SOME STUDIES IN MULTINOMIAL SEQUENCE WITH DOUBLE DEPENDENCE BETWEEN SUCCESSIVE OBSERVATIONS

by P.V. Krishna Iyer, F. N. A. *and* S. Lipton, *University of Queensland, St. Lucia, 4067, Australia*

(*Received 4 November 1971*)

The paper is concerned with the distributions of the random variables Y, Z and W which represent the number of *aaa*, *aba* and *abb* transitions respectively, in a second order multinomial Markovian sequence, i.e. a Markovian sequence in which the transition probabilities depend on the two preceding observations. The probabilities for the *aaa*, *aba* and *abb* transitions and also those for *aa*, *ab*, *a*, *b*, etc. at any stage of the sequence have been calculated for given initial conditions. The approximate probabilities for the states *a*, *b*, etc. reduce to two sets for the second-order chain. For the *k*th order chain they reduce to *k*-sets. The asymptotic values of the expected values, variances and covariances of Y, Z, and W have been obtained by two methods: (i) by using the solution for the joint probability generating function of Y, Z, and W, and (ii) by using the transition probabilities at any stage for given initial conditions.

Using the asymptotic values of the variances and covariances for the (ijk) transitions, the expected value and the variance of ψ^2 defined by

$$\psi^2 = \sum \frac{(n_{ijk}-m_{ijk})^2}{m_{ijk}}$$

have been determined. The distribution of this statistic tends to the χ^2-form with A^2/B degrees of freedom. Hence a test of hypothesis that a given sequence follows a second order chain defined by a $(k^2 \times k^2)$ matrix (p_{ijk}) can be made with the χ^2 table by taking the observed $\psi^2 = (A\,\psi^2/B)$ with (A^2/B) degrees of freedom, where A and B are as given below:

$$A = \sum \frac{\sigma^2_{ijk}}{m_{ijk}}$$

$$B = \sum_{lmn} \sum_{ijk} \frac{\sigma^2_{ijk.\,lmn}}{m_{ijk.\,lmn}}.$$

The asymptotic values of A and B have also been given.

I. Introduction

In handling observations arising in Meteorology, Physics, Geology, Economics, Transport etc., sequences in which successive observations are not random frequently occur. The observations are correlated and the correlation is often of a complicated nature. When such data is quantitative, they are analysed by Harmonic, Spectral or Autoregressive methods. Sequences of a qualitative nature, on the other hand, can be analysed using the results established for Markovian sequences. Most of the data met with in practice can be reduced to multinomial sequences and therefore the results of investigations into simple and complex Markov chains are extremely useful in analysing sequences observed in the course of many scientific investigations.

A simple Markov chain is defined by the transition probability matrix (p_{ij}) where p_{ij} is the conditional probability that state i is immediately succeeded by state j. The problem in simple chains is to estimate the transition probabilities (p_{ij}) for a given sequence of observations. This naturally involves not only the estimation of (p_{ij}) but also the complete distribution of the number of (ij) transitions with the moments. Much work has been carried out on these aspects by Whittle (1955), Billingsley (1961), Goodman (1961) and others. This work is quite extensive in the case of simple chains. However, for chains where the dependence extends to more than one observation, existing studies are very limited. In this paper we propose to discuss the following problems in relation to a second order Markov chain:

(i) The distribution of transitions (ijk) and (ij) between successive observations.

(ii) Asymptotic results for conditional and unconditional probabilities for different states.

(iii) The probability for transitions like (ijk) when i, j and k are separated by l and m observations.

(iv) The evaluation of the first two moments of these distributions.

II. Distribution of the Number of Transitions Like (ijk)

For convenience and simplicity we shall consider a second order chain with two states, say, a and b. The results for k-state chains can be derived along the same lines as for two state chains. The second order chain is defined by the matrix

$$
\begin{array}{cc|cccc}
 & & \multicolumn{4}{c}{\text{Succeeding two observations}} \\
 & & aa & ab & ba & bb \\
\hline
 & aa & p_{aaa}\ (p_1) & p_{aab}\ (q_1) & 0 & 0 \\
\text{Previous two observations} & ab & 0 & 0 & p_{aba}\ (p_2) & p_{abb}\ (q_2) \\
 & ba & p_{baa}\ (p_3) & p_{bab}\ (q_3) & 0 & 0 \\
 & bb & 0 & 0 & p_{bba}\ (p_4) & p_{bbb}\ (q_4)
\end{array}
\qquad (1)
$$

The conditional probabilities for $(p(a/aa), p(b/aa), p(a/ab), p(b/bb)$ etc., are written as (p_1, q_1); (p_2, q_2); (p_3, q_3); and (p_4, q_4): they are the transition conditional probabilities for aaa, aab; aba, abb; baa, bab; bba, bbb as explained by the symbols above.

