The Annals of Mathematical Statistics
1969, Vol. 40, No. 1, 116–126

ON DETECTING CHANGES IN THE MEAN OF NORMAL VARIATES[1]

BY L. A. GARDNER, JR.

Sperry Rand Research Center

A test statistic for the hypothesis of no change in the mean within the first n observations against general alternatives, having prescribed prior distribution, is derived. In the special case of a uniform temporal distribution of at most one change (hereafter abbreviated by AMOC), we prove that the statistic has the same limiting distribution as Smirnov's ω_n^2. Critical values for finite n, obtained by numerical integration, are presented.

1. Introduction and summary. The problem considered is that of detecting changes in the mean of independent unit variance normal random variables when the times of change are assigned an *a priori* distribution. Two situations are considered: The unknown amounts of change are (A) arbitrary, or (B) successively plus and minus the same unknown quantity. Model B is appropriate in certain problems involving angular tracking of an evading target.

In calculating the likelihood of the observations, conditioned on an arbitrary sequence of change indices, we assign the nuisance parameters, viz. the initial mean level and the amount(s) of change, normal probability distributions. Their respective variances are then allowed to approach infinity and zero at appropriate points in the argument. Let $x_1, x_2, \cdots, x_n$ be the first n observations, and let ω_i equal 1 or 0 according to whether there is or is not a change in the mean between x_i and x_{i+1}. In the indicated fashion we find (Sections 2, 3) that the log-likelihood ratios are, up to additive constants,

$$\text{(1.1A)} \qquad \Lambda_{\boldsymbol{\omega}}(x_1, \cdots, x_n) = \sum_{j=1}^{n-1} \omega_j \left[\sum_{i=j}^{n-1} (x_{i+1} - \bar{x}_n)\right]^2$$

$$\text{(1.1B)} \qquad = \left[\sum_{j=1}^{n-1} \left(\sum_{i=1}^{j} (-1)^{\Omega_i} \omega_i\right)(x_{j+1} - \bar{x}_n)\right]^2$$

wherein

$$\Omega_i = 1 + \omega_1 + \cdots + \omega_i,$$

and $\bar{x}_n$ is the arithmetic mean of the first n observations. These ratios are for testing no change against a specified sequence $\boldsymbol{\omega} = (\omega_1, \cdots, \omega_{n-1})$ of change times under Models A and B, respectively. As required, the statistics are translation invariant. A test of the hypothesis of change in the mean at no point, against a set of alternatives $\{\boldsymbol{\omega}\}$ having assigned nonzero prior probabilities, rejects the hypothesis for large values of

$$\text{(1.2)} \qquad Q_n = \sum_{\{\boldsymbol{\omega}\}} p(\boldsymbol{\omega}) \Lambda_{\boldsymbol{\omega}}(x_1, \cdots, x_n).$$

Received 8 January 1968.

[1] This research was supported by the U. S. Naval Ordnance Laboratory under Contract No. N60921-68-C-0020.

Under the null hypothesis, Q_n is a quadratic form in n independent standardized normal variates. As such, its distribution is a mixture of chi-square distributions (cf. Robbins and Pitman [7]) which, except in the case of AMOC, is too complicated to handle. (The distribution theory for the Bayesian test based on the weighted sum of likelihood ratios is even more difficult and hence not considered.)

Over and above derivation of (1.1), we do in fact restrict attention to AMOC. Here we can get useable results. Suppose the change occurs, if it does, at the index k with prior probability p_k. Then (1.1) and (1.2) combine and reduce to the single statistic

$$Q_n = \sum_{k=1}^{n-1} p_k [\sum_{j=k}^{n-1} (x_{j+1} - \bar{x}_n)]^2, \tag{1.3}$$

because for AMOC Model A and Model B are identical.

Before continuing, we point out that our technique for deriving (1.1) follows that of Chernoff and Zacks [2]. (See also [5]. A different approach to the problem is taken in [1], [8] and [9]). They assume, however, that the sign of the change is known beforehand, which is a presumption we do not wish to make. This makes a world of difference from the standpoint of distribution theory. If we knew the change would result in (say) a positive increase in the mean value, then the appropriate statistic is (1.3) with the square deleted. For a uniform prior distribution it then agrees with the statistic given in Equation (8.9) of [2]. Without the square in (1.3) the distribution theory is clearly trival. On the other hand, calculation of the distribution of (1.3) as written is far from being a straightforward matter, even under the null hypothesis.

