SEQUENTIAL TESTS FOR NORMAL MARKOV SEQUENCE

R. M. PHATARFOD

(Received 14 August 1969)

Communicated by P. D. Finch

1. Introduction

This is a sequel to the author's (Phatarfod [9]) paper in which an analogue of Wald's Fundamental Identity (F.I.) for random variables defined on a Markov chain with a finite number of states was derived. From it the sampling properties of sequential tests of simple hypotheses about the parameters occurring in the transition probabilities were obtained. In this paper we consider the case of continuous Markovian variables. We restrict our attention to the practically important case of a Normal Markov sequence $X_0, X_1, X_2, \cdots$ such that

$$X_r - m = \rho(X_{r-1} - m) + Y_r, \quad (|\rho| < 1, \ r = 1, 2, 3, \cdots) \tag{1.1}$$

the Y_r being independent normal variables with mean zero and variance σ^2.

For such a sequence of observations, constructing an S.P.R.T. of the simple hypothesis $m = m_0$ against the simple alternative $m = m_1$ when σ^2 and ρ are known and deriving its sampling properties present no particular problems, since the logarithm of the likelihood ratio can be written as

$$\log \frac{p(X_0, X_1, X_2, \cdots, X_n | m_1)}{p(X_0, X_1, X_2, \cdots, X_n | m_0)} \sim Z_1 + Z_2 + Z_3 + \cdots + Z_n$$

where

$$Z_r = \frac{(m_0^2 - m_1^2)(1-\rho)^2}{2\sigma^2} + \frac{(m_1 - m_0)(1-\rho)}{\sigma^2}(X_r - \rho X_{r-1}).$$

The Z_r's being mutually independent random variables, Wald's theory of sequential analysis of independent observations holds. The test is carried out by plotting $\sum_{r=1}^{n} (X_r - \rho X_{r-1})$ against n. The same is true of a sequential test of $\sigma^2 = \sigma_0^2$ against $\sigma^2 = \sigma_1^2$ when m and ρ are known.

However, the test one would normally like to perform for the sequence X_0, $X_1, X_2, \cdots$ is about the parameter ρ. Consider, first, the case of the simple hypothesis $H_0 : \rho = \rho_0$ (usually $\rho_0 = 0$) against $H_1 : \rho = \rho_1$ ($\rho_1 > \rho_0$, say), with known values of m and σ^2, (here taken to be 0 and 1 respectively). We will assume that X_0 has the stationary distribution of the sequence. We then have the S.P.R.T. as follows: *Continue sampling while*

$$\log B < Z_0+Z_1+\cdots+Z_n < \log A, \tag{1.2}$$

where

$$\begin{aligned} Z_0 &= \tfrac{1}{2}\log\frac{1-\rho_1^2}{1-\rho_0^2}+\frac{(\rho_1^2-\rho_0^2)X_0^2}{2}, \\ Z_r &= (\rho_1-\rho_0)X_{r-1}X_r+\frac{\rho_0^2-\rho_1^2}{2}X_{r-1}^2, \end{aligned} \tag{1.3}$$

and accept H_0 or H_1 according as the left-hand or the right-hand inequality is the first not satisfied. Relation (1.2) reduces to

$$\begin{aligned} \frac{1}{\rho_1-\rho_0}\left[\log B+\tfrac{1}{2}\log\frac{1-\rho_0^2}{1-\rho_1^2}\right] &< \sum_{r=1}^{n} X_{r-1}X_r-\frac{\rho_0+\rho_1}{2}\sum_{2}^{n} X_{r-1}^2 \\ &< \frac{1}{\rho_1-\rho_0}\left[\log A+\tfrac{1}{2}\log\frac{1-\rho_0^2}{1-\rho_1^2}\right], \end{aligned} \tag{1.4}$$

and the test is carried out by plotting $\sum_{r=1}^{n} X_{r-1}X_r$ against $\sum_{r=2}^{n} X_{r-1}^2$. The constants A and B are given by, using Wald's [12] argument (which remains valid even if the observations are not independent), $A \sim (1-\varepsilon_1)/\varepsilon_0$, $B \sim \varepsilon_1/(1-\varepsilon_0)$, where $\varepsilon_0, \varepsilon_1$ are the probabilities of the first and second kinds of error respectively.