Suppose we are concerned with the joint distribution for transitions like *aaa*, *aab* and *abb*. Let $P_{aa}(n;r,s,u)$, $P_{ab}(n;r,s,u)$, $P_{ba}(n;r,s,u)$, $P_{bb}(n;r,s,u)$ be the probabilities for r, s and u transitions like *aaa*, *aba* and *abb* respectively from n observations, where the last two observations of the sequence on the right side are as given in the suffix attached to P's. Then the following difference equations hold:

$$\left.\begin{aligned} P_{aa}(n;r,s,u) &= p_1P_{aa}(n-1, r-1, s, u) + p_3P_{ba}(n-1;r,s,u) \\ P_{ab}(n;r,s,u) &= q_1P_{aa}(n-1, r, s, u) + q_3P_{ba}(n-1;r,s,u) \\ P_{ba}(n;r,s,u) &= p_2P_{ab}(n-1, r, s-1, u) + p_4P_{bb}(n-1;r,s,u) \\ P_{bb}(n;r,s,u) &= q_2P_{ab}(n-1, r, s, u-1) + q_4P_{bb}(n-1;r,s,u). \end{aligned}\right\} \tag{2}$$

Define the generating function for the joint distribution for *aaa, aba* and *abb* by

$$\phi(t_1,t_2,t_3;n) = \Sigma\, P(n;r,s,u)\, t_1^r\, t_2^s\, t_3^u. \tag{3}$$

We may write this for convenience as $\phi(n)$ and note that

$$\phi(t_1,t_2,t_3;n) = \phi_{aa}(t_1,t_2,t_3;n) + \phi_{ab}(t_1,t_2,t_3;n) + \phi_{ba}(t_1,t_2,t_3;n) + \phi_{bb}(t_1,t_2,t_3;n). \tag{4}$$

Multiplying eqns. (2) by $t_1^r\ t_2^s\ t_3^u$ and adding, we obtain

$$\left.\begin{aligned} \phi_{aa}(n) &= p_1t_1\phi_{aa}(n-1) + p_3\phi_{ba}(n-1) \\ \phi_{ab}(n) &= q_1\phi_{aa}(n-1) + q_3\phi_{ba}(n-1) \\ \phi_{ba}(n) &= p_2t_2\phi_{ab}(n-1) + p_4\phi_{bb}(n-1) \\ \phi_{bb}(n) &= q_2t_3\phi_{ab}(n-1) + q_4\phi_{bb}(n-1) \end{aligned}\right\} \tag{5}$$

This can be written in matrix form and reduces to

$$\begin{pmatrix} \phi_{aa}(n) \\ \phi_{ab}(n) \\ \phi_{ba}(n) \\ \phi_{bb}(n) \end{pmatrix} = \begin{pmatrix} p_1t_1 & 0 & p_3 & 0 \\ q_1 & 0 & p_3 & 0 \\ 0 & p_2t_2 & 0 & p_4 \\ 0 & q_2t_3 & 0 & q_4 \end{pmatrix} \begin{pmatrix} \phi_{aa}(n-1) \\ \phi_{ab}(n-1) \\ \phi_{ba}(n-1) \\ \phi_{bb}(n-1) \end{pmatrix} \tag{6}$$

It may be noted that the column vector

$$[\phi_{ij}(n)] = A^{n-2}\,[\phi_{ij}(2)] \tag{7}$$

where A represents the 4×4 matrix in (6).

If $\lambda_1, \lambda_2, \lambda_3$ and λ_4 are the roots of the determinantal equation of the matrix A, the general solution for $\phi_{ij}(n)$ is given by

$$\phi_{ij}(n) = \sum_{1}^{4} \alpha_r \lambda_r^n \tag{8}$$

where the α's are determined so as to satisfy given initial conditions. As the roots of the determinantal equation cannot be determined readily, we depend on (7) to obtain the generating function.

The cumulants can be obtained by using the usual methods for evaluating cumulants from generating functions. We substitute e^{θ_1}, e^{θ_2} and e^{θ_3} for t_1, t_2 and t_3 in (8), differentiate $\log \phi_{ij}(n)$ with respect to θ_1, θ_2, and θ_3 and then put $\theta_1=0$, $\theta_2=0$ and $\theta_3=0$.

$$\text{i.e.}\quad k_{rsu} = \left(\frac{\partial^{r+s+u}\log\phi(\theta_1,\theta_2,\theta_3;n)}{\partial^r\theta_1\,\partial^s\theta_2\,\partial^u\theta_3}\right)_{\theta_i=0} \sim \left(\frac{\partial^{r+s+u}\log\lambda_1}{\partial^r\theta_1\,\partial^s\theta_2\,\partial^u\theta_3}\right)_{\theta_i=0} \tag{9}$$

where $i=1,2,3$
and λ_1 is the largest root of the determinantal equation for the 4×4 matrix A.

The procedure has been explained in detail by Krishna Iyer and Kapoor (1954) and it can be easily established that the cumulants tend to be linear functions in n. This follows from the main property that one of the roots of the determinantal equation is equal to 1 when $t_1 = t_2 = t_3 = 1$; while the other roots are less than one.

