) are A and A^{-1} where A is defined by equation (9). Since $v_k > 1$ for each k, $L > 1$. Hence, $L = A$ which completes the proof of the lemma.

LEMMA 1.2

$$\lim_{r\to\infty} \sum_{i=1}^{r} \xi_i^{(r)} = (A - 1)^{-1}.$$

PROOF. Substituting $n = r$ and $n = r + 1$ in equation (6), we obtain

$$\xi_i^{(r)}/\xi_i^{(r-1)} = \xi_r^{(r)}, \qquad i = 1, 2, \cdots r - 1.$$

Hence,

$$\sum_{i=1}^{r} \xi_i^{(r)} = \xi_r^{(r)} + \xi_r^{(r)} \sum_{i=1}^{r-1} \xi_i^{(r-1)}. \tag{11}$$

By Lemma 1.1 and equation (6), $\xi_i^{(r)} > 0$ for each i and r. Also

$$\xi_r^{(r)}/\xi_{r-1}^{(r-1)} = [(v_{r-1} - 1)/v_{r-1}(v_r - 1)][v_{r-2}(v_{r-1} - 1)/(v_{r-2} - 1)] < 1$$

if and only if

$$v_{r-2} > v_{r-1}. \tag{12}$$

Equation (12) is always true by virtue of Lemma 1.1. Hence, $\{\xi_r^{(r)}\}$ is a monotone decreasing sequence. Also, $\xi_r^{(r)} > 0$ for each r. Therefore, $\{\xi_r^{(r)}\}$ converges to a limit. Using Lemma 1.1 and the relation $\xi_r^{(r)} = (v_{r-1} - 1)/v_{r-1}(v_r - 1)$, we obtain

$$\lim_{r\to\infty} \xi_r^{(r)} = A^{-1} \tag{13}$$

Consider now the relation between $V(\hat{\mu}_{n,r})$ and $V(\hat{\mu}_{n,r+1})$. $\hat{\mu}_{n,r+1}$ is the BLUE based on the last $(r+1)$ observations. $\hat{\mu}_{n,r}$, the BLUE based on the last r observations, is also a linear unbiased estimate based on last $(r+1)$ observations. Therefore, $V(\hat{\mu}_{n,r}) \geqq V(\hat{\mu}_{n,r+1})$ which from (5) leads to

$$\sum_{i=1}^{r-1} \xi_i^{(r-1)} \leqq \sum_{i=1}^{r} \xi_i^{(r)}. \tag{14}$$

Hence, $\{\sum_{i=1}^{r} \xi_i^{(r)}\}$ is a monotone increasing sequence. Using the expression given in equation (11) repeatedly, we obtain

$$\sum_{i=1}^{r} \xi_i^{(r)} = \xi_r^{(r)} + \xi_r^{(r)}\xi_{r-1}^{(r-1)} + \cdots + \xi_r^{(r)}\xi_{r-1}^{(r-1)} \cdots \xi_1^{(1)}.$$

Using the fact $\{\xi_r^{(r)}\}$ is monotone decreasing and $\xi_1^{(1)} = (1+c)^{-1} < 1$ we obtain

$$\sum_{i=1}^{r} \xi_i^{(r)} \leqq \xi_1^{(1)}(1-\xi_1^{(1)})^{-1} \leqq c^{-1}. \tag{15}$$

Equation (15) shows that the sequence $\{\sum_{i=1}^{r} \xi_i^{(r)}\}$ is bounded above by c^{-1}. Since, $\{\sum_{i=1}^{r} \xi_i^{(r)}\}$ is monotone increasing it must converge to a limit. From equations (11) and (13), it is easy to see that this limit is $(A-1)^{-1}$ which completes the proof of the lemma.

Lemma 1.2 shows the result of using a finite number of recent observations for the estimation of the present value of the mean when there is an infinite sequence of observations $\{x_n\}$ and the present time-point is at infinity. It follows from Lemma 1.2 that the sequence $\{V(\hat{\mu}_{n,r})\}$ decreases to the quantity $\lambda^2(1-A^{-1})$ with increasing r. Thus, if we choose a finite number r_0 which ensures that $V(\hat{\mu}_{n,r_0})$ is sufficiently close to $\lambda^2(1-A^{-1})$ then by including an additional number of finite observations from the end and using the corresponding BLUE, we cannot substantially improve its variance. The remaining $n-r_0$ observations, however, contain information about c and can, therefore, be used for its estimation. This will enable us to obtain an estimate of μ_n when nothing is known about λ^2, σ^2 and P. We first describe a procedure of estimating c in the following lemma:

LEMMA 1.3. *Suppose, $\{\epsilon_j\}$ and $\{Z_i\}$ are each sequences of identically distributed random variables. Let*

$$y_j = x_j - x_{j+1} \qquad (j = 1, 2, \cdots, n-1),$$

$$u_j = y_j - y_{j+1} = x_j - 2x_{j+1} + x_{j+2} \qquad (j = 1, 2, \cdots, n-2). \tag{16}$$

$$S_1^2 = (n-1)^{-1}\sum_{j=1}^{n-1} y_j^2, \qquad S_2^2 = (n-2)^{-1}\sum_{j=1}^{n-1} u_j^2.$$

Then, $(6S_1^2 - 2S_2^2)(S_2^2 - 2S_1^2)^{-1}$ converges to c almost surely.

PROOF. From (1)

$$y_j = J_jZ_j + \epsilon_j - \epsilon_{j+1} \qquad (j = 1, 2, \cdots, n-1).$$

It is easy to see that because of the assumptions made in the lemma $\{y_j\}$ is a 1-dependent marginally strictly stationary sequence (see Tainiter [6] for definition). It also follows that [6]

$$(n-1)^{-1}\sum_{j=1}^{n-1} y_j^2 = S_1^2 \to 2\lambda^2 + \sigma^2P \quad \text{a.s.} \tag{17}$$

Likewise $\{u_j\}$ is a 2-dependent marginally strictly stationary sequence and

(18) $$(n-2)^{-1}\sum_{j=1}^{n-2} u_j^2 = S_2^2 \rightarrow 6\lambda^2 + 2\sigma^2 P \quad \text{a.s.}$$

The proof of the lemma is immediate from equations (17) and (18).