Let

$$Q_n = \boldsymbol{\varepsilon}' \mathbf{Q} \boldsymbol{\varepsilon}$$

be a non-negative quadratic form in independent normal variates $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$, each with zero mean and unit variance. The distribution of Q_n is clearly invariant under the rotation which diagonalizes the symmetric matrix $\mathbf{Q}$. Thus

$$Q_n =_D \sum_{k=1}^{r} \beta_k \epsilon_k^2 \tag{1.4}$$

($X =_D Y$ means X and Y have the same distribution) where $\beta_1, \cdots, \beta_r$ are the strictly positive eigenvalues of $\mathbf{Q}$ and r is its rank, both of which generally depend on n (cf. Kendall and Stuart [6], Section 15.11). In particular, the matrix of (1.3) is

$$\mathbf{Q} = n^{-2} \sum_{k=1}^{n-1} \mathbf{a}_k \qquad \mathbf{a}_k{}' \mathbf{a}_k = p_k^{\frac{1}{2}} \begin{bmatrix} n-k \\ \vdots \\ n-k \\ -k \\ \vdots \\ -k \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix} a \\ a \\ a \end{matrix}}\right\} k \\ \left.\vphantom{\begin{matrix} a \\ a \\ a \end{matrix}}\right\} n-k \end{matrix} \tag{1.5}$$

where we may assume without loss of generality that no p_k vanishes. It is shown (Section 4) that

$$\beta_k = 1/\gamma_k \tag{1.6}$$

where $\gamma_1, \cdots, \gamma_{n-1}$ are the eigenvalues of a certain symmetric positive definite tridiagonal matrix. If $\varphi_k(\lambda)$ denotes $p_1 p_2 \cdots p_k$ times the characteristic polynomial of the leading k by k submatrix of said matrix, then we find a second order difference equation

$$\Delta^2 \varphi_k - (\lambda p_{k+2} - 4)\varphi_{k+1} = 0, \qquad \varphi_0 = 1, \qquad \varphi_1 = \lambda p_1 - 2, \tag{1.7}$$

where Δ is the first forward difference operator. The desired γ's in (1.6) are the roots of $\varphi_{n-1}(\lambda) = 0$.

Although (1.7) has normal Sturm Liouville form, and hence inequalities on its eigenvalues obtainable (cf. Fort [3], Chapter X), we cannot write down an explicit formula for the γ's except when p_k is independent of k (uniform prior distribution). Taking $p_k = 1/n$, rather than $1/(n-1)$, we readily obtain (Section 5)

$$\gamma_k = 4n \cos^2 k\pi/2n \qquad (k = 1, 2, \cdots, n-1). \tag{1.8}$$

We now scale (1.3) and introduce the sequence of positive random variables

$$Y_n = 6n(n^2 - 1)^{-1} Q_n. \tag{1.9}$$

This normalization makes $\mathcal{E} Y_n = 1$ for all n, as can be seen by direct evaluation of $\mathcal{E} Q_n = \text{tr } \mathbf{Q}$ with $\mathbf{Q}$ given by (1.5). Combining (1.4), with $r = n - 1$, (1.6) and (1.8) we get

$$Y_n =_D 6n^2[\pi^2(n^2 - 1)]^{-1} \sum_{k=1}^{n-1} [(k\pi/2n)^{-1} \cos (k\pi/2n)]^{-2} k^{-2} \epsilon_k^2, \tag{1.10}$$

after dividing and multiplying the summand by $k^2\pi^2/n^2$. Since $\epsilon_1^2, \cdots, \epsilon_n^2$ are independent χ_1^2 variates, the νth cumulant of (1.10) is

$$\begin{aligned} \varkappa_\nu(Y_n) &= 2^{\nu-1}(\nu - 1)!\,(6n^2[\pi^2(n^2 - 1)]^{-1})^\nu \\ &\quad \cdot \sum_{k=1}^{n-1} [(k\pi/2n)^{-1} \cos (k\pi/2n)]^{-2\nu} k^{-2\nu} \\ &= \tfrac{1}{2}(\nu - 1)!\,(3(n^2 - 1)^{-1})^\nu \sum_{k=1}^{n-1} (\cos k\pi/2n)^{-2\nu}, \end{aligned} \tag{1.11}$$

with $\varkappa_1(Y_n) \equiv 1$. When $n = 2$, (1.11) reduces to $2^{\nu-1}(\nu - 1)!$, so

$$Y_2 =_D \chi_1^2.$$

We have not been able to find a closed form expression for the sum appearing in the first line of (1.11) for general ν and n. We can, however, prove that it converges to $\zeta(2\nu)$ as $n \to \infty$ for every fixed $\nu \geqq 1$, where ζ is the Riemann Zeta function (i.e., the limiting value is the same as that obtained when we set the multiplier of $1/k^{2\nu}$ to unity). It follows that under the null hypothesis,

$$Y_n = 6(n^2 - 1)^{-1} \sum_{k=1}^{n-1} [\sum_{j=k}^{n-1} (x_{j+1} - \bar{x}_n)]^2 \tag{1.12}$$

converges in distribution to a random variable Y uniquely determined by the cumulants

$$\varkappa_\nu(Y) = 2^{\nu-1}(\nu - 1)!\,(6\pi^{-2})^\nu \zeta(2\nu). \tag{1.13}$$