The expected values, variances and covariances for the number of *aaa, aba* and *abb* transitions obtained by differentiating the determinantal equation and using eqn. (9) are given below for large values of n :

Let Y, Z and W stand for the number of *aaa, aba* and *abb* transitions in a sequence of n observations.

Then

$$E(Y) = n\psi \qquad \qquad (10)$$

$$\begin{aligned} V(Y) &= n\psi\,\{1+2p_1+2p_1^2+2p_1^3-7\psi\} \\ &\quad +2n\psi\theta[\eta\psi+p_1^2\,(p_1^2-p_3p_4+2p_1^2p_4-p_1p_3p_4-p_1^2p_2)] \end{aligned} \qquad (11)$$

$$E(Z) = n\xi \qquad \qquad (12)$$

$$\begin{aligned} V(Z) &= n\xi\,\{1+2p_2q_3+2p_2p_3q_1-7\xi\} \\ &\quad +2n\xi\theta(\eta\xi+p_2(p_2-p_4+2p_1p_4+p_1p_3-p_1p_2-2p_3p_4+2p_1p_3p_4 \\ &\quad -p_1^2p_3-p_1p_2p_3+p_1^2p_2p_3+p_1p_3^2p_4-2p_1^2p_3p_4)] \end{aligned} \qquad (13)$$

$$\begin{aligned} \mathrm{Cov}\,(Y, Z) &= n\{-7\psi\xi+\psi\,(q_1p_2+p_1q_1p_2)+\xi(p_3p_1+p_3p_1^2)\} \\ &\quad +n[2\psi\xi\eta\theta+p_2\psi\theta(p_1^2-p_3p_4+2p_1^2p_4-p_1^3-p_1^2p_2 \\ &\quad +p_1^3p_2+p_1^2p_3p_4-2p_1^3p_4)+p_1\xi\theta(p_4-p_2+2p_2p_3 \\ &\quad -2p_3p_4+p_1^2p_3-p_3^2p_4+2p_1^2p_3p_4-p_1^2p_2p_3-p_1p_3^2p_4) \\ &\quad +p_1p_2p_4\theta^2(p_2-p_4)] \end{aligned} \qquad (14)$$

$$E(W) = n\zeta \qquad \qquad (15)$$

$$\begin{aligned} V(W) &= n\zeta\{1+2p_4q_2q_3-7\zeta\} \\ &\quad +2n\zeta\theta(\eta\zeta+p_4(2p_2-2p_4-2p_1p_2-2p_2p_3+p_3p_4-p_2^2+2p_1p_4 \\ &\quad +2p_2p_4+2p_1p_2p_3-2p_1p_2p_4-p_2p_3p_4+p_2^2p_3+p_1p_2^2+p_3^2p_4 \\ &\quad -2p_1p_3p_4-p_2p_3^2p_4-p_1p_2^2p_3+2p_1p_2p_3p_4)] \end{aligned} \qquad (16)$$

$$\begin{aligned} \mathrm{Cov}\,(Y, W) &= n\{-7\psi\zeta+\psi(p_1q_1q_2+q_1q_2)+\zeta p_1p_3p_4\} \\ &\quad +n[2\psi\zeta\eta\theta+p_1\psi\theta(2p_4-2p_2+p_2^2-p_1^2+2p_1p_4-2p_1p_2+2p_1^2p_2 \\ &\quad +p_2^2p_1+p_1p_2p_3-2p_1^2p_4-2p_1p_2p_4+2p_1^2p_2p_4-p_1p_2^2p_3-p_1^2p_2^2) \\ &\quad +p_4\zeta\theta(p_1^2p_3+p_1p_2p_3-2p_1p_3p_4+2p_1^2p_3p_4-2p_1p_2p_3p_4) \\ &\quad +p_1p_4\theta^2(p_2p_3^2p_4-p_1p_2p_3p_4+p_1^2p_3-p_1p_2^2-p_3^2p_4+p_2p_3^2)] \end{aligned} \qquad (17)$$

$$\begin{aligned} \mathrm{Cov}(Z, W) &= n\{-7\xi\zeta+\xi(q_2q_3+p_3q_1q_2)+\zeta p_2p_4q_3\} \\ &\quad +n[2\xi\zeta\eta\theta+p_2\xi\theta(2p_4-p_2-p_3+p_1p_2+p_2p_3-2p_1p_4+p_1p_2p_3 \\ &\quad +p_1p_3^2-2p_1p_3p_4+2p_1^2p_3p_4-p_1p_2p_3^2-p_1^2p_2p_3)+p_4\zeta\theta(p_2^2+p_1p_2 \\ &\quad -2p_2p_4+2p_1p_2p_4-p_1p_2p_3-p_1p_2^2+2p_2p_3p_4-2p_2^2p_3+p_1p_2^2p_3 \\ &\quad +p_2^2p_3^2-2p_1p_2p_3p_4)+p_2p_4q_4\theta^2(p_1+p_2-2p_4+2p_1p_4-2p_1p_2 \\ &\quad +2p_1p_3p_4-p_1^2p_3-p_1p_2p_3+2p_1^2p_2p_3-2p_1^2p_2p_4)] \end{aligned} \qquad (18)$$

where $\theta=\{p_4(q_1+p_3)+q_1(q_2+p_4)\}^{-1}$
$\eta=(3p_4-p_1-2p_2+3p_1p_2+3p_3p_4-5p_1p_4-p_2p_3)$
$\psi=P_aP(a \mid a)p_1, \quad \xi=P_aPp(b \mid a)p_2, \quad \zeta=P_bP(b \mid b)p_4$
and P_a, $P(a \mid a)$, $P(b \mid b)$ are as given later by eqn. (21).

