

## MULTIMODALITY OF THE MARKOV BINOMIAL DISTRIBUTION

MICHEL DEKKING \* \*\* AND  
DERONG KONG, \* \*\*\* *Delft University of Technology*

### Abstract

We study the shape of the probability mass function of the Markov binomial distribution, and give necessary and sufficient conditions for the probability mass function to be unimodal, bimodal, or trimodal. These are useful to analyze the double-peaking results of a reactive transport model from the engineering literature. Moreover, we give a closed-form expression for the variance of the Markov binomial distribution, and expressions for the mean and the variance conditioned on the state at time  $n$ .

*Keywords:* Markov binomial distribution; log-concavity; double peaking in kinetic transport

2010 Mathematics Subject Classification: Primary 60J10  
Secondary 60J20

### 1. Introduction

The Markov binomial distribution occurs in diverse applications. Examples are weather forecasting, stock market trends, DNA matching, quality control (cf. [13]), and biometrics (cf. [3]; see also [4]). In 1924 Markov [11] showed that, under certain conditions, a Markov binomial distribution is asymptotically normally distributed. Later in 1953 Dobrušin [6] studied some other limit distributions of a Markov binomial distribution. In 1960 Edwards [7] rediscovered the Markov binomial distribution in connection with work on the human sex ratio. More recently, many authors studied its distribution and moments (cf. [8], [9], and [15]), and its approximations by compound Poisson distributions and binomial distributions (cf. [1], [2], and [16]).

Our interest in the possible lack of unimodality of the Markov binomial distribution arose from [12], where the authors deduced from simulations a somewhat surprising behaviour of double peaking in the concentration of the aqueous part of a solute undergoing kinetic adsorption and moving by advection and dispersion. In our paper [5], we explained this behaviour rigorously using the multimodality properties derived in the present paper.

Let  $\{Y_k, k \geq 1\}$  be a Markov chain on the two states  $\{S, F\}$  with initial distribution  $\nu = (\nu_S, \nu_F)$  and transition matrix

$$P = \begin{bmatrix} P(S, S) & P(S, F) \\ P(F, S) & P(F, F) \end{bmatrix} = \begin{bmatrix} 1 - a & a \\ b & 1 - b \end{bmatrix}, \quad (1.1)$$

where we assume that  $0 < a, b < 1$  throughout the paper. The *Markov binomial distribution* (MBD) is defined, for  $n \geq 1$ , as the distribution of the random variable which counts the number

Received 22 February 2011; revision received 15 August 2011.

\* Postal address: Delft University of Technology, Faculty EW1, PO Box 5031, 2600 GA Delft, The Netherlands.

\*\* Email address: f.m.dekking@math.tudelft.nl

\*\*\* Email address: d.kong@tudelft.nl

of successes in  $n$  experiments with the two outcomes success and failure:

$$K_n = \sum_{k=1}^n \mathbf{1}_{\{Y_k=S\}}.$$

We say that  $K_n$  is a  $\text{Bin}(n, a, b, \nu)$  distributed random variable. Clearly, the MBD generalizes the binomial distribution, where  $a + b = 1$  and  $(\nu_S, \nu_F) = (b, a)$ .

In Section 2 we will give an explicit formula for the variance of an MBD. This was not given in [15], and only implicitly in [9] and [13]. By introducing the notion of ‘excentricity’ we can write down tractable formulae for the expectation and the variance of an MBD. For the application to the reactive transport model, we need a bit more, namely the variances conditioned on the state of the chain at time  $n$ . Expressions for these formulae will be computed in Section 3.

In Section 4 we will give a closed formula for the probability mass function  $f_n$  of  $K_n$ , and we study its shape. The probability mass function  $f_n$  was implicitly given in [8], [9], and [15], but the closed formula presented here is helpful to study its shape. Surprisingly, the shape can be unimodal, bimodal, or trimodal. We show in particular that, when  $a + b \geq 1$ , the probability mass function of  $K_n$  is unimodal, and that the probability mass function of  $K_n$  restricted to the interval  $[1, n - 1]$  is always unimodal.

In Section 5 we give formulae for the probability mass functions of  $K_n$ , conditional on the state at time  $n$ . Here again our interest arises from the fact that in the reactive transport model of [12] the authors considered the behaviour of the concentration of the aqueous part of a solute, which corresponds to conditioning at the state of the chain at time  $n$  (aqueous  $\sim$  success, adsorbed  $\sim$  failure).