Stated in other terms

$$Y =_D 6\pi^{-2} \sum_{k=1}^{\infty} k^{-2} \epsilon_k^2. \tag{1.14}$$

Fortunately, the distribution of (1.14) is tabulated. In fact, it is precisely that of the limiting distribution of Smirnov's ω_n^2-criterion, normalized to have mean one (see von Mises [10], Chapter IX, Section 7). This statistic, it will be recalled, is used to test whether or not observations $z_1, \cdots, z_n$ are drawn from a population with a prescribed continuous cumulative distribution, say G. The formula is

$$\omega_n^2 = (2n)^{-1} + 6 \sum_{k=1}^{n} [G(z_k) - (2k-1)/2n]^2. \tag{1.15}$$

It is somewhat remarkable that the limiting distribution of Y_n is identical to that of ω_n^2. Letting $f_n(y)$ denote the density of (1.12) and

$$\int_{c_n(\alpha)}^{\infty} f_n(y)\, dy = \alpha,$$

we can thus obtain $c_2(\alpha)$ and $c_\infty(\alpha)$ from tables of the percentage points of the χ_1^2 distribution and Table VIII of [10], respectively.

We are left with the problem of finding critical values, c_n, for finite values of n. Using (1.11), the cumulant generating function of (1.12) is

$$\Psi_n(t) = \sum_{\nu \geq 1} \kappa_\nu(Y_n)(it)^\nu(\nu!)^{-1} = -\tfrac{1}{2} \sum_{k=1}^{n-1} \log\,(1 - it\alpha_k^{-1}) \tag{1.16}$$

wherein

$$\alpha_k = \tfrac{1}{3}(n^2 - 1) \cos^2 k\pi/2n \tag{1.17}$$

depends on n. The density function

$$f_n(y) = (2\pi)^{-1} \int_{-\infty}^{\infty} e^{-ity + \Psi_n(t)}\, dt$$

written in its real form is

$$f_n(y) = \pi^{-1} \int_0^\infty \prod_{k=1}^{n-1} (1 + t^2\alpha_k^{-2})^{-\frac{1}{4}} \cos\,(ty - \tfrac{1}{2} \sum_{k=1}^{n-1} \tan^{-1} t\alpha_k^{-1})\, dt. \tag{1.18}$$

The density (1.18) was computed by numerical integration for selected values of n in the range $3 \leqq n \leqq 20$.[2] The tails of $-\log f_n$ were then fit by a quartic at y-points so chosen to give an oscillatory constant peak amplitude error curve in approximating f_n. The results were integrated to yield the corresponding critical values. Figure 1 is a curve drawn between these points, and suffices for purposes of application. We note the approach to limiting values is extremely rapid.

Figure 2 was obtained by drawing smooth curves through observed rejection frequencies in 250 Monte Carlo runs. δ is the amount of change taking place between the kth and $(k+1)$st observation.

2. The distribution of the observations given the points of change. The observations are

$$x_i = \mu_i + \epsilon_i, \qquad (i = 1, 2, \cdots, n), \tag{2.1}$$

[2] The author wishes to express thanks to Mr. Robert Church of Sperry Rand Research Center for doing the numerical work. Graphs of these functions are available upon request.

or in (column) vector notation

$$\mathbf{x} = \boldsymbol{\mu} + \boldsymbol{\varepsilon}$$

where the ϵ's are independently and identically distributed as a $\mathfrak{N}(0, 1)$ variate. Let $\omega_1, \cdots, \omega_{n-1}$ be the indicators defined in Section 1. Under Model A, let