It will be noted that the contribution of the terms included in the square bracket [ ] for the variances and covariances given above is very small compared to the other terms and can be neglected for all practical purposes.

III. Asymptotic Probabilities for the States *aa*, *ab*, *ba* and *bb*

The difference equations for the probabilities for the states *aa*, *ab*, etc. are similar to those given in (2). Defining P_{aa} to be the probability that the lth and $l-1$th observations are in states a and a etc., we note

$$\left.\begin{aligned}
P_{aa}(l)&=p_1P_{aa}(l-1)+p_3P_{ba}(l-1),\\
P_{ab}(l)&=q_1P_{aa}(l-1)+q_3P_{ba}(l-1),\\
P_{ba}(l)&=p_2P_{ab}(l-1)+p_4P_{bb}(l-1),\\
P_{bb}(l)&=q_2P_{ab}(l-1)+q_4P_{bb}(l-1).
\end{aligned}\right\} \tag{19}$$

When l is large $P_{aa}(l)=P_{aa}(l-1)$ and we find

$$\frac{P_{aa}}{p_3p_4}=\frac{P_{ab}}{p_4q_1}=\frac{P_{ba}}{p_4q_1}=\frac{P_{bb}}{q_1q_2}. \tag{20}$$

Asymptotically $P_{ab}=P_{ba}$. Further, the asymptotic conditional probabilities are

$$\left.\begin{aligned}
&P_{aa}=P(a \mid a)=\frac{p_3}{1-p_1+p_3}; \qquad P_{ab}=P(b \mid a)=\frac{1-p_1}{1-p_1+p_3}\\
&P_{aa}=P(a \mid b)=\frac{p_4}{1-p_2+p_4}; \qquad P_{bb}=P(b \mid b)=\frac{1-p_2}{1-p_2+p_4}\\
&\text{also } P_a=\frac{p_4(1-p_1+p_4)}{p_4(1-p_1+p_3)+(1-p_1)(1-p_2+p_4)}\\
&P_b=\frac{(1-p_1)(1-p_2+p_4)}{p_4(1-p_1+p_3)+(1-p_1)(1-p_2+p_4)}.
\end{aligned}\right\} \tag{21}$$

These results have been given by Bateman (1948).

IV. Determination of $P_{aa}(l)$, $P_{ab}(l)$, $P_{ba}(l)$ and $P_{bb}(l)$ for given Initial Conditions

Putting equation. (19) in matrix form, we note that

$$\begin{pmatrix} P_{aa}(l)\\ P_{ab}(l)\\ P_{ba}(l)\\ P_{bb}(l)\end{pmatrix}=\begin{pmatrix} p_1 & 0 & p_3 & 0\\ q_1 & 0 & q_3 & 0\\ 0 & p_2 & 0 & p_4\\ 0 & q_2 & 0 & q_4\end{pmatrix}\begin{pmatrix} P_{aa}(l-1)\\ P_{ab}(l-1)\\ P_{ba}(l-1)\\ P_{bb}(l-1)\end{pmatrix} \quad .. \tag{22}$$

Defining the 4×4 matrix on the right hand side to be $\triangle$, we obtain

$$(P_{ij}(l))=\triangle^{l-2}(\theta)$$

where (θ) is the column vector representing the initial probabilities for the first two observations.

The general solution depends on the roots of the determinantal equation

$$|\ \lambda-\triangle\ |=0 \qquad (23)$$

which is equal to

$$(\lambda-1)\ [(\lambda+q_1)\ (\lambda-q_4)\ \lambda+\lambda\ \{q_1(q_2-q_4)+p_2(q_1-q_3\}$$
$$-(q_1-q_3)\ (p_2q_4-q_2p_4)]=0. \qquad (24)$$

If the constant term is small, we may take the roots of the above equation to be 1, $(-q_1+\delta_1)$, $(q_4+\delta_2)$, 0, where δ_1 and δ_2 are small and can be determined easily by solving the quadratic resulting from (24) making the constant term zero. It may be noted that one of the roots is negative and the other is positive.