We are now in a position to prove the main result of this section. This is given in the following theorem.

THEOREM 1.1. *Suppose $r_0 < n$ is a fixed positive integer. Let*

(19) $$\begin{aligned} S_{1r_0}^2 &= (n-r_0-1)^{-1}\sum_{i=1}^{n-r_0-1}(x_i - x_{i+1})^2, \\ S_{2r_0}^2 &= (n-r_0-2)^{-1}\sum_{i=1}^{n-r_0-2}(x_i - 2x_{i+1} + x_{i+2})^2, \\ \hat{c} &= (6S_{1r_0}^2 - 2S_{2r_0}^2)(S_{2r_0}^2 - 2S_{1r_0}^2)^{-1}. \end{aligned}$$

Let the sequence $\{\hat{v}_k\}$ be defined in the same way as the sequence $\{v_k\}$ replacing c by $\hat{c}$ in equation (7). *Likewise, let us define a sequence $\hat{\xi}_i^{(r_0-1)}$ $(i = 1, 2, \cdots, r_0 - 1)$ replacing n by r_0 and v_k by $\hat{v}_k$ in equation* (6). *Let*

(20) $$\begin{aligned} a_j &= (\hat{\xi}_j^{(r_0-1)}+)(1 + \textstyle\sum_{j=1}^{r_0-1}\hat{\xi}_j^{(r_0-1)}+)^{-1}, \qquad j = 1, 2, \cdots, r_0 - 1, \\ a_{r_0} &= (1 + \textstyle\sum_{j=1}^{r_0-1}\hat{\xi}_j^{(r_0-1)}+)^{-1} \end{aligned}$$

where $\hat{\xi}_j^{(r_0-1)}+ = \max(0, \hat{\xi}_j^{(r_0-1)})$, *so that* $\sum_{j=1}^{r_0} a_j = 1$. *Consider now the estimate* t_{n,r_0} *of* μ_n *defined as*

(21) $$t_{n,r_0} = \sum_{j=1}^{r_0} a_j x_{n-r_0+j}; \quad n = r_0 + 1, r_0 + 2, \cdots.$$

Then,

(i) $E(t_{n,r_0}) = \mu_n$.
(ii) *For each finite n,* $V(t_{n,r_0}) \leqq \lambda^2 + \sigma^2 P(r_0 - 1)$.
(iii) $\lim_{n\to\infty} V(t_{n,r_0}) = \lambda^2[1 + \sum_{j=1}^{r_0-1}\xi_j^{(r_0-1)}]^{-1}$

which is the variance of the BLUE *based on the last r_0 observations.*

PROOF. First of all, we observe that $S_{1r_0}^2$ and $S_{2r_0}^2$ are functions of the random variables ϵ_j $(j = 1, 2, \cdots, n - r_0)$, J_j $(j = 1, 2, \cdots, n - r_0 - 1)$, Z_j $(j = 1, 2, \cdots, n - r_0 - 1)$. Consequently, a_j $(j = 1, 2, \cdots, r_0)$ defined in equation (20) are also functions of the same random variables. x_{n-r_0+j} $(j = 1, 2, \cdots, r_0)$ are functions of the random variables ϵ_j $(n - r_0 + 1, \cdots, n)$, J_j $(n - r_0 + 1, \cdots, n - 1)$, Z_j $(n - r_0 + 1, \cdots, n - 1)$. Hence, a_j $(j = 1, 2, \cdots, r_0)$ is independent of x_{n-r_0+j} $(j = 1, 2, \cdots, r_0)$. From (20) and (21), we obtain

$$E(t_{n,r_0}) = \mu_n \sum_{j=1}^{r_0} E(a_j) = \mu_n$$

which proves the first part of the theorem.

(22) $$\begin{aligned} V(t_{n,r_0}) = E[\textstyle\sum_{j=1}^{r_0} a_j(x_{n-r_0+j} - \mu_n)]^2 &= \textstyle\sum_{j=1}^{r_0}[\lambda^2 + \sigma^2 P(r_0 - j)]E(a_j^2) \\ &\quad + 2\textstyle\sum_{j=2}^{r_0}\sum_{j'=1}^{j-1}\sigma^2 P(r_0 - j)E(a_j a_{j'}). \end{aligned}$$

Using the condition $\sum_{j=1}^{r_0} E(a_j^2) \leqq \sum_{j=1}^{r_0} E(a_j) = 1$ we obtain

$$
\begin{aligned}
V(t_{n,r_0}) &\leqq \lambda^2 + \sigma^2 P r_0 \textstyle\sum_{j=1}^{r_0} E(a_j^2) - \sigma^2 P \sum_{j=1}^{r_0} jE(a_j^2) \\
&\quad + 2\sigma^2 P r_0 \textstyle\sum_{j=2}^{r_0} \sum_{j'=1}^{j-1} E(a_j a_{j'}) - 2\sigma^2 P \sum_{j=2}^{r_0} \sum_{j'=1}^{j-1} jE(a_j a_{j'}) \\
&= \lambda^2 + \sigma^2 P r_0 E(\textstyle\sum_{j=1}^{r_0} a_j)^2 - \sigma^2 P[\sum_{j=1}^{r_0} jE(a_j^2) + 2\sum_{j=2}^{r_0} \sum_{j'=1}^{j-1} jE(a_j a_{j'})] \\
&\leqq \lambda^2 + \sigma^2 P r_0 - \sigma^2 P E[\textstyle\sum_{j=1}^{r_0} j^{\frac{1}{2}} a_j]^2 \leqq \lambda^2 + \sigma^2 P(r_0 - 1).
\end{aligned}
$$