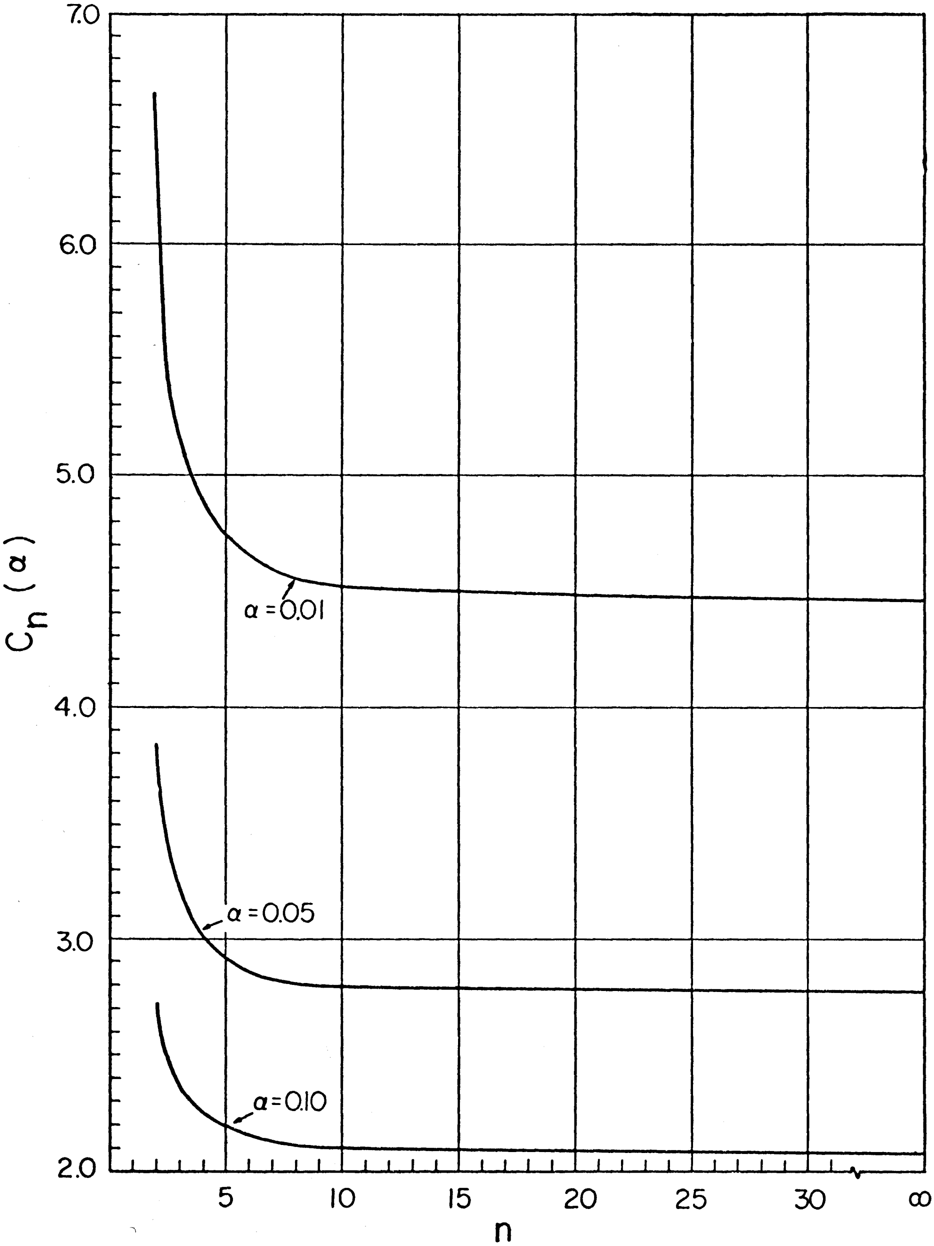

FIG. 1. Critical values of the statistic (1.12)

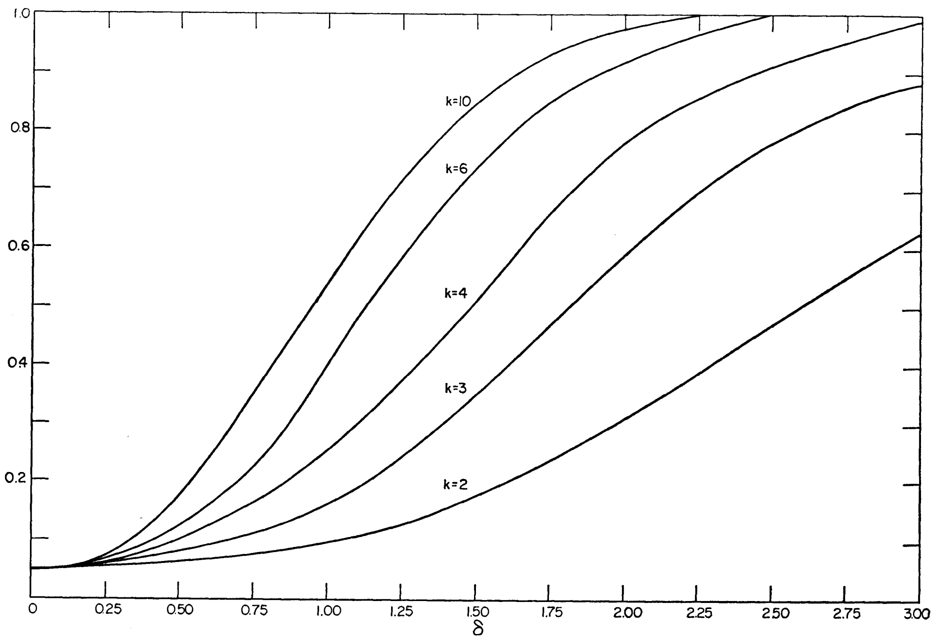

FIG. 2. The power function of the test using the statistic (1.12) for $n = 20$ and $\alpha = .05$