The sequence $\{Z_r\}$ is no longer independent, and to derive the O.C. and A.S.N. functions of the test one needs an analogue of Wald's Fundamental Identity for the sequence $\{Z_r\}$ defined on a Markov sequence $\{X_r\}$. An examination of the proof of the analogue for finite-state Markov chains given by Phatarfod [9], shows that the following generalization holds:

Let $X_0, X_1, X_2, \cdots$ be a Markov chain (discrete or continuous state space), on which is defined a sequence of random variables $\{Z_r\}$, $Z_r = h(X_{r-1}, X_r)$. Let S_N denote the cumulative sum $Z_0+Z_1+Z_2+\cdots+Z_N$, and let n be the least positive integer such that S_N does not lie in the open interval (b, a), $(b < 0, a > 0)$. If the m.g.f. (for real θ in an interval (θ_1, θ_2) around zero) of S_N can be written for large N as

$$M_N(\theta) \sim C(\theta)\lambda_1^N(\theta), \tag{1.5}$$

with $\lambda_1(0) = 1$, $\lambda_1''(0) > \lambda_1'(0)^2$, then

$$E[\exp(\theta S_n)\lambda_1^{-n}(\theta)d(\theta|X_n)] = C(\theta), \tag{1.6}$$

for all real θ in (θ_1, θ_2) such that $\lambda_1(\theta) \geqq 1$. (The functions $C(\theta)$ and $d(\theta)$ have the relation $E[d(\theta|X_0)] = C(\theta)$.)

The proof is identical to the proof given in the earlier paper. The only requirement needed is that $P[n > N] \to 0$ as $N \to \infty$. This follows automatically, because (1.5) implies that S_N is asymptotically normal with mean $Nm = N\lambda_1'(0)$ and variance $N\sigma^2 = N[\lambda_1''(0)-\lambda_1'(0)^2]$, and hence

https://doi.org/10.1017/S1446788700010296 Published online by Cambridge University Press

$$\begin{aligned} P[n > N] &\leqq \Pr[a < S_N < b] \\ &= \Phi\left(\frac{a-Nm}{\sqrt{N}\sigma}\right) - \Phi\left(\frac{b-Nm}{\sqrt{N}\sigma}\right) \\ &\to 0 \quad \text{as } N \to \infty. \end{aligned}$$

A considerable amount of work has been done to establish conditions under which result (1.5) holds. The problem is essentially of generalizing results of Perron and Frobenius for non-negative matrices. The earliest such generalization seems to be due to Jentzsch [6] for integral operators with a positive kernel. More recently results have been given by Krein and Rutman [8], Birkhoff [1], Karlin [7], Chun [2], Harris [5] and Vere-Jones [11]. However, it is not immediately known how to determine the quantity $\lambda_1(\theta)$, the 'largest eigen value', and it seems to the present author, that this can be determined only by ad hoc methods.

As in the case of the analogue for finite state Markov chains, $\lambda_1(\theta)$ takes the role of the m.g.f. $M(\theta)$ in the independence case, and has the usual properties e.g. $\lambda_1'(0) = E(Z)$ etc. Moreover, for the case considered above, it is found that $\lambda_1(\theta)$ is a convex function of θ having a minima near zero. This gives a unique non-zero real solution θ_0 for $\lambda_1(\theta) = 1$.

From the Identity (1.6), one can now readily obtain the usual results for the random walk process $\{S_N\}$. Unfortunately, except in special cases, only approximate expressions can be obtained. This is due to the fact that we ignore the excess of S_n over the boundaries, and the terms $C(\theta)$ and $d(\theta|X_n)$ which depend on the distribution of X_0, and the value of X_n respectively.