The general solution for $P_{ab}(l)$ takes the form

$$P_{ab}(l)=P_{ab}+a_1\lambda_1^{l-2}+a_2\lambda_2^{l-2} \qquad (25)$$

where λ_1 and λ_2 stand for $(-q_1+\delta_1)$ and $(q_4+\delta_2)$ respectively and P_{ab} is the asymptotic probability for two successive observations to be a and b as $l\to\infty$. Using the initial conditions and putting $l=2$ and 3 we get

$$\begin{aligned}\phi_1 &= (\theta_{ab}-P_{ab}) = a_1+a_2,\\ \phi_2 &= (\psi_{ab}-P_{ab}) = a_1\lambda_1+a_2\lambda_2,\end{aligned} \qquad (26)$$

where θ_{ab} is the probability that the first two observations are in states a and b, and ψ_{ab} is the probability that the 2nd and 3rd observations are in states a and b for given initial conditions.

Since

$$(P_{ab}(l)-P_{ab}=\frac{\phi_2}{(\lambda_1-\lambda_2)}(\lambda_1^{l-2}-\lambda_2^{l-2})-\frac{\phi_1\lambda_1\lambda_2}{\lambda_1-\lambda_2}(\lambda_1^{l-3}-\lambda_2^{l-3}) \qquad (27)$$

and

$$P_{ab}(l)+P_{aa}(l)=P_a(l-1) \qquad (28)$$

it can be easily seen that

$$(P^a(l)-P_a)=C_1(\lambda_1^{l-1}-\lambda_2^{l-1})-C_2(\lambda_1^{l-2}-\lambda_2^{l-2}). \qquad (29)$$

Since one of the roots is negative and the other positive,

$$\left.\begin{aligned}&(P_a(l)-P_a)\sim C_1'\lambda_1^{l-1}+C_2'\ |\ \lambda_2\ |^{\ l-1} \text{ if } l \text{ is even}\\ \text{and}\\ &(P_a(l)-P_a)\sim C_1'\lambda_1^{l-2}+C_2'\ |\ \lambda_2\ |^{\ l-2} \text{ if } l \text{ is odd}\end{aligned}\right\} \qquad (30)$$

where C_1' and C_2' are as obtained from (29). When $\lambda_1 = |\lambda_2|$, the solution finally takes the form $D\,|\lambda_1|^{l-1}$ or $D\,|\lambda_1|^{l-2}$ according as l is even or odd.

V. Determination of $P_a(r)$ and $P_b(r)$ by a Different Approach

The solution for $P_a(l)$ obtained above depends on the nature of l (odd or even). We shall now obtain the same solution by a different approach. The difference equations satisfied by $P_a(r)$ and $P_b(r)$, the probabilities that the r^{th} observation (from the left end) are in states a and b are as follows:

$$\left.\begin{aligned} P_a(r) &= P_a(r-2)\{p_aP(a_{r-1} \mid a_{r-2}) + p_2P(b_{r-1} \mid a_{r-2})\} \\ &\quad + P_b(r-2)\{p_3P(a_{r-1} \mid b_{r-2}) + p_4P(b_{r-1} \mid b_{r-2})\} \\ P_b(r) &= P_a(r-2)\{q_1P(a_{r-1} \mid a_{r-2}) + q_2P(b_{r-1} \mid a_{r-2})\} \\ &\quad + P_b(r-2)\{q_3P(a_{r-1} \mid b_{r-2}) + q_4P(b_{r-1} \mid b_{r-2})\} \end{aligned}\right\} . \tag{31}$$

From the previous discussions it can be seen that

$$\left.\begin{aligned} p(a_r \mid a_{r-1}) &\sim P_{aa} + \alpha_1\lambda_1^{r-1} + \beta_1\lambda_2^{r-1} \\ p(b_r \mid a_{r-1}) &\sim P_{ab} + \alpha_3\lambda_1^{r-1} + \beta_3\lambda_2^{r-1} \\ p(a_r \mid b_{r-1}) &\sim P_{ba} + \alpha_2\lambda_1^{r-1} + \beta_2\lambda_2^{r-1} \\ p(b_r \mid b_{r-1}) &\sim P_{bb} + \alpha_4\lambda_1^{r-1} + \beta_4\lambda_2^{r-1} \end{aligned}\right\} \tag{32}$$

where P_{aa}, P_{ab}, P_{ba}, and P_{bb} are the asymptotic transition probabilities and α and β are constants which depend on the initial conditions.

Equations (31) reduce to the form

$$\begin{pmatrix} (E^2-\theta_1) & -\theta_2 \\ -\theta_3 & (E^2-\theta_4) \end{pmatrix} \qquad \begin{pmatrix} P_a(r) \\ P_b(r) \end{pmatrix} = 0 \tag{33}$$

where θ_1, θ_2 ; θ_3 and θ_4 are the coefficients of $P_a(r-2)$ and $P_b(r-2)$ in eqn. (31).
Now $\theta_1 + \theta_3 = 1$, and $\theta_2 + \theta_4 = 1$.
Therefore $P_a(r)$ and $P_b(r)$ satisfy the equation