$\delta_1, \cdots, \delta_{n-1}$ be the corresponding amounts of change. Then for $i = 1, 2, \cdots, n - 1$,

$$\text{(2.2A)} \qquad \begin{aligned} \mu_{i+1} &= \mu_i + \omega_i \delta_i \\ &= \mu + \textstyle\sum_{k=1}^{i} \omega_k \delta_k \end{aligned}$$

wherein $\mu = \mu_1$. For Model B, the equation for the means becomes

$$\text{(2.2B)} \qquad \mu_{i+1} = \mu + \delta \textstyle\sum_{k=1}^{i} \omega_k (-1)^{\Omega_k}$$

Here the levels are successively μ, $\mu + \delta$, μ, $\mu + \delta$, $\cdots$ with $\delta \gtrless 0$. We assign (fictitious) normal distributions

$$\text{(2.3)} \qquad \mathcal{L}(\mu) = \mathfrak{N}(0, \tau^2) \qquad \mathcal{L}(\delta_i \text{ or } \delta) = \mathfrak{N}(0, \sigma^2)$$

and assume the ϵ's, μ and δ's or δ are mutually independent.

Until otherwise stated, calculations will be based on the indicators $\omega = (\omega_1, \cdots, \omega_{n-1})$ being held fixed. Then, conditioned also on the initial mean level μ, we have from (2.1)–(2.3) $\mathcal{E}\mathbf{x} = \mu\mathbf{e}$, where $\mathbf{e}$ is the n-vector of all 1's. This is true for either model. We have

$$\begin{aligned} x_1 - \mu &= \epsilon_1 \\ x_i - \mu &= \epsilon_i + \begin{cases} \sum_{k=1}^{i-1} \omega_k \delta_k & \text{(A)} \\ \delta \sum_{k=1}^{i-1} \omega_k (-1)^{\Omega_k} & \text{(B)} \end{cases} \qquad (i = 2, 3, \cdots, n). \end{aligned}$$

In both cases $\mathcal{E}(x_1 - \mu)^2 = 1$ and $\mathcal{E}(x_i - \mu)(x_1 - \mu) = 0$ for $i \geqq 2$. Furthermore, for $i, j = 2, 3, \cdots, n$,

$$\mathcal{E}(x_i - \mu)(x_j - \mu) = \delta_{ij} + \begin{cases} \sigma^2 \sum_{k=1}^{\min(i-1,j-1)} \omega_k & \text{(A)} \\ \sigma^2 \sum_{k=1}^{i-1} \omega_k(-1)^{\Omega_k} \sum_{k=1}^{j-1} \omega_k(-1)^{\Omega_k} & \text{(B)} \end{cases}$$

wherein δ_{ij} is Kronecker's delta. We thus obtain

$$\mathcal{L}(\mathbf{x} \mid \omega, \mu) = \mathfrak{N}(\mu \mathbf{e}, \mathbf{I} + \sigma^2 \mathbf{B}) \tag{2.4}$$

where

$$\mathbf{B} = \begin{cases} \mathbf{CC}' & \text{(A)} \\ \mathbf{cc}' & \text{(B)} \end{cases} \tag{2.5}$$

and

$$\mathbf{C} = \begin{bmatrix} 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & \omega_1 & & & & \\ \cdot & \omega_1 & \omega_2 & & 0 & \\ \cdot & \omega_1 & \omega_2 & \omega_3 \cdot & & \\ \cdot & \cdot & \cdot & \cdot & \cdot\, \omega_{n-2} & \\ 0 & \omega_1 & \omega_2 & \omega_3 & \omega_{n-2} & \omega_{n-1} \end{bmatrix} \tag{2.6A}$$

$$\mathbf{c} = \begin{bmatrix} 0 \\ c_2 \\ c_3 \\ \vdots \\ c_{n-1} \end{bmatrix} \qquad c_i = \sum_{k=1}^{i-1} \omega_k(-1)^{\Omega_k}. \tag{2.6B}$$