This proves part (ii) of the theorem. Finally, from Lemma 1.3

$$
a_j \to (\xi_j^{(r_0-1)})(1 + \textstyle\sum_{j=1}^{r_0-1} \xi_j^{(r_0-1)})^{-1} \quad \text{a.s.} \qquad (j = 1, 2, \cdots, r_0 - 1)
$$

$$
a_{r_0} \to (1 + \textstyle\sum_{j=1}^{r_0-1} \xi_j^{(r_0-1)})^{-1} \quad \text{a.s.}
$$

Further, $0 \leqq a_j^2 \leqq 1$, $0 \leqq a_j a_{j'} \leqq 1$. Hence, by the dominated convergence theorem

$$
\begin{aligned}
\lim_{n\to\infty} E(a_j^2) &= (\xi_j^{(r_0-1)})^2 (1 + \textstyle\sum_{j=1}^{r_0-1} \xi_j^{(r_0-1)})^{-2} \\
&\qquad\qquad (j = 1, 2, \cdots, r_0 - 1) \\
\lim_{n\to\infty} E(a_{r_0}^2) &= (1 + \textstyle\sum_{j=1}^{r_0-1} \xi_j^{(r_0-1)})^{-2} \\
\lim_{n\to\infty} E(a_j a_{j'}) &= (\xi_j^{(r_0-1)} \xi_{j'}^{(r_0-1)})(1 + \textstyle\sum_{j=1}^{r_0=1} \xi_j^{(r_0-1)})^{-2} \\
&\qquad\qquad (j = 2, 3, \cdots, r_0 - 1; j' = 1, 2, \cdots, j - 1) \\
\lim_{n\to\infty} E(a_{r_0} a_{j'}) &= \xi_{j'}^{(r_0-1)} (1 + \textstyle\sum_{j=1}^{r_0-1} \xi_j^{(r_0-1)})^{-2} \\
&\qquad\qquad (j' = 1, 2, \cdots, r_0 - 1).
\end{aligned} \tag{23}
$$

From (22) and (23), we obtain

$$
\begin{aligned}
\lim_{n\to\infty} V(t_{n,r_0}) &= (1 + \textstyle\sum_{j=1}^{r_0-1} \xi_j^{(r_0-1)})^{-2} [\sum_{j=1}^{r_0-1} V(x_{n-r_0+j})(\xi_j^{(r_0-1)})^2 + V(x_n) \\
&\quad + 2\textstyle\sum_{j=2}^{r_0-1} \sum_{j'=1}^{j-1} \operatorname{cov}(x_{n-r_0+j}, x_{n-r_0+j'})(\xi_j^{(r_0-1)} \xi_{j'}^{(r_0-1)})] \\
&= (1 + \textstyle\sum_{j=1}^{r_0-1} \xi_j^{(r_0-1)})^{-2} V[\sum_{j=1}^{r_0-1} \xi_j^{(r_0-1)} x_{n-r_0+j} + x_n] \\
&= \lambda^2 [\textstyle\sum_{j=1}^{r_0-1} \xi_j^{(r_0-1)} + 1]^{-1}.
\end{aligned}
$$

This completes the proof of the theorem.

Theorem 1.1 describes a procedure of estimating the present value of the mean when n is large and nothing is known about the parameters given in equation (2). The theorem asserts that for each finite positive integer r_0, it is possible to construct an estimate t_{n,r_0} of the present value of the mean which is unbiased for each n and whose variance is asymptotically equal to the variance of the BLUE based on the last r_0 observations. We also observe from Lemma 1.2 that if the value of r_0 is increased successively, then after a certain stage, the inclusion of an additional number of finite observations will not improve the asymptotic variance of t_{n,r_0} substantially. The final choice of r_0 is, thus, left to the statistician depending upon how closely he wishes to approach the minimum variance $\lambda^2(1 - A^{-1})$.

2. Estimation of the various time points of change. In this section, we consider another important aspect of the problem, namely, the estimation of all such pairs of points between which changes have taken place. It will be assumed throughout this section that the various probabilities of change are equal, $\{\epsilon_j\}$ and $\{Z_j\}$ defined in (1) are each sequences of identically distributed random variables and certain distributions (to be specified later) are known. The method which we are going to develop uses an idea similar to that of Tainiter [6].

Let

$$y_j = x_j - x_{j+1} = \epsilon_j - \epsilon_{j+1} + J_j Z_j \qquad (j = 1, 2, \cdots, n-1). \tag{24}$$

Each J_j in equation (24) can assume two values, namely, one and zero. If $J_j = 1$, there is a change between the time points j and $(j + 1)$, while if $J_j = 0$ there is no change between these two points. The problem of estimating the location of all such pairs of points between which changes have occurred can thus be reduced to the problem of testing at each stage whether J_j is zero or one. We also observe that under the assumptions made at the beginning of this section $\{y_i\}$ is a marginally strictly stationary one-dependent sequence of random variables. The problem considered here can be reduced in the framework of empirical Bayes procedure. For detail, we refer to Tainiter [6].

Consider a class C of decision rules in which the decision at the jth stage depends on y_j and y_{j-1} $(j \geqq 2)$ and for $j = 1$ it depends only on y_1. Then it follows [6] that there is a sequence of decision rules which minimizes the Bayes risk at each stage. The minimum value of this Bayes risk (also known as Bayes envelope function) for $j \geqq 2$ is independent of j and will be denoted by $R(P)$.

