Res. Lett. Inf. Math. Sci., 2005, Vol. **8**, pp 203-207
Available online at http://iims.massey.ac.nz/research/letters/

Some properties of transition matrices for chain binomial models

G. JONES[1]

[1] *Institute of Information Sciences & Technology*
Massey University, Palmerston North, New Zealand †

A chain binomial model is a Markov chain with a transition matrix whose rows are binomial probabilities. Two such chains are presented and illustrated with possible applications. The paper will focus in particular on some interesting properties of the transition matrices.

1 Introduction

A chain binomial model comprises a sequence of random variables $\{X_t\}$ such that the conditional distribution $X_t|X_{t-1}$ is binomial distribution whose parameters are functions of X_{t-1}. It is clear from this definition that X_t has the Markov property, and that the sequence forms a Markov chain whose transition matrix P consists of rows of binomial probabilities. We consider here two examples where P has some interesting and unusual spectral properties, i.e. the eigenvalues and eigenvectors follow a simple pattern. These properties can be easily verified, but a constructive proof of the results awaits discovery.

We first consider a finite chain based on a simple infection model. Our second example has an infinite state space and is based on the negative binomial distribution.

2 An infection model

Suppose a population of n individuals susceptible to a certain disease. Let X_t be the number still uninfected at time t, for $t = 0, 1, 2, ...$, and let p_t be the probability that an individual is infected at time t. Such a model of infection was considered by Greenwood (1949) with p_t proportional to the number of recently infected, and therefore infective, individuals. Their chain stopped when there were no new infectives. This model is suitable for certain diseases, like measles, which have a short fixed period of infectivity. Some results for this model using Markov chain theory were given by Gani and Jerwood (1971).

Jones et al. (2000) considered a simpler situation in which disease is caused by a fixed source of infection, so that $p_t = p$ a constant. Here the chain stops only when all

† Email address: g.jones@massey.ac.nz

individuals have become infected, i.e. when $X_t = 0$. Taking $n = 8$ for definiteness and writing $1 - p = q$, the transition matrix P is

	0	1	2	3	4	5	6	7	8
0	1	0	0	0	0	0	0	0	0
1	p	q	0	0	0	0	0	0	0
2	p^2	$2pq$	q^2	0	0	0	0	0	0
3	p^3	$3p^2q$	$3pq^2$	q^3	0	0	0	0	0
4	p^4	$4p^3q$	$6p^2q^2$	$4pq^3$	q^4	0	0	0	0
5	p^5	$5p^4q$	$10p^3q^2$	$10p^2q^3$	$5pq^4$	q^5	0	0	0
6	p^6	$6p^5q$	$15p^4q^2$	$20p^3q^3$	$15p^2q^4$	$6pq^5$	q^6	0	0
7	p^7	$7p^6q$	$21p^5q^2$	$35p^4q^3$	$35p^3q^4$	$21p^3q^4$	$7pq^6$	q^7	0
8	p^8	$8p^7q$	$28p^6q^2$	$56p^5q^3$	$70p^4q^4$	$56p^3q^5$	$28p^2q^6$	$8pq^7$	q^8

The N-step ahead matrix P^N, giving the transition probabilities for a period of N "days", is commonly derived from the canonical decomposition

$$P = E \Lambda E^{-1}$$

where E is the matrix of right-eigenvectors of P and Λ is the diagonal matrix of eigenvalues, giving

$$P^N = E \Lambda^N E^{-1}$$

For the above P we find that E is

	0	1	2	3	4	5	6	7	8
0	1	0	0	0	0	0	0	0	0
1	1	1	0	0	0	0	0	0	0
2	1	2	1	0	0	0	0	0	0
3	1	3	3	1	0	0	0	0	0
4	1	4	6	4	1	0	0	0	0
5	1	5	10	10	5	1	0	0	0
6	1	6	15	20	15	6	1	0	0
7	1	7	21	35	35	21	7	1	0
8	1	8	28	56	70	56	28	8	1

i.e. E is a matrix of binomial coefficients. Moreover the inverse E^{-1} is simply

	0	1	2	3	4	5	6	7	8
0	1	0	0	0	0	0	0	0	0
1	-1	1	0	0	0	0	0	0	0
2	1	-2	1	0	0	0	0	0	0
3	-1	3	-3	1	0	0	0	0	0
4	1	-4	6	-4	1	0	0	0	0
5	-1	5	-10	10	-5	1	0	0	0
6	1	-6	15	-20	15	-6	1	0	0
7	-1	7	-21	35	-35	21	-7	1	0
8	1	-8	28	-56	70	-56	28	-8	1

because the minor of E_{ij} is E_{ji}. These results may be verified for any given n, but I have not been able to find a general proof. It is easily seen however that the eigenvalues of P are 1, q, q^2, q^3, ..., q^8, because the eigenvalues of a triangular matrix are the diagonal elements.