We now integrate out the unknown μ. Let $\varphi = \varphi(\mu \mid 0, \tau^2)$ be the normal density function, with mean 0 and variance τ^2, evaluated at μ. Then from (2.4), with all integrations over the whole real axis,

$$\mathcal{E}\mathbf{x} = \int \mathcal{E}(\mathbf{x} \mid \mu)\varphi \, d\mu = \mathbf{e} \int \mu\varphi \, d\mu = \mathbf{0},$$

and

$$\begin{aligned} \mathcal{E}(\mathbf{xx}') &= \int \mathcal{E}(\mathbf{xx}' \mid \mu)\varphi \, d\mu \\ &= \int [\mathcal{E}(\mathbf{x} - \mu\mathbf{e})(\mathbf{x} - \mu\mathbf{e})' + \mu\mathbf{ee}'\mu]\varphi \, d\mu \\ &= \mathbf{I} + \sigma^2 \mathbf{B} + \tau^2 \mathbf{ee}'. \end{aligned}$$

There results

$$\mathcal{L}(\mathbf{x} \mid \omega) = \mathfrak{N}(\mathbf{0}, \mathbf{\Sigma}) \tag{2.7}$$

$$\mathbf{\Sigma} = \mathbf{\Sigma}_0 + \sigma^2 \mathbf{B} \qquad \mathbf{\Sigma}_0 = \mathbf{I} + \tau^2 \mathbf{ee}'.$$

The dependence on ω is via $\mathbf{B}$ given by (2.5) and either (2.6A) or (2.6B). The zero subscript pertains to the no change case when $\mathbf{B}$ is the zero matrix.

3. The likelihood ratio for a specified alternative. The statistic used to test the hypothesis of no change, i.e., $\omega = \mathbf{0}$, against a specified alternative set of change indices, $\boldsymbol{\omega}$, is obtained as a monotone function of the likelihood ratio

$$\lambda = L(\mathbf{x} \mid \boldsymbol{\omega})/L(\mathbf{x} \mid \mathbf{0}) \propto e^{\frac{1}{2}\mathbf{x}'(\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}^{-1})\mathbf{x}}, \tag{3.1}$$

after letting $\tau^2 \to \infty$ and expanding to first order in σ^2 as $\sigma^2 \to 0$. The positive constant of proportionality in (3.1) depends on $\boldsymbol{\omega}$ but not on the observations. (If we were to consider the weighted sum of likelihood ratios, rather than their logarithms, then this constant would have to be retained.)

We carry through the calculation for Model A. The result for Model B follows as a special case. Under either model, Woodbury's formula (Householder [4], p. 124) gives

$$\boldsymbol{\Sigma}_0^{-1} = \mathbf{I} - (n + \tau^{-2})^{-1}\mathbf{ee}' \to \mathbf{I} - n^{-1}\mathbf{ee}' \tag{3.2}$$

as $\tau^2 \to \infty$. With $\mathbf{B} = \mathbf{CC}'$, (2.7) and the same formula yield

$$\boldsymbol{\Sigma}^{-1} = \boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_0^{-1}\mathbf{CMC}'\boldsymbol{\Sigma}_0^{-1} \tag{3.3}$$

where

$$\mathbf{M} = (\mathbf{C}'\boldsymbol{\Sigma}_0^{-1}\mathbf{C} + \sigma^{-2}\mathbf{I})^{-1}.$$

Setting $\mathbf{A} = -\mathbf{C}'\boldsymbol{\Sigma}_0^{-1}\mathbf{C}$, and assuming $\mathbf{A}$'s largest eigenvalue in absolute value is less than unity, we have the convergent series

$$\mathbf{M} = \sigma^2(\mathbf{I} - \sigma^2\mathbf{A})^{-1} = \sigma^2(\mathbf{I} + \sigma^2\mathbf{A} + \sigma^4\mathbf{A}^2 + \cdots).$$

Substituting this and the limiting expression (3.2) into (3.3), we obtain

$$\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}^{-1} = \sigma^2(\mathbf{I} - n^{-1}\mathbf{ee}')\mathbf{CC}'(\mathbf{I} - n^{-1}\mathbf{ee}') + 0(\sigma^2\mathbf{ee}'),$$

as $\sigma^2 \to 0$. The rejection region corresponds to large values of λ in (3.1), or equivalently large values of twice the exponent,

$$\Lambda = \|\mathbf{C}'(\mathbf{I} - n^{-1}\mathbf{ee}')\mathbf{x}\|^2 \tag{3.4}$$

after dropping the σ^2. This result holds true under Model B if we replace $\mathbf{C}$ by the vector $\mathbf{c}$. (The expansion step is not needed because $\mathbf{M}$ is then a scalar.) After substituting (2.6A) and (2.6B) into (3.4), we see that (1.1A) and (1.1B) obtain.