THEOREM 2.1. *Suppose P is unknown. Let us define the function $h(y)$ as*

$$\begin{aligned} h(y) &= (1 - P\{Y_j \geqq K \mid J_j = 0\}) \\ &\quad \cdot (P\{Y_j \geqq K \mid J_j = 1\} - P\{Y_j \geqq K \mid J_j = 0\})^{-1} \\ &= (1 - P\{\epsilon_j - \epsilon_{j+1} \geqq K\}) \\ &\quad \cdot (P\{\epsilon_j - \epsilon_{j+1} + Z_j \geqq K\} - P\{\epsilon_j - \epsilon_{j+1} \geqq K\})^{-1} \quad \text{if} \quad y \geqq K, \\ h(y) &= -P\{y_j \geqq K \mid J_j = 0\} \\ &\quad \cdot (P\{y_j \geqq K \mid J_j = 1\} - \{y_j \geqq K \mid J_j = 0\})^{-1} \\ &= -P\{\epsilon_j - \epsilon_{j+1} \geqq K\} \\ &\quad \cdot (P\{\epsilon_j - \epsilon_{j+1} + Z_j \geqq K\} - P\{\epsilon_j - \epsilon_{j+1} \geqq K\})^{-1} \quad \text{if} \quad y < K, \end{aligned} \tag{25}$$

where K is some real constant for which

$$P(\epsilon_j - \epsilon_{j+1} + Z_j \geqq K) \neq P(\epsilon_j - \epsilon_{j+1} \geqq K).$$

Let

$$\begin{aligned} \tilde{P}_j &= \textstyle\sum_{k=1}^{j} [h(y_k)]/j, \\ P_j(y_1, y_2, \cdots, y_j) &= 0 \qquad \text{if} \quad \tilde{P}_j \leqq 0 \\ &= \tilde{P}_j \qquad \text{if} \quad 0 < \tilde{P}_j < 1 \\ &= 1 \qquad \text{if} \quad \tilde{P}_j \geqq 1. \end{aligned} \tag{26}$$

Define a sequence of decision rules $\{d_j^*\}$ *as follows*:

$$(27)\qquad d_j^*(y_j) = 1 \quad \text{if} \quad \varphi_{P_j}(y_j) \geqq 0\,(\textit{conclude } J_j = 1)$$
$$= 0 \quad \text{if} \quad \varphi_{P_j}(y_j) < 0\,(\textit{conclude } J_j = 0) \qquad (j \geqq 2)$$

where a_0 *and* a_1 *are the losses corresponding to the wrong decisions,*

$$(28)\quad \varphi_{P_j}(y_j) = P_j a_1[(1 - P_j)f_1^0(y_j \mid y_{j-1}) + P_j f_1^1(y_j \mid y_{j-1})] - (1 - P_j)a_0[(1 - P_j)f_0^0(y_j \mid y_{j-1}) + P_j f_0^1(y_j \mid y_{j-1})]$$

and $f_\rho^{\,i}(y_j \mid y_{j-1})$ = *conditional density of* y_j *given* y_{j-1} *when* $J_j = \rho$ *and* $J_{j-1} = i$ $(i, \rho = 0, 1)$. *Let* $R_j^*(P)$ *be the risk of the procedure at the* jth *stage. Then*

$$\lim_{j\to\infty} R_j^*(P) = R(P).$$

PROOF. The proof of the theorem is essentially the same as the theorem given in Section 4 of Tainiter [6] with $r = 1$ and is, therefore, omitted.

Theorem 2.1 outlines a procedure to test for posible change at each stage when P is unknown but the distributions of the sequences $\{\epsilon_j\}$ and $\{Z_j\}$ are known. In this case we can compute the sequence $\{\varphi_{P_j}(y_j)\}$ as given in equation (28) for each $j \geqq 2$. The theorem asserts that if we apply these tests successively at each stage, the risk will converge to $R(P)$ which is the Bayes envelope function in the class C consisting of all decision rules with the property that the decision about J_j depends on y_j and y_{j-1} and P is known.

We now consider an example where the sequence $\{\varphi_{P_j}(y_j)\}$ defined in (28) can be reduced to a more explicit form. Let ϵ_j be normally distributed with mean 0 and variance λ^2 and Z_j be normally distributed with mean 0 and variance σ^2. These assumptions are similar to those made by Chernoff and Zacks [1].

Let

$$(29)\qquad \Phi(x) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{x} e^{-t^2/2}\,dt.$$

From (25), we obtain the sequence $\{h(y_j)\}$ as:

$$h(y_j) = \Phi(K(2\lambda^2)^{-\frac{1}{2}})/[\Phi(K(2\lambda^2)^{-\frac{1}{2}}) - \Phi(K(2\lambda^2 + \sigma^2)^{-\frac{1}{2}})] \quad \text{if} \quad y_j \geqq K$$
$$(30)\qquad = [\Phi(K(2\lambda^2)^{-\frac{1}{2}}) - 1]\,[\Phi(K(2\lambda^2)^{-\frac{1}{2}}) - \Phi(K(2\lambda^2 + \sigma^2)^{-\frac{1}{2}})]^{-1} \quad \text{if} \quad y_j < K$$

where K is some real constant for which $\Phi(K(2\lambda^2)^{-\frac{1}{2}}) \neq \Phi(K(2\lambda^2 + \sigma^2)^{-\frac{1}{2}})$ i.e., $K \neq 0$. From a practical point of view, K can either be taken as a small negative number or a large positive number so that we can confine ourselves to one of the two definitions of $h(y_j)$ given in (30). From (26) and (30), we can compute the sequence $\{P_j\}$. Also,

$$f_0^0(y_j \mid y_{j-1}) \text{ is } N(-\tfrac{1}{2}y_{j-1}, \tfrac{3}{2}\lambda^2),$$
$$f_0^1(y_j \mid y_{j-1}) \text{ is } N(-\lambda^2(2\lambda^2 + \sigma^2)^{-1}y_{j-1}, \lambda^2(3\lambda^2 + \sigma^2)(2\lambda^2 + \sigma^2)^{-1}),$$
$$(31)\quad f_1^0(y_j \mid y_{j-1}) \text{ is } N(-\tfrac{1}{2}y_{j-1}, \tfrac{3}{2}\lambda^2 + \sigma^2),$$
$$f_1^1(y_j \mid y_{j-1}) \text{ is } N(-\lambda^2(2\lambda^2 + \sigma^2)^{-1}y_{j-1}, (3\lambda^2 + \sigma^2)(\lambda^2 + \sigma^2)(2\lambda^2 + \sigma^2)^{-1}).$$

Using the sequence $\{P_j\}$ and equation (31) we obtain the sequence of tests $\{d_j^*\}$ defined in (27) to be applied at each successive stage.