3 A chain negative binomial model

Suppose that a gambler is able to play a game in which he wins \$1 with probability p and loses \$1 with probability $1 - p = q$. He may play as many times as he wishes, and so resolves to keep playing each day until he as made a profit of \$1, with any debts incurred to be paid off the following day. Let X_{t-1} be the debt incurred on day t-1, to be paid off from his winnings on day t. Then on day t he must keep playing until he has won X_{t-1} +1 games, so

$$X_t \mid X_{t-1} \sim \text{Negative Binomial}(X_{t-1}+1, p)$$

This has been proposed by Jones and Lai (2002) as a possible model for the debt burden of a business. They show that if $p < ½$ then the unconditional mean and variance of X_t increase without limit, but if $p > ½$ then X_t converges to a geometric distribution with "probability of success" parameter

$$p_s = 2 - \frac{1}{p}$$

If we regard the sequence $\{X_t\}$ as a Markov chain, then the transition matrix P consists of rows of negative binomial probabilities as shown below.

$$
\begin{array}{c|ccccccccc|}
 & 0 & 1 & 2 & 3 & 4 & 5 & \dots & j & \dots \\
\hline
0 & p & pq & pq^2 & pq^3 & pq^4 & pq^5 & \dots & pq^j & \dots \\
1 & p^2 & 2p^2q & 3p^2q^2 & 4p^2q^3 & 5p^2q^4 & 6p^2q^5 & \dots & \dots & \dots \\
2 & p^3 & 3p^3q & 6p^3q^2 & 10p^3q^3 & 15p^3q^4 & 21p^3q^5 & \dots & \dots & \dots \\
3 & p^4 & 4p^4q & 10p^4q^2 & 20p^4q^3 & 35p^4q^4 & 56p^4q^5 & \dots & \dots & \dots \\
4 & p^5 & 5p^5q & 15p^5q^2 & 35p^5q^3 & 70p^5q^4 & 126p^5q^5 & \dots & \dots & \dots \\
5 & p^6 & 6p^6q & 21p^6q^2 & 56p^6q^3 & 126p^6q^4 & 252p^6q^5 & \dots & \dots & \dots \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\
i & p^{i+1} & \dots & \dots & \dots & \dots & \dots & \dots & \binom{i+j}{j} p^{i+1}q^j & \dots \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots
\end{array}
$$

To examine the canonical decomposition of P we begin, following the example of the previous section, by defining the matrix of coefficients E as

$$
\begin{array}{c|ccccccccc|}
 & 0 & 1 & 2 & 3 & 4 & 5 & \dots & j & \dots \\
\hline
0 & 1 & 1 & 1 & 1 & 1 & 1 & \dots & 1 & \dots \\
1 & 1 & 2 & 3 & 4 & 5 & 6 & \dots & \dots & \dots \\
2 & 1 & 3 & 6 & 10 & 15 & 21 & \dots & \dots & \dots \\
3 & 1 & 4 & 10 & 20 & 35 & 56 & \dots & \dots & \dots \\
4 & 1 & 5 & 15 & 35 & 70 & 126 & \dots & \dots & \dots \\
5 & 1 & 6 & 21 & 56 & 126 & 252 & \dots & \dots & \dots \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\
i & 1 & \dots & \dots & \dots & \dots & \dots & \dots & \binom{i+j}{j} & \dots \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots
\end{array}
$$

If we now denote the columns of E by e_0, e_1, e_2, e_3, ..., then we find that

$$
\begin{aligned}
Pe_0 &= e_0 \\
Pe_1 &= e_0 + q_s e_1 \\
Pe_2 &= e_0 + 2q_s e_1 + (q_s)^2 e_2 \\
Pe_3 &= e_0 + 3q_s e_1 + 3(q_s)^2 e_2 + (q_s)^3 e_3 \qquad \text{etc.}
\end{aligned}
$$

so the right-eigenvectors of P are given by

$$v_0 = e_0$$
$$v_1 = e_0 - p_s e_1$$
$$v_2 = e_0 - 2p_s e_1 + (p_s)^2 e_2$$
$$v_3 = e_0 - 3p_s e_1 + 3(p_s)^2 e_2 - (p_s)^3 e_3 \qquad \text{etc.}$$

and the eigenvalues of P are 1, q_s, $(q_s)^2$, $(q_s)^3$, ...

To derive the left-eigenvectors of P, we first define P_s to be the matrix P with p replaced by p_s. Now denote the rows of P_s by r_0, r_1, r_2, r_3, ..., and we find that the right-eigenvectors of P are given by

$$u_0 = r_0$$
$$u_1 = r_0 - r_1$$
$$u_2 = r_0 - 2r_1 + r_2$$
$$u_3 = r_0 - 3r_1 + 3r_2 - r_3 \qquad \text{etc.}$$

Again, these relationships may be easily verified. The author would be interested to see a general proof.

References

Gani J. & Jerwood D (1971) Markov chain methods in chain binomial epidemic models. *Biometrics* 27:591-604.

Greenwood M (1949). The infectiousness of measles. *Biometrika* 36:1-8.

Jones G, Lai C. D. & Rayner J. C. W (2000). Some results for a chain binomial model. *Math. Scientist* 25:94-99.

Jones G. & Lai C. D (2002). Some parametric chain models. *Math. Scientist* 27:117-125.