4. The AMOC case. We now restrict attention to $\mathbf{Q}_n$ as defined in (1.3). It is not difficult to see that the matrix of this quadratic form is indeed (1.5), which we write in the form

$$n^2\mathbf{Q} = \mathbf{AA}' \qquad \mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_{n-1}].$$

To find the eigenvalues of $\mathbf{Q}$, we introduce another n by $n - 1$ matrix

$$\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_{n-1}] \qquad \mathbf{b}_k = p_k^{-\frac{1}{2}}\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ -1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}\begin{matrix} \left.\vphantom{\begin{matrix}0\\ \vdots\\ 0\\ 1\end{matrix}}\right\} k \\ \left.\vphantom{\begin{matrix}-1\\ 0\\ \vdots\\ 0\end{matrix}}\right\} n-k \end{matrix}$$

where of course we are assuming $p_1 p_2 \cdots p_{n-1} \neq 0$. From the definition of $\mathbf{a}_k$ in (1.5) we see that $\mathbf{A}'\mathbf{B} = n\mathbf{I}$, where $\mathbf{I}$ is here the $(n-1)$-dimensional identity. Thus

$$\mathbf{B}'\mathbf{Q}\mathbf{B} = \mathbf{I}. \tag{4.1}$$

We easily find that $\mathbf{B}'\mathbf{B}$ is a symmetric positive definite tridiagonal matrix:

$$\mathbf{B}'\mathbf{B} = \begin{bmatrix} 2p_1^{-1} & -(p_1 p_2)^{-\frac{1}{2}} & & & & \\ -(p_1 p_2)^{-\frac{1}{2}} & 2p_2^{-1} & -(p_2 p_3)^{-\frac{1}{2}} & & \mathbf{0} & \\ & -(p_2 p_3)^{-\frac{1}{2}} & 2p_3^{-1} & & & \\ & & & \ddots & & \\ & \mathbf{0} & & & 2p_{n-2}^{-1} & -(p_{n-2}\, p_{n-1})^{-\frac{1}{2}} \\ & & & & -(p_{n-2}\, p_{n-1})^{-\frac{1}{2}} & 2p_{n-1}^{-1} \end{bmatrix} \tag{4.2}$$

This can be diagonalized by a pure rotation, i.e.,

$$\mathbf{B}'\mathbf{B} = \mathbf{R}\boldsymbol{\Gamma}\mathbf{R}' \qquad \mathbf{R}'\mathbf{R} = \mathbf{I}$$

where $\boldsymbol{\Gamma} = \operatorname{diag}[\gamma_1, \gamma_2, \cdots, \gamma_{n-1}]$. Letting

$$\mathbf{C} = \mathbf{B}\mathbf{R}\boldsymbol{\Gamma}^{-\frac{1}{2}},$$

which is n by $n-1$, this is the same thing as

$$\mathbf{C}'\mathbf{C} = \mathbf{I}. \tag{4.3}$$

We return to (4.1) and post (resp. pre) multiply both sides by $\mathbf{R}\boldsymbol{\Gamma}^{-\frac{1}{2}}$ (resp. $\boldsymbol{\Gamma}^{-\frac{1}{2}}\mathbf{R}'$) to obtain

$$\mathbf{C}'\mathbf{Q}\mathbf{C} = \boldsymbol{\Gamma}^{-1}. \tag{4.4}$$

Now we note that the sum of the components of $\mathbf{a}_k$, given in (1.5), as well as of $\mathbf{b}_k$ is zero for any $k = 1, 2, \cdots, n-1$. Consequently, if we introduce the unit n-vector

$$\mathbf{u} = \begin{bmatrix} n^{-\frac{1}{2}} \\ n^{-\frac{1}{2}} \\ \vdots \\ n^{-\frac{1}{2}} \end{bmatrix},$$

we have $\mathbf{u}'\mathbf{A} = \mathbf{0}$ and hence $\mathbf{u}'\mathbf{Q} = \mathbf{0}$. Furthermore, $\mathbf{u}'\mathbf{B} = \mathbf{0}$ and hence $\mathbf{u}'\mathbf{C} = \mathbf{0}$ by definition of $\mathbf{C}$. Defining the n by n matrix

$$\mathbf{U} = [\mathbf{C}, \mathbf{u}],$$

(4.4) and (4.3) therefore respectively yield

$$\mathbf{U}'\mathbf{Q}\mathbf{U} = \begin{bmatrix} \mathbf{C}'\mathbf{Q}\mathbf{C} & \mathbf{C}'\mathbf{Q}\mathbf{u} \\ \mathbf{u}'\mathbf{Q}\mathbf{C} & \mathbf{u}'\mathbf{Q}\mathbf{u} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Gamma}^{-1} & 0 \\ 0 & 0 \end{bmatrix},$$

and

$$\mathbf{U}'\mathbf{U} = \begin{bmatrix} \mathbf{C}'\mathbf{C} & \mathbf{C}'\mathbf{u} \\ \mathbf{u}'\mathbf{C} & \mathbf{u}'\mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & 0 \\ 0 & 1 \end{bmatrix}.$$

These two equations prove that the $n - 1$ nonzero eigenvalues of $\mathbf{Q}$ are the reciprocals of those of (4.2).