$$(\lambda^2-1)\,[\lambda^2-(\theta_3-\theta_4)] = 0. \tag{34}$$

It may be noted that

$$\begin{aligned} (\theta_3-\theta_4) &\sim (p_{aa}+\alpha_1\lambda_1^{r-1}+\beta_1\lambda_2^{r-1})\,q_1 + (p_{ab}+\alpha_3\lambda_1^{r-1}+\beta_3\lambda_2^{r-1})q_2 \\ &\quad -(p_{ba}+\alpha_2\lambda_1^{r-1}+\beta_2\lambda_2^{r-1})\,q_3 - (p_{bb}+\alpha_4\lambda_1^{r-1}+\beta_4\lambda_2^{r-1})q_4 \\ &\sim (p_{aa}q_1+p_{ab}q_2-p_{ba}q_3-p_{bb}q_4) + A\lambda_1^{r-1} + B\lambda_2^{r-1} \quad \ldots \quad \ldots \end{aligned} \tag{35}$$

where A and B are small quantities less than one. Therefore for all practical purposes the contributions from λ_1^{r-1} and λ_2^{r-1} can be omitted and $(\theta_3 - \theta_4)$ may be taken to be $(p_{aa}q_1+p_{ab}q_2-p_{ba}q_3-p_{bb}q_4)$. The solution of the difference eqn. (33) can be written as one of two expressions according as r is odd or even.

For r even $(P_a(r)-P_a) = (P_a'-P_a)\xi^{r/2}$,

For r odd $(P_a(r)-P_a) = (P_a'-P_a)\,(P_{aa}'-P_{ba}')\xi^{\frac{r-1}{2}}$, $\quad \ldots \quad \ldots$ (36)

where P_a' and P_{aa}', P_{ba}' are the initial probabilities for first observations to be a and the transition probabilities for the second observations to be a respectively and $\xi^2=(\theta_3-\theta_4)$.

A closer solution will be given by

$$\left.\begin{aligned}(P_a(r)-P_a) &\sim (P_a'-P_a)\xi^{\frac{r}{2}}\prod_{s=1}^{\frac{r}{2}}\{1+\tfrac{1}{2}(\gamma_1\lambda_1^s+\gamma_2\lambda_2^s)\} \\ &\sim (P_a'-P_a)\xi^{\frac{r}{2}}\left(1+\tfrac{1}{2}\ \frac{\gamma_1\lambda_1}{1-\lambda_1}+\frac{\gamma_2\lambda_2}{1-\lambda_2}\right)\end{aligned}\right\} \quad \cdots \quad \cdots \quad (37)$$

[where the γ's are small quantities compared with ξ.]
A similar solution can also be given for r odd.

The importance of this approach lies in the fact that this indicates that for a 2nd order k-state Markov Chain the solution can be reduced to k sets and will depend on $r=\text{mod}\ |\ k\ |$, where k is the order of dependence between the observations.

VI. Moments for the Distributions of the Number States and Transitions

It has already been pointed out that the cumulants for these distributions can be obtained from the solution for the probability generating functions. But the solution depends on the roots of the determinantal equation and this cannot be determined for k-state and higher order chains. Consequently, it is not possible to obtain the exact expression for the cumulants in any manageable form. The asymptotic values, however, can be determined on the basis of the largest root of the determinantal equation using the procedure outlined by Krishna Iyer and Kapoor (1954) mentioned earlier. This can also be calculated by finding the expectations of the different terms involved in the moments as explained below:

Taking the distribution for the number of transitions like aba for the second-order two state Markov Chain, we may note that a transition is obtained from three consecutive points and the total number of transitions is the sum of $(n-2)$ random variables like $z_{r,r+1,r+2}$ which take values 1 if the r^{th}, $(r+1)^{th}$ and $(r+2)^{th}$ observations together yield one aba transition and zero otherwise i.e. r^{th} observation is a, $(r+1)^{th}$ is b and $(r+2)^{th}$ observation is a. Assuming that Z represents the total number of transitions like aba in a sequence of n observations

$$Z=z_{123}+z_{234}+\cdots\cdots z_{n-2,n-1,n} \quad \cdots \quad \cdots \quad (38)$$

where

$z_{r,r+1,r+2}=1$ if r, $r+1$, and $(r+2)^{th}$ observations are a, b and a respectively.

$z_{r,r+1,+2}=0$ if r, $r+1$ and $(r+2)^{th}$ observations do not give a transition aba,

$$E(Z) = \sum_{r=1}^{n-2} E(z_{r,\,r+1,\,r+2}). \qquad .. \qquad .. \qquad (39)$$

The probability for the r^{th} and $(r+1)^{\text{th}}$ observation to be a and b under any initial conditions has already been determined and is

$$P_r(ab) = P_{ab} + \alpha_1 \lambda_1^{r-2} + \alpha_2 \lambda_2^{r-2}. \qquad .. \qquad .. \qquad (40)$$

Therefore $E(z_{r,\,r+1,\,r+2}) = (P_{ab} + \alpha_1 \lambda_1^{r-2} + \alpha_2 \lambda_2^{r-2}) p_2$.