Putting $\theta = \theta_0$ (whenever such a value exists) in (1.6), the probability of the random walk terminating at b before a is given by,

$$(1.7) \qquad P_b \sim \frac{e^{a\theta_0}-1}{e^{a\theta_0}-e^{b\theta_0}}.$$

The average duration of the random walk process is given by (in a manner similar to that in the case of finite state Markov chains (Phatarfod [10]),

$$(1.8) \qquad E(n) \sim \frac{bP_b+aP_a}{\lambda_1'(0)}, \qquad (\lambda_1'(0) \neq 0).$$

In § 2 we show that (1.5) holds for the Markov chain (1.1) with Z_r's as in (1.3). The O.C. and A.S.N. functions of the test are obtained from (1.7) and (1.8) respectively. In § 3 we consider a Cox-type composite hypothesis test of $H_0 : \rho = \rho_0$ against $H_1 = \rho = \rho_1$ when values of m and σ^2 are not known.

2. Test of the simple hypothesis $\rho = \rho_0$

Consider the sequence $X_0, X_1, X_2, \cdots$ given by

$$X_r = \rho X_{r-1} + Y_r \qquad (|\rho| < 1, \ r = 1, 2, 3, \cdots)$$

the Y_r being independent normal variables with mean zero and variance unity.

With the Z_r's as defined in (1.3), and $S_N = Z_0+Z_1+Z_2+\cdots+Z_N$, we have for real θ,

$$M_N(\theta) = E[\exp(\theta S_N)] = C\int \exp\{-\tfrac{1}{2}[x_0^2+x_N^2+a\sum_{r=1}^{N-1}x_r^2+2b\sum_{r=1}^{N}x_r x_{r-1}]\} \times dx_0\,dx_1\cdots dx_N,$$

where

$$C = \left(\frac{1-\rho_1^2}{1-\rho_0^2}\right)^{\theta/2}\cdot\frac{(1-\rho^2)^{\frac{1}{2}}}{(2\pi)^{(n+1)/2}},\qquad a = \theta(\rho_1^2-\rho_0^2)+\rho^2+1,\quad b = (\rho_0-\rho_1)\theta-\rho.$$

The expression in the square brackets in the integrand is a quadratic form in $x_0, x_1, x_2, \cdots, x_N$ and hence, $M_N(\theta)$ may be written as (for values of θ such that A_{N+1} is positive definite),

$$M_N(\theta) = \left(\frac{1-\rho_1^2}{1-\rho_0^2}\right)^{\theta/2}\frac{(1-\rho^2)^{\frac{1}{2}}}{|A_{N+1}|^{\frac{1}{2}}},\tag{2.1}$$

where $A_{N+1} = (a_{ij})$, the matrix of the quadratic form is given by

$$\begin{array}{ll} a_{ii} = 1(i = 0, N),\quad a_{ii} = a(i \neq 0, N); & (i, j = 0, 1, 2, \cdots, N) \\ a_{ij} = b(|i-j| = 1),\quad a_{ij} = 0 & \text{otherwise.} \end{array}$$

To determine $|A_{N+1}|$, we note

$$|A_{N+1}| = |D_N|-b^2|D_{N-1}|,$$

where D is a matrix similar to A, with 'a' instead of 1 as the element in the first row and column. For $|D_N|$ we have

$$|D_N| = a|D_{N-1}|-b^2|D_{N-2}|,\tag{2.2}$$

with

$$|D_1| = 1,\qquad |D_2| = a-b^2.\tag{2.3}$$

From (2.2) and (2.3) we obtain,

$$|D_N| = \frac{1}{(\mu_1-\mu_2)}[(1-\mu_2)\mu_1^N-(1-\mu_1)\mu_2^N],\tag{2.4}$$

and hence

$$|A_{N+1}| = \frac{1}{(\mu_1-\mu_2)}[(1-\mu_2)^2\mu_1^N-(1-\mu_1)^2\mu_2^N],\tag{2.5}$$

where μ_1, μ_2 are the roots of the equation $x^2-ax+b^2 = 0$; we take

$$\mu_1 = \tfrac{1}{2}[a+(a^2-4b^2)^{\frac{1}{2}}],\quad \mu_2 = \tfrac{1}{2}[a-(a^2-4b^2)^{\frac{1}{2}}].$$