It is not difficult to see that $p_1 p_2 \cdots p_k$ times the characteristic polynomial of the leading k by k submatrix of (4.2) indeed satisfies the recurrence relation (1.7) (see Householder [4], p. 174).

5. Uniform prior distribution. When p_k is independent of k, the eigenvalues of (4.2) are easy to compute. (Indeed, they are proportional to those of a matrix consisting of 2's on the main diagonal and -1's on the first upper and lower diagonals.) Taking (say) $p_k = 1/n$ in (1.7), the difference equation is simply

$$\varphi_{k+2} - ((\lambda/n) - 2)\varphi_{k+1} + \varphi_k = 0, \qquad \varphi_0 = 1, \qquad \varphi_1 = (\lambda/n) - 2.$$

This is recognized as the recurrence relation which defines the Chebyshev polynomials of the second kind in the variable

$$\cos\theta = \lambda(2n)^{-1} - 1, \tag{5.1}$$

viz.

$$\varphi_k(\lambda) = \sin(k+1)\theta/\sin\theta \qquad (k \geqq 0). \tag{5.2}$$

The roots of $\varphi_{n-1} = 0$ are $\theta_k = k\pi/n$ $(k = 1, 2, \cdots, n-1)$. This combines with the transformation (5.1) and $\cos\theta = 2\cos^2\frac{1}{2}\theta - 1$ to establish (1.8), and consequently (1.10) and (1.11).

It remains to prove that as $n \to \infty$

$$S_{\nu,n} = \textstyle\sum_{k=1}^{n-1}[(k\pi/2n)^{-1}\cos(k\pi/2n)]^{-2\nu}k^{-2\nu} \to \zeta(2\nu) \tag{5.3}$$

as claimed, and thus establish (1.13). The value of the sum is unaltered if we replace the cosine by the sine. Thus, setting $\delta_n = \pi/2n$,

$$S_{\nu,n} = \textstyle\sum_{k=1}^{m-1}(k\delta_n/\sin k\delta_n)^{2\nu}k^{-2\nu} + \sum_{k=m}^{n-1}(k\delta_n/\sin k\delta_n)^{2\nu}k^{-2\nu}, \tag{5.4}$$

where the integers $m = m(n)$ are at our disposal. We choose them going to infinity with n in such a way that

$$m/n \to 0 \qquad m^2/n \to \infty.$$

Since

$$1 < k\delta_n/\sin k\delta_n < \tfrac{1}{2}\pi$$

for all $1 \leqq k \leqq n - 1$ and all n, the second sum in (5.4) is bounded above by

$$(n - m)m^{-2\nu}(\pi/2)^{2\nu} \leqq \text{const.}\ nm^{-2} = 0(1)$$

for every fixed $\nu \geqq 1$. The value of the first sum in (5.4) always falls between

$$(\delta_n/\sin\delta_n)^{2\nu}\textstyle\sum_{k=1}^{m-1}k^{-2\nu} \quad \text{and} \quad (m\delta_n/\sin m\delta_n)^{2\nu}\sum_{k=1}^{m-1}k^{-2\nu}$$

Since both δ_n and $m\delta_n$ tend to 0 as $n \to \infty$, the quantities multiplying the common sum both approach unity. In other words, both bounds approach $\zeta(2\nu)$, which proves (5.3).

REFERENCES

[1] Bather, J. A. (1967). On a Quickest Detection Problem. *Ann. Math. Statist.* **38** 711–724.

[2] 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.

[3] Fort, T. (1948). *Finite Differences and Difference Equations in the Real Domain.* Oxford Univ. Press, London.

[4] Householder, A. S. (1964). *The Theory of Matrices in Numerical Analysis.* Blaisdell (Ginn and Co.), N. Y.

[5] 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.

[6] Kendall, M. G. and Stuart, A. (1958). *The Advanced Theory of Statistics. Vol. 1.* Hafner, N. Y.

[7] Robbins, H. and Pitman, E. J. G. (1949). Application of the Method of Mixtures to Quadratic Forms in Normal Variates. *Ann. Math. Statist.* **20** 552–560.

[8] Shiryeav, A. N. (1963). On Optimum Methods in Quickest Detection Problems. *Theor. Probability Appl.* **8** 22–46.

[9] Shiryeav, A. N. (1965). Some Exact Formulas in a "Disorder" Problem. *Theor. Probability Appl.* **10** 348–354.

[10] von Mise's, R. (1964). *Mathematical Theory of Probability and Statistics.* Academic Press, N. Y.