3. Test procedure for possible changes at unknown time-points. The present section is devoted to the generalization of a problem studied by Page [5], Chernoff and Zacks [1], Kander and Zacks [2]. The problem can be described as follows: Given observations on independent random variables $x_1, x_2, \cdots, x_n$ (taken at consecutive time points) with density $f(x, \theta_i)$ $(i = 1, 2, \cdots, n)$, it is required to test the hypothesis $H_0: \theta_1 = \theta_2 = \cdots = \theta_n = \theta$ (known) against the alternative

$$(32)\qquad H_1: \begin{cases} \theta_1 = \theta_2 = \cdots = \theta_{m_1} = \theta \\ \theta_{m_1+1} = \theta_{m_1+2} = \cdots = \theta_{m_2} = \theta + \delta \\ \vdots \\ \theta_{m_r+1} = \theta_{m_r+2} = \cdots = \theta_n = \theta + r\delta \end{cases}$$

where r is known but $m_1, m_2, \cdots, m_r$ and δ are unknown. The hypothesis H_1 has the following interpretation: During the period 1 to n the parameter θ has changed r-times between the time points $(m_1, m_1 + 1), \cdots, (m_r, m_r + 1)$. The points $m_1, m_2, \cdots, m_r$ as well as the amount of change are unknown. However, if the parameter θ changes between any two points the amount of change is fixed and is equal to δ. Obviously, $1 \leqq m_1 < m_2 < \cdots < m_r < n$.

In the present study, we make the following four assumptions regarding the density $f(x, \theta)$, the fixed amount of change δ, and the unknown time points $m_1, m_2, \cdots, m_r$.

1. $f(x, \theta) = h(x) \exp [\Psi_1(\theta)U(x) + \Psi_2(\theta)]$.
2. $\Psi_1(\theta)$ is monotone increasing in θ and $\Psi_1'(\theta)$, $\Psi_2'(\theta)$, $\Psi_1''(\theta)$, $\Psi_2''(\theta)$ are finite over the whole range of θ.
3. Terms of order $o(\delta)$ can be neglected (where $o(\delta)$ denotes all such terms for which $[o(\delta)]\delta^{-1} \to 0$ as $\delta \to 0$).
4. $(m_1, m_2, \cdots, m_r)$ is a random vector with the following probability distribution

$$(33)\qquad \begin{aligned} P(m_1) &= (n - r)^{-1}, & m_1 &= 1, 2, \cdots, n - r, \\ &= 0, & &\text{otherwise.} \end{aligned}$$

$$(34)\qquad \begin{aligned} &P(m_j \mid m_{j-1}, m_{j-2}, \cdots, m_1) \\ &\quad = P(m_j \mid m_{j-1}) \\ &\quad = \begin{cases} (n - m_{j-1} - r + j - 1)^{-1}, & m_j = m_{j-1} + 1, \cdots, n - r + j - 1, \\ 0, & \text{otherwise} \quad (j = 2, 3, \cdots, r). \end{cases} \end{aligned}$$

The first three assumptions are similar to those made by Kander and Zacks [2] while the fourth assumption is a natural extension of an assumption also given in [2]. Its interpretation is that each possible location of jth jump is equally likely once the location of the $(j - 1)$st jump is fixed.

In various references mentioned at the beginning of this section, it has been assumed that $r = 1$. The results of the present study is valid for any finite r and includes the result given in [2] as a special case.

The joint density of $x_1, x_2, \cdots, x_n$ and $m_1, m_2, \cdots, m_r$ under H_1 (using the four assumptions) after some simple calculations is given by:

$$\begin{aligned} f(x_1, \cdots, x_n, m_1, \cdots, m_r) \\ = P(m_1)[\textstyle\prod_{j=2}^{r} P(m_j \mid m_{j-1})][\prod_{i=1}^{n} h(x_i)][\exp\{\sum_{i=1}^{n} g(x_i, \theta)\}] \\ \cdot[1 + \delta \textstyle\sum_{i=m_1+1}^{m_2} g'(x_i, \theta) + \cdots + r\delta \sum_{i=m_r+1}^{n} g'(x_i, \theta) + o(\delta)], \end{aligned} \tag{35}$$

where $g(x_i, \theta) = \Psi_1(\theta)U(x_i) + \Psi_2(\theta)$.

The marginal density of $x_1, x_2, \cdots, x_n$ under H_1 can be obtained by summing equation (35) over $m_1, m_2, \cdots, m_r$. The marginal density of $x_1, x_2, \cdots, x_n$ under H_0 is

$$[\textstyle\prod_{i=1}^{n} h(x_i)][\exp\{\sum_{i=1}^{n} g(x_i, \theta)\}]. \tag{36}$$

From (35) and (36) we observe that the likelihood ratio can be expressed as

$$\begin{aligned} \textstyle\sum_{m_1=1}^{n-r} \sum_{m_2=m_1+1}^{n-r+1} \cdots \sum_{m_r=m_{r-1}+1}^{n-1} [P(m_1)][\prod_{j=2}^{r} P(m_j \mid m_{j-1})] \\ \cdot[1 + \delta \textstyle\sum_{i=m_1+1}^{m_2} g'(x_i, \theta) + \cdots + r\delta \sum_{i=m_r+1}^{n} g'(x_i, \theta) + o(\delta)]. \end{aligned} \tag{37}$$