The following properties of $\mu_1(\theta)$, $\mu_2(\theta)$ can be easily verified:

(1) $\mu_1(\theta)$, $\mu_2(\theta)$ are real and unequal for all θ in (θ_1, θ_2) where

$$\theta_1 = -\frac{(1+\rho)^2}{(\rho_1-\rho_0)(2+\rho_0+\rho_1)} (< 0), \qquad \theta_2 = \frac{(1-\rho)^2}{(\rho_1-\rho_0)(2-\rho_0-\rho_1)} (> 0).$$

$\mu_1(\theta)$, $\mu_2(\theta)$ are real and equal for $\theta = \theta_1, \theta_2$.

(2) $a > 0$ for all θ in (θ_1, θ_2), and hence $\mu_1(\theta) > 0$, $\mu_2(\theta) \geqq 0$, $\mu_1(\theta) > \mu_2(\theta)$ for all θ in (θ_1, θ_2).

(3) $\mu_1(\theta)$ is concave and $\mu_2(\theta)$ is convex in (θ_1, θ_2).

(4) The equation $\mu_1(\theta) = 1$ has two solutions $\theta = 0$ and

$$\theta = \theta_0 = \frac{\rho_0+\rho_1-2\rho}{\rho_1-\rho_0}, \qquad \left(\rho \neq \frac{\rho_0+\rho_1}{2}\right),$$

unless $|\rho-\rho_0-\rho_1| > 1$, in which case $\theta_0 = 0$ is the only solution of $\mu_1(\theta) = 1$. For $|\rho-\rho_0-\rho_1| = 1$, the matrix A_{N+1} is singular when $\theta = \theta_0$.

We will restrict θ in (θ_1, θ_2) and the values of ρ such that $|\rho-\rho_0-\rho_1| < 1$. For such a range of values for θ and ρ, we have $\mu_1(\theta)$, $\mu_2(\theta)$ real, $\mu_1(\theta) > 0$, $0 \leqq \mu_2(\theta) < 1$, $\mu_1(\theta) > \mu_2(\theta)$, and a non-zero real solution θ_0 of $\mu_1(\theta) = 1$, except when $\rho = (\rho_0+\rho_1)/2$.

From (2.4) and (2.5) it can be seen that $|D_r| > 0$ $(r = 1, 2, \cdots, N)$ and $|A_{N+1}| > 0$, and hence the matrix A_{N+1} is positive definite. Further, from the properties of $\mu_1(\theta)$, $\mu_2(\theta)$ given above, we have from (2.1), (2.5) as $N \to \infty$,

$$M_N(\theta) \sim \left(\frac{1-\rho_1^2}{1-\rho_0^2}\right)^{\theta/2} \frac{(1-\rho^2)^{\frac{1}{2}}(\mu_1-\mu_2)^{\frac{1}{2}}}{(1-\mu_2)} \cdot \mu_1(\theta)^{-N/2} = C(\theta)\lambda_1^N(\theta)$$

where

$$\lambda_1(\theta) = \mu_1^{-\frac{1}{2}}(\theta).$$

It can be easily seen that $(\lambda_1''(0)-\lambda_1'(0)^2) > 0$, and hence the Identity (1.6) and therefore results (1.7), (1.8) apply. We have the O.C. function as, (putting $a = \log A$, $b = \log B$),

$$L(\rho) \sim \frac{A^{\theta(\rho)}-1}{A^{\theta(\rho)}-B^{\theta(\rho)}},$$

where

$$\theta(\rho) = \frac{\rho_1+\rho_0-2\rho}{\rho_1-\rho_0}.$$

The limitation on the values of ρ is not of serious consequence, as the range of permissible values of ρ include the interval (ρ_0, ρ_1) and beyond. It is interesting to note that the O.C. curve is identical to that of a test for the mean of a normal distribution in the independent observations case.