Substituting the above value in $E(Z)$ we get

$$E(Z) = (n-2)\, P_{ab} p_2 + p_2 \left\{ \alpha_1 \frac{1-\lambda_1^{n-1}}{1-\lambda_1} + \alpha_2 \frac{1-\lambda_2^{n-1}}{1-\lambda_2} \right\}$$

$$\sim (n-2)\, P_{ab} p_2 + p_2 \left\{ \frac{\alpha_1}{1-\lambda_1} + \frac{\alpha_2}{1-\lambda_2} \right\}, \qquad .. \qquad .. \qquad (41)$$

$$E(Z^2) = \sum_{r=1}^{n-2} E(z^2_{r,\,r+1,\,r+2}) + 2 \sum_{r=1}^{n-2} \sum_{s=2}^{n-2} E(z_{r,\,r+1,\,r+2},\ z_{r+s,\,r+s+1,\,r+s+2})$$

$$= \sum_{r=1}^{n-2} E(z_{r,\,r+1,\,r+2}) + 2 \sum \sum E(z_{r,\,r+1,\,r+2},\ z_{r+s,\,r+s+1,\,r+s+2}) \quad .. \qquad (42)$$

we have already evaluated $\sum E(z_{r,\,r+1,\,r+2})$ in (41).

To evaluate $E(z_{r,\,r+1,\,r+2},\ z_{r+s,\,r+s+1,\,r+s+2})$ we use the general expression for $P_r(aba)$ for given initial conditions. It may, however, be mentioned that the basis for this evaluation is $P_s(ab \mid (ab)_r)$ which is obtained by modifying the initial conditions in (26).

The asymptotic expected values and the variances and covariance for the number of *aaa* (Y) and *aba* (Z) transitions are given below:

$$E(Y) \sim nP(aa)p_1 = n\psi, \qquad .. \qquad .. \qquad .. \qquad .. \qquad (43)$$

where $P(aa)$ is the asymptotic probability for *aa* transitions discussed in Section III.

$$V(Y) \sim n\psi\Big(1 + 2p_1 + 2p_1^2 + 2p_1^3 - 7\psi\ \frac{2\alpha_1 p_1 \lambda_1^4}{1-\lambda_1} + \frac{2\alpha_1 p_1 \lambda_2^4}{1-\lambda_2}\Big), \qquad (44)$$

$$E(Z) \sim nP(ab)p_2 = n\xi, \qquad .. \qquad .. \qquad .. \qquad .. \qquad (45)$$

where $P(ab)$ is the asymptotic probability for the ab ransition.

$$V(Z) \sim n\xi \left(1+2q_3p_2+2p_3q_1p_2-7\xi+ \frac{2\beta_1p_2\lambda_1^4}{1-\lambda_1} + \frac{2\beta_2p_2\lambda_2^4}{1-\lambda_2} \right) \tag{46}$$

$$\text{Cov}(YZ) \sim n\Bigg[-7\psi\xi+q_1p_2\psi+p_3p_1\xi+p_3p_1^2\xi+p_1q_1p_2\psi+\psi p_2 \left(\frac{\gamma_1\lambda_1^4}{1-\lambda_1}+\frac{\gamma_2\lambda_2^4}{1-\lambda_2}\right)$$
$$+ \xi p_1 \left(\frac{\delta_1\lambda_1^4}{1-\lambda_1} + \frac{\delta_2\lambda_2^4)}{1-\lambda_2}\right)\Bigg] \tag{47}$$

where

$$a_1=[(1-P_{aa})\lambda_2-(p_1-P_{aa})] /(\lambda_2-\lambda_1)$$
$$a_2=[(1-P_{aa})\lambda_1-(p_1-P_{aa})] /(\lambda_1-\lambda_2)$$
$$\beta_1=[(1-P_{ab})\lambda_2-(p_2-P_{ab})] /(\lambda_2-\lambda_1)$$
$$\beta_2=[(1-P_{ab})\lambda_1-(p_2-P_{ab})] /(\lambda_1-\lambda_2)$$
$$\gamma_1= \frac{P_{ab}(\lambda_2-1)}{(\lambda_1-\lambda_2)}, \quad \gamma_2 = \frac{P_{ab}(1-\lambda_1)}{(\lambda_1-\lambda_2)}.$$
$$\delta_1= \frac{P_{aa}(\lambda_2-1)}{(\lambda_1-\lambda_2)}, \quad \delta_2= \frac{P_{aa}(1-\lambda_1)}{(\lambda_1-\lambda_2)}.$$

Comparing eqns. (43), (44), (45), (46) and (47) with (10), (11), (12), (13), and (14) we note that asymptotically the difference between the respective values is very small and negligible. Similar results can be obtained for the variable W as given in Section II. It may be mentioned that the results given above and in Section II can be made applicable for a k -state Markov-chain by putting P_{aaa}, P_{aba}, P_{baa}, P_{bba} for p_1, p_2, p_3 and p_4 and P_{aab}, P_{abb}, P_{bab} and P_{bbb} for q_1, q_2, q_3 and q_4.