Now,

$$g'(x_i, \theta) = \Psi_1'(\theta)U(x_i) + \Psi_2'(\theta).$$

By assumption 2, $\Psi_1'(\theta) > 0$ and by assumption 3 terms of order $o(\delta)$ can be neglected. Hence, from (37), the quantity

$$\begin{aligned} T = \textstyle\sum_{m_1=1}^{n-r} \sum_{m_2=m_1+1}^{n-r+1} \cdots \sum_{m_r=m_{r-1}+1}^{n-1} [P(m_1)][\prod_{j=2}^{r} P(m_j \mid m_{j-1})] \\ [\textstyle\sum_{i=m_1+1}^{m_2} U_i + 2\sum_{i=m_2+1}^{m_3} U_i + \cdots + r\sum_{i=m_r+1}^{n} U_i] \end{aligned} \tag{38}$$

(where $U_i = U(x_i)$] is a monotone increasing or decreasing function of the likelihood ratio according as δ is positive or negative. We propose T as a test statistic for testing H_0 against H_1.

The expression for T given in (38) cannot be computed directly since m_1, $m_2, \cdots, m_r$ are unknown. We first develop a procedure which will enable us to express T in terms of n, r and U_j ($j = 1, 2, \cdots, n$). Next we shall study the asymptotic distribution of T (with fixed r) as n becomes infinite.

Let us define a sequence of functions $W_j(s, l, i)$ ($j = 2, \cdots, r$) by

$$\begin{aligned} W_2(s, l, i) = (s - l)^{-1}i + (s - l - 1)^{-1}(i - 1) \\ + \cdots + (s - l - i + 1)^{-1}, \\ i = 1, 2, \cdots, s - l; \quad s, l = 0, 1, 2, \cdots, \text{with } s > l, \end{aligned}$$

$$W_{k+1}(s, l, i) = (s - l)^{-1}W_k(s + 1, l + 1, i) + (s - l - 1)^{-1} \tag{39}$$

$$\cdot W_k(s+1, l+2, i-1) + \cdots + (s-l-i+1)^{-1}$$
$$\cdot W_k(s+1, l+i, 1),$$
$$i = 1, 2, \cdots, s-l; \quad k = 2, 3 \cdots.$$

The proof of the following three lemmas can be verified in a straightforward way.

LEMMA 3.1. *Let l, s, t be positive integers with $t > s > l$. Then*

$$(40) \quad \sum_{m=l+1}^{s} \sum_{i=m+1}^{t} U_i = \sum_{i=1}^{s-l} iU_{l+i+1} + (s-l)(U_{s+2} + \cdots + U_t).$$

(If $t = s + 1$, the second term on the right hand side of (40) is zero).

LEMMA 3.2. *Let l, s be positive integers with $s > l$. Then*

$$\sum_{m=l+1}^{s} (s-m+1)^{-1} \sum_{i=1}^{s-m+1} iU_{m+1+i} = \sum_{i=1}^{s-l} W_2(s, l, i) U_{l+2+i}.$$

LEMMA 3.3. *Let l, s be positive integers with $s > l$. Then*

$$\sum_{m=l+1}^{s} (s-m+1)^{-1} \sum_{i=1}^{s-m+1} W_k(s+1, m, i) U_{m+k+i}$$
$$= \sum_{i=1}^{s-l} W_{k+1}(s, l, i) U_{l+k+1+i}.$$

THEOREM 3.1.

$$(41) \quad (n-r)T = \sum_{j=1}^{n-r} jU_{j+1} + \sum_{j=2}^{r} \sum_{i=1}^{n-r} W_j(n-r, 0, i) U_{j+i}$$
$$+ (n-r) \sum_{j=1}^{r-1} jU_{n-r+j+1}.$$

PROOF. From (33) and (38) we obtain

$$(42) \quad \begin{aligned} (n-r)T &= \sum_{m_1=1}^{n-r} (n-m_1-r+1)^{-1} \sum_{m_2=m_1+1}^{n-r+1} \cdots (n-m_{r-1}-1)^{-1} \\ &\quad \cdot \sum_{m_r=m_{r-1}+1}^{n-1} [\sum_{i=m_1+1}^{n} U_i + \sum_{i=m_2+1}^{n} U_i + \cdots + \sum_{i=m_j+1}^{n} U_i \\ &\quad + \cdots + \sum_{i=m_r+1}^{n} U_i] \\ &= T_1 + T_2 + \cdots + T_j + \cdots + T_r \end{aligned}$$

where

$$\begin{aligned} T_j &= \sum_{m_1=1}^{n-r} (n-m_1-r+1)^{-1} \sum_{m_2=m_1+1}^{n-r+1} \cdots (n-m_{r-1}-1)^{-1} \\ &\quad \cdot \sum_{m_r=m_{r-1}+1}^{n-1} \sum_{i=m_j+1}^{n} U_i \\ &= \sum_{m_1=1}^{n-r} (n-m_1-r+1)^{-1} \sum_{m_2=m_1+1}^{n-r+1} \cdots (n-m_{j-1}-r+j-1)^{-1} \\ &\quad \cdot \sum_{m_j=m_{j-1}+1}^{n-r+j-1} \sum_{i=m_j+1}^{n} U_i. \end{aligned}$$

Substituting $m = m_j$; $l = m_{j-1}$, $s = n - r + j - 1$, $t = n$ and using Lemma 3.1, we obtain

$$(43) \quad \sum_{m_j=m_{j-1}+1}^{n-r+j-1} \sum_{i=m_j+1}^{n} U_i = \sum_{i=1}^{n-r-m_{j-1}+j-1} iU_{m_{j-1}+i+1}$$
$$+ (n-r-m_{j-1}+j-1)[U_{n-r+j+1} + \cdots + U_n].$$