https://doi.org/10.1017/S1446788700010296 Published online by Cambridge University Press

For the A.S.N. we have,

$$E_\rho(n) = \frac{2[L(\rho)\log B+(1-L(\rho))\log A]}{-\mu_1'(0)}, \qquad \left(\rho \neq \frac{\rho_0+\rho_1}{2}\right).$$

Note

$$-\mu_1'(0)/2 = E_\rho(Z) = \frac{1}{2(1-\rho^2)}[(\rho_0^2-\rho_1^2)+2(\rho_1-\rho_0)].$$

For $\rho = (\rho_0+\rho_1)/2$, a formula can be obtained by differentiating (1.6) twice w.r.t. 'θ' and putting $\theta = 0$. Also since $\theta(\rho)$ and $L(\rho)$ are monotonic functions of ρ, the test given above can be used to test composite hypotheses of the form $\rho \leqq \rho'$ against $\rho > \rho'$, where m and σ^2 are known.

3. Test of composite hypothesis $\rho = \rho_0$

We will now consider the general model given by (1.1) and derive a Cox-type (Cox [3]) test of the composite hypothesis $H_0 : \rho = \rho_0$ against $H_1 : \rho = \rho_1$ $(\rho_1 > \rho_0)$.

The likelihood of the observations $X_0, X_1, X_2, \cdots, X_n$ assuming X_0 has the stationary distribution of the sequence is given by

$$p(x_0, x_1, \cdots, x_n) = \text{const}\cdot\exp\left\{-\frac{1}{2\sigma^2}[(x_0-m)^2+(x_n-m)^2+(1+\rho^2)\times\sum_{r=1}^{n-1}(x_r-m)^2-2\rho\sum_{r=1}^{n}(x_{r-m})(x_{r-1}-m)]\right\}.$$

Writing

$$\bar{x} = (\tfrac{1}{2}x_0+x_1+x_2+\cdots\tfrac{1}{2}x_n)/n,$$

$$S^2 = \tfrac{1}{2}(x_0-\bar{x})^2+\sum_{r=1}^{n-1}(x_r-\bar{x})^2+\tfrac{1}{2}(x_n-\bar{x})^2,$$

and

$$r_n = \sum_{r=1}^{n}(x_r-\bar{x})(x_{r-1}-\bar{x})/S^2,$$

the expression in the square brackets above can be written as

$$\frac{(1-\rho^2)}{2}\{(x_0-m)^2+(x_n-m)^2\}+(1+\rho^2)S^2-2\rho r_n S^2+n(\bar{x}-m)^2(1-\rho)^2.$$

The quantities x_0, x_n, $\bar{x}$, S^2 and r_n, therefore form a set of jointly sufficient statistics for the parameters m, σ^2 and ρ. For large n we may ignore x_0 and x_n and conclude that $\bar{x}$, S^2 and r_n are asymptotically jointly sufficient for m, σ^2 and ρ. Moreover, the distribution of r_n as given by Daniels [4] is independent of m and σ^2. Furthermore the transformation $y_i = ax_i+b$ $((i = 0, \cdots, n)$ satisfies condition (iv) of Cox'

https://doi.org/10.1017/S1446788700010296 Published online by Cambridge University Press