VII. Tests of Hypothesis of Double Dependence

To test the hypothesis that a given sequence of observations belong to a multinomial Markovian sequence with double dependence between successive observations defined by a $(k^2\times k^2)$ matrix (p_{ijk}), as given in Section II, we may use the statistic

$$\psi^2 = \sum \frac{(n_{ijk} - m_{ijk})^2}{m_{ijk}} \tag{48}$$

where n_{ijk} and m_{ijk} are the observed and expected number of ijk transitions in the sequence. This statistic is an extension of the usual χ^2-test and can be approximated to the likelihood ratio test by maximisation of the probabilities. Patankar (1955) has shown that the distribution of this statistic tends to the χ^2-form with A^2/B degrees of freedom and $(A\psi^2)/B$ as the observed value, where A and B are given by

$$E(\psi^2) = \sum \frac{\sigma^2_{ijk}}{m_{ijk}} = A, \tag{49}$$

$$V(\psi^2) \approx 2 \sum_{lmn} \sum_{ijk} \frac{\sigma^2_{ijk.lmn}}{m_{ijk} \, . m_{lmn}} = 2B. \tag{50}$$

In the above expressions

$$\sigma^2_{ijk} = E(n_{ijk} - m_{ijk})^2, \tag{51}$$

$$\sigma_{ijk.lmn} = E(n_{ijk} - m_{ijk})(n_{lmn} - m_{lmn}). \tag{52}$$

This can be easily seen by taking the expected value and variance of $(A\psi^2)/B$.

$$E\{(A\psi^2)/B\} = A^2/B \quad \ldots \quad \ldots \quad \ldots \quad \ldots \tag{53}$$

$$V\{(A\psi^2)/B\} = \frac{2A^2}{B} \cdot \quad \ldots \quad \ldots \quad \ldots \quad \ldots \tag{54}$$

They are the same as that of a χ^2-distribution with A^2/B degrees of freedom.

The variance and covariance for the number of ijk-transitions have been given in Sections II and VI. Using them, A and B can be calculated, the observed χ^2 and its degrees of freedom can be determined and the probability that the given sequence has its parameters defined by (p_{ijk}) can be evaluated. It may be mentioned that for a sequence in which the observations are independent

$$E(\psi^2) = (k^3 - 1), \; V(\psi^2) = 2(k^3 + 2k^2 + 2k - 5). \quad \ldots \quad \ldots \quad \ldots \tag{55}$$

If they are not independent, they are given by Krishna Iyer and Samarasimhudu (1968).

$$E(\psi^2) \simeq k^3 - 5 + 2 \sum p_{iii} + 2 \sum \sum p_{jij} \, p_{iji} + 2 \sum_i \sum_s \left(1 - \frac{s}{n}\right) P_{ii}^{(s+2)} \tag{56}$$

$$\begin{aligned} V(\psi^2) \simeq 2[k^3 + 2k^2 + 2 \sum_{mij} \frac{P_m \, p_{mi} \, p^2_{mij}}{p_i \, p_{ij}} - 25 \\ + 8 \sum p_{iii} + 6 \sum_{ij} p_{iji} \, p_{jij} + 4 \sum p^2_{iii} \\ + 4 \sum_{i \neq j \neq m} p_{ijm} \, p_{jmi} \, p_{mij} + 2 \sum_{i \neq j \neq m \neq r} p_{jmr} \, p_{mri} \, p_{vij} \, p_{ijm} \\ + 4 \sum_i \sum_s \left(1 - \frac{s}{n}\right) \left\{ P_{ii}^{(s+1)} + P_{ii}^{(s+2)} + P_{ii}^{(s+3)} + P_{ii}^{(s+4)} \right\} \\ + 4 \sum_i \sum_{s_1, s_2} \left(1 - \frac{s_1}{n}\right) \left(1 - \frac{s_2}{n}\right) P_{ii}^{(s_1 + s_2 + 4)} \,] \end{aligned} \tag{57}$$

where $P_{ii}^{(s)} = p_{ii}^{(s)} - P_i$ and $p_{ii}^{(s)}$ represents the usual conditional probability.

It may be noted that the contribution of the terms involving $\sum \sum$ in the above expressions is very small compared to k, the number of states.

References

Bateman, G. (1948). On the power function of the longest run as a test for randomness in a sequence of alternatives. *Biometrika*, **35**, 97-112.

Billingsley, P. (1961). Statistical methods in Markov chain. *Ann. math. Statist.*, **32**, 12-40.

Goodman, L.A. (1961). *Ann. math. Statist.*, **32**, 49.

Krishna Iyer, P.V., and Kapoor, M.N. (1954). Probability distributions arising from points on a line. *Biometrika*, **41**, 553-54.

Krishna Iyer, P.V., and Samarasimhudu (1968). Contributions in Statistics and Agricultural Sciences. Indian Society of Agricultural Statistics, p. 99.

Patankar, V.N. (1954). The goodness of fit of frequency distributions obtained from stochastic processes. *Biometrika*, **41**, 450-462.

Whittle, P. (1955). Some distribution and moment formulae for the Markov chain. *Jl R. statist. Soc.*, B **17**, 235-42.