From (42) and (43) we obtain

$$T_j = \sum_{m_1=1}^{n-r} (n - m_1 - r + 1)^{-1} \cdot \sum_{m_2=m_1+1}^{n-r+1} \cdots \sum_{m_{j-1}=m_{j-2}+1}^{n-r+j-2} (n - m_{j-1} - r + j - 1)^{-1} \cdot \sum_{i=1}^{n-m_{j-1}-r+j-1} iU_{m_{j-1}+i+1} + (n - r)[U_{n-r+j+1} + \cdots + U_n]. \quad (44)$$

Let $m = m_{j-1}$; $l = m_{j-2}$; $s = n - r + j - 2$. Then applying Lemma 3.2, we obtain

$$\sum_{m_{j-1}=m_{j-2}+1}^{n-r+j-2} (n - m_{j-1} - r + j - 1)^{-1} \sum_{i=1}^{n-m_{j-1}-r+j-1} iU_{m_{j-1}+i+1} = \sum_{i=1}^{n-r-m_{j-2}+j-2} W_2(n - r + j - 2, m_{j-2}, i)U_{m_{j-2}+2+i}.$$

It is clear that we can apply the above technique repeatedly and use Lemma 3.3 till all the summations of equation (44) are exhausted. Thus, we finally end up with

$$T_j = \sum_{i=1}^{n-r} W_j(n - r, 0, i)U_{j+i} + (n - r)[U_{n-r+j+1} + \cdots + U_n], \quad j = 2, 3, \cdots, r; \quad (45)$$
$$T_1 = \sum_{m_1=1}^{n-r} \sum_{i=m_1+1}^{n} U_i .$$

Let $m_1 = m$; $l = 0$; $s = n - r$; $t = n$; then applying Lemma 3.1 we obtain

$$T_1 = \sum_{j=1}^{n-r} jU_{j+1} + (n - r)[U_{n-r+2} + \cdots + U_n]. \quad (46)$$

Summing equation (45) from $j = 2, \cdots, r$ and adding the expression for T_1 from (46), we obtain

$$(n - r)T = \sum_{j=1}^{n-r} jU_{j+1} + \sum_{j=2}^{r} \sum_{i=1}^{n-r} W_j(n - r, 0, i)U_{j+i} + (n - r) \sum_{j=1}^{r-1} jU_{n-r+j+1} \quad (47)$$

which completes the proof of the theorem.

Lemma 3.4. $W_j (s, l, i) \leqq s - l$ $(1 \leqq i \leqq s - l; j \geqq 2)$.

Proof. The lemma can be proved by induction.

Lemma 3.5.

$$\sum_{i=1}^{s-l} W_j(s, l, i) = 2^{-j}(s - l)(s - l + 2^j - 1) \qquad (j = 2, 3, \cdots).$$

Proof. The lemma can be proved by induction.

Lemma 3.6. *Let* $E(U_i) = \mu$ *under* H_0. *Then, when* H_0 *is true*

$$E(T) = \mu[\tfrac{1}{2}(n + r^2 - 1) + 2^{-r}(n - r - 1)(2^{r-1} - 1)]. \quad (48)$$

Proof. The proof of this lemma can be obtained in a straightforward way by the application of Lemma 3.5 in equation (41).

We are now in a position to prove the asymptotic normality of the test statistic T. This is given in the following theorem.

Theorem 3.2. *Suppose* m_1, m_2, $\cdots$, m_r *are the points where the parameter* θ

has changed and δ *is the fixed amount of change occurring at these points. Let the mean and the variance of* U_j $(j = 1, 2, \cdots, n)$ *for given* $m_1, m_2, \cdots, m_r$ *and* δ *be as follows:*

$$
\begin{aligned}
E(U_j) &= \mu, & j &= 1, 2, \cdots, m_1, \\
&= \mu_1, & j &= m_1 + 1, \cdots, m_2, \\
&\vdots & &\vdots \\
&= \mu_r, & j &= m_r + 1, \cdots, n.
\end{aligned} \tag{49}
$$

$$
\begin{aligned}
V(U_j) &= \sigma^2, & j &= 1, 2, \cdots, m_1, \\
&= \sigma_1^2, & j &= m_1 + 1, \cdots, m_2, \\
&\vdots & &\vdots \\
&= \sigma_r^2, & j &= m_r + 1, \cdots, n.
\end{aligned} \tag{50}
$$

Suppose, $c_1, c_2 . \cdots, c_{n-1}$ *are a sequence of constants defined as*

$$
\begin{aligned}
c_1 &= (n - r)^{-1}, \\
c_2 &= (n - r)^{-1}[2 + W_2(n - r, 0, 1)], \\
c_3 &= (n - r)^{-1}[3 + W_2(n - r, 0, 2) + W_3(n - r, 0, 1)], \\
&\vdots \\
c_{n-r} &= (n - r)^{-1}[(n - r) + W_2(n - r, 0, n - r - 1) \\
&\quad + \cdots + W_r(n - r, 0, n - 2r + 1)], \\
c_{n-r+1} &= (n - r)^{-1}[(n - r) + W_2(n - r, 0, n - r) \\
&\quad + \cdots + W_r(n - r, 0, n - 2r + 2)], \\
c_{n-r+2} &= (n - r)^{-1}[2(n - r) + W_3(n - r, 0, n - r) \\
&\quad + \cdots + W_r(n - r, 0, n - 2r + 3)], \\
&\vdots \\
c_{n-1} &= (n - r)^{-1}[(n - r)(r - 1) + W_r(n - r, 0, n - r)].
\end{aligned} \tag{51}
$$

Let

$$\mu^* = \mu \sum_{j=1}^{m_1-1} c_j + \mu_1 \sum_{j=m_1}^{m_2-1} c_j + \cdots + \mu_r \sum_{j=m_r}^{n-1} c_j, \tag{52}$$

$$\sigma^{2*} = \sigma^2 \sum_{j=1}^{m_1-1} c_j^2 + \sigma_1^2 \sum_{j=m_1}^{m_2-1} c_j^2 + \cdots + \sigma_r^2 \sum_{j=m_r}^{n-1} c_j^2 \tag{53}$$

and $E|U_i - E(U_i)|^3 \leqq \beta < \infty$ *for all* i. *Then* $(T - \mu^*)\sigma^{*-1}$ *is asymptotically normal with mean* 0 *and variance* 1.