Theorem. Hence a test can be constructed based on the 'observations' $r_2, r_3, \cdots, r_n$. The asymptotic distribution of r_n as given by Daniels has the p.d.f. (ignoring terms of order $O(n^{-\frac{3}{2}})$),

$$L(r_n) \sim \frac{\Gamma(\frac{1}{2}N+\frac{3}{2})}{2\pi^{\frac{1}{2}}\Gamma(\frac{1}{2}N)[N(1-\rho)-(1+\rho)]} \frac{(1-r_n)(1-r_n^2)^{\frac{1}{2}N-1}}{(1-2\rho r_n+\rho^2)^{\frac{1}{2}(N-1)}} \tag{3.1}$$

where

$$N = n + \frac{\rho^2}{(1-\rho^2)}.$$

The test can be written: Continue sampling while

$$\log B < \log \frac{L(r_n|\rho_1)}{L(r_n|\rho_0)} < \log A. \tag{3.2}$$

Substituting $L(r_n)$ from (3.1) in (3.2), using Stirling's approximation, and ignoring terms $O(n^{-2})$, the middle term in (3.2) reduces to

$$\log\frac{1-\rho_0}{1-\rho_1} + \frac{\frac{5}{2}(\rho_1^2-\rho_0^2)+2(\rho_1-\rho_0)(1+\rho_0\rho_1)}{n(1-\rho_1^2)(1-\rho_0^2)}$$
$$+\tfrac{1}{2}(N_0-1)\log(1-2\rho_0 r_n+\rho_0^2)-\tfrac{1}{2}(N_1-1)\log(1-2\rho_1 r_n+\rho_1^2)$$
$$+\left(\frac{N_1-N_0}{2}\right)\log(1-r_n^2),$$

where

$$N_i = n+\frac{\rho_i^2}{1-\rho_i^2}, \qquad (i = 0, 1).$$

For fairly large n, we have (3.2) approximately as

$$\log\frac{B(1-\rho_1)}{(1-\rho_0)} < \frac{\frac{5}{2}(\rho_1^2-\rho_0^2)+2(\rho_1-\rho_0)(1+\rho_0\rho_1)}{n(1-\rho_1^2)(1-\rho_0^2)}$$
$$+\frac{n}{2}\log\frac{(1-2\rho_0 r_n+\rho_0^2)}{(1-2\rho_1 r_n+\rho_1^2)} < \log\frac{A(1-\rho_1)}{(1-\rho_0)},$$

from which acceptance numbers r_n^+ and rejection numbers r_n^- may be calculated.

Acknowledgement

I am grateful to Professor M. S. Bartlett for suggesting the problem to me.

https://doi.org/10.1017/S1446788700010296 Published online by Cambridge University Press

References

[1] G. Birkhoff, 'Extensions of Jentzsch's theorem', *Trans. Amer. Math. Soc.* 85 (1957), 219–227.

[2] K. L. Chung, *Markov chains with stationary transition probabilities* (Springer-Verlag, Berlin, 1960).

[3] D. R. Cox, 'Sequential tests for composite hypotheses', *Proc. Camb. Phil. Soc.* 48 (1952), 290–299.

[4] H. E. Daniels, 'The approximate distribution of serial correlation coefficients', *Biometrika* 43 (1956), 169–185.

[5] T. E. Harris, *The Theory of Branching Processes* (Springer-Verlag, Berlin, 1963).

[6] R. Jentzsch, 'Über Integralgleichungen mit positivem Kern', *J. Reine Angew, Math.* 141 (1912), 235–244.

[7] S. Karlin, 'Positive Operators', *J. Math. Mech.* 8 (1957), 907–938.

[8] M. G. Krein and M. A. Rutman, 'Linear operators leaving invariant a cone in a Banach space', *Amer. Math. Soc. Translation No.* 26, 1950.

[9] R. M. Phatarfod, 'Sequential analysis of dependent observations I', *Biometrika* 52 (1965), 157–165.

[10] R. M. Phatarfod, 'A note on the cumulation sums of Markovian variables', *J. Aust. Math. Soc.* 5 (1965), 285–287.

[11] D. Vere-Jones, 'Ergodic properties of non-negative matrices-I', *Pacific. J. Math.* 22 (1967) 361–386.

[12] A. Wald, *Sequential Analysis* (J. Wiley, New York, 1947).

Monash University
Clayton, Victoria

https://doi.org/10.1017/S1446788700010296 Published online by Cambridge University Press