PROOF. From (41) and (51) we observe that

$$T = \sum_{j=1}^{n-1} c_j U_{j+1}.$$

Let $y_j = c_j U_{j+1}, j = 1, 2, \cdots, n-1$. Then, the sequence $\{y_j\}$ are independently distributed. Let $\sigma^2_{\min} = \min(\sigma^2, \sigma_1^2, \cdots, \sigma_r^2)$. Then

$$\begin{aligned} \alpha^2 &= \textstyle\sum_{j=1}^{n-1} V(y_j) \geqq \sigma^2_{\min} \sum_{j=1}^{n-1} c_j^2 \\ &\geqq \sigma^2_{\min} \textstyle\sum_{j=1}^{n-r} j^2 (n-r)^{-2} \qquad (54) \\ &= \sigma^2_{\min}(n-r)^{-1}(n-r+1)(2n-2r+1)/6. \end{aligned}$$

Also,

$$\rho^3 = \textstyle\sum_{j=1}^{n-1} E|y_j - E(y_j)|^3 = \sum_{j=1}^{n-1} c_j^3 E|U_{j+1} - E(U_{j+1})|^3 \leqq \beta \sum_{j=1}^{n-1} c_j^3.$$

By Lemma 3.4 the sequence $\{c_j\}$ defined in (51) is uniformly bounded by r. Therefore,

$$\rho^3 \leqq \beta r^3(n-1). \qquad (55)$$

From (54) and (55), we obtain $\rho/\alpha \leqq O(n^{-1/6})$. Hence, $\lim_{n\to\infty} \rho/\alpha = 0$. Thus, the sequence $\{y_i\}$ satisfies Liapounoff's condition ([3], P. 275), and the proof of the theorem is complete.

The following two corollaries are immediate:

Corollary 3.1. *Under H_0, $[T - E(T)]\sigma^{-1}[\sum_{j=1}^{n-1} c_j^2]^{-\frac{1}{2}}$ is asymptotically normal with mean 0 and variance 1 where σ^2 is the common value of $\sigma_1^2, \sigma_2^2, \ldots \sigma_r^2$, given in equation (50) and $E(T)$ is given in equation (48).*

Corollary 3.2. *If $\delta > 0$, the asymptotic power function of a test of size ϵ is given by*

$$P(m_1, m_2, \cdots, m_r, \delta) = 1 - \Phi(\mu(\textstyle\sum_{j=1}^{n-1} c_j)\sigma^{*-1} - \mu^* \sigma^{*-1} + \sigma(\sum_{j=1}^{n-1} c_j^2)^{\frac{1}{2}} t_\epsilon \sigma^{*-1})$$

where

$$\Phi(t_\epsilon) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{t_\epsilon} e^{-x^2/2}\,dx = 1 - \epsilon.$$

We now investigate the rate of convergence of the distribution of the statistic T to the normal distribution in the following theorem.

Theorem 3.3. *Let $F_n(z) = P\{(T - \mu^*)\sigma^{*-1} \leqq z\}$. Then $|F_n(z) - \Phi(z)| = O(n^{-\frac{1}{2}})$, where $\Phi(z)$ is defined in equation (29).*

Proof. The proof of the theorem can be verified from a result given in Loève ([3], P. 288).

Finally, we consider an example. Let $f(x, \theta)$ be the density of the normal distribution with mean θ and variance 1. Then, we can write

$$f(x, \theta) = h(x) \exp[\Psi_1(\theta) U(x) + \Psi_2(\theta)]$$

where,

$$h(x) = (2\pi)^{-\frac{1}{2}} e^{-x^2/2}; \qquad \Psi_1(\theta) = \theta; \qquad \Psi_2(\theta) = -\tfrac{1}{2}\theta^2; \qquad U(x) = x.$$

Hence, $T = \sum_{j=1}^{n-1} c_j X_{j+1}$.

In this case, the distribution of T is normal for any sample size. Further, it follows from (52) and (53) that,

$$\mu^* = \theta \textstyle\sum_{j=1}^{n-1} c_j + \Delta,$$

where

$$\Delta = \delta\{\sum_{j=m_1}^{m_2-1} c_j + 2 \sum_{j=m_2}^{m_3-1} c_j + \cdots + r \sum_{j=m_r}^{n-1} c_j\}$$

and

$$\sigma^{*2} = \sum_{j=1}^{n-1} c_j^2.$$

It may be pointed out that if we substitute $r = 1$ in (41), we obtain the test-statistic suggested by Kander and Zacks [2] for detecting one change.

Acknowledgment. The author likes to thank professor Innis G. Abrahamson for many useful discussions he had with her throughout the writing of this dissertation. He is indebted to the late professor E. J. Gumbel for his help and encouragement.

REFERENCES

[1] 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.

[2] Kander, Z. and Zacks, S. (1966). Test procedures for possible changes in parameters occurring at unknown time points. *Ann. Math. Statist.* **37** 1196–1210.

[3] Loève, M. (1960). *Probability Theory.* Van Nostrand, New York.

[4] Mustafi, Chandan K. (1967). The inverse of a certain matrix, with an application. *Ann. Math. Statist.* **38** 1289–1292.

[5] Page, E. (1955). A test for a change in a parameter occurring at an unknown point. *Biometrika* **42** 523–526.

[6] Tainiter, M. (1966). Sequential hypothesis tests for r-dependent marginally stationary processes. *Ann. Math. Statist.* **37** 90–97.