### 2. The variance of the Markov binomial distribution

Let  $(\pi_S, \pi_F)$  be the stationary distribution of the chain  $\{Y_k, k \geq 1\}$ . We have

$$\pi_S = \frac{b}{a + b}, \quad \pi_F = \frac{a}{a + b}.$$

In fact, diagonalizing  $P$  yields, for  $n = 0, 1, 2, \dots$ ,

$$P^n = \begin{bmatrix} \pi_S & \pi_F \\ \pi_S & \pi_F \end{bmatrix} + \gamma^n \begin{bmatrix} \pi_F & -\pi_F \\ -\pi_S & \pi_S \end{bmatrix}, \tag{2.1}$$

where  $\gamma = 1 - a - b$  is the second largest eigenvalue of  $P$ . Note that, for  $1 \leq k \leq n$ ,

$$P_\nu(Y_k = S) = \nu_S P^{k-1}(S, S) + \nu_F P^{k-1}(F, S) = \pi_S \left( 1 - \left( 1 - \frac{\nu_S}{\pi_S} \right) \gamma^{k-1} \right),$$

and, similarly,

$$P_\nu(Y_k = F) = \nu_S P^{k-1}(S, F) + \nu_F P^{k-1}(F, F) = \pi_F \left( 1 - \left( 1 - \frac{\nu_F}{\pi_F} \right) \gamma^{k-1} \right).$$

It thus appears useful to define the *excentricities*  $\varepsilon_S$  and  $\varepsilon_F$  of an initial distribution  $\nu$  by

$$\varepsilon_\tau := \varepsilon_\tau(\nu) = 1 - \frac{\nu_\tau}{\pi_\tau} \quad \text{for } \tau \in \{S, F\}.$$

Both quantities measure the deviation of the initial distribution  $\nu$  from the stationary distribution  $\pi$ . Using them, we can rewrite  $P_\nu(Y_k = S)$  and  $P_\nu(Y_k = F)$  as

$$P_\nu(Y_k = S) = \pi_S(1 - \varepsilon_S \gamma^{k-1}), \quad P_\nu(Y_k = F) = \pi_F(1 - \varepsilon_F \gamma^{k-1}). \tag{2.2}$$

Moreover, the expectation of  $K_n$  is given by (note that  $\gamma < 1$  since  $a + b > 0$ )

$$E_\nu[K_n] = \sum_{k=1}^n E_\nu[\mathbf{1}_{\{Y_k=S\}}] = \sum_{k=1}^n P_\nu(Y_k = S) = \pi_S \left( n - \varepsilon_S \frac{1 - \gamma^n}{1 - \gamma} \right). \tag{2.3}$$

The expectation of  $K_n$  is particularly simple if we start in the equilibrium distribution, since in this case  $\varepsilon_S = 0$ .

Obtaining  $\text{var}_\nu(K_n)$  is more involved, because of correlations. A closed formula for the variance in the following proposition has been obtained by several authors (see, for example [2] and [13]). Here it is particularly compact by our use of excentricities.

**Proposition 2.1.** *For any  $\text{Bin}(n, a, b, \nu)$  distributed random variable  $K_n$ , we have*

$$\begin{aligned} \text{var}_\nu(K_n) = \pi_S \left\{ n \frac{\pi_F(1 + \gamma)}{1 - \gamma} + \frac{\gamma(\varepsilon_S(\pi_S - \pi_F) - 2\pi_F) - \varepsilon_S(\pi_F - \nu_S)}{(1 - \gamma)^2} + n\gamma^n \frac{2\varepsilon_S(\pi_F - \pi_S)}{1 - \gamma} \right. \\ \left. + \gamma^n \left( \frac{\varepsilon_S(\pi_S - \pi_F)}{1 - \gamma} + 2 \frac{\gamma\pi_F + \varepsilon_S(\pi_F - \nu_S)}{(1 - \gamma)^2} \right) - \gamma^{2n} \frac{\pi_S \varepsilon_S^2}{(1 - \gamma)^2} \right\}. \end{aligned}$$

*Proof.* Since  $\text{var}_\nu(K_n) = E_\nu[K_n^2] - (E_\nu[K_n])^2$ , using (2.3), it suffices to calculate

$$\begin{aligned} E_\nu[K_n^2] &= E_\nu \left[ \left( \sum_{k=1}^n \mathbf{1}_{\{Y_k=S\}} \right)^2 \right] \\ &= \sum_{k=1}^n P_\nu(Y_k = S) + 2 \sum_{1 \leq i < j \leq n} P_\nu(Y_i = S, Y_j = S) \\ &= E_\nu[K_n] + 2 \sum_{1 \leq i < j \leq n} P_\nu(Y_i = S, Y_j = S). \end{aligned}$$

Thus, we only need to calculate

$$\begin{aligned} P_\nu(Y_i = S, Y_j = S) &= P_\nu(Y_j = S \mid Y_i = S) P_\nu(Y_i = S) \\ &= (\pi_S + \pi_F \gamma^{j-i}) \pi_S (1 - \varepsilon_S \gamma^{i-1}) \\ &= \pi_S (\pi_S + \pi_F \gamma^{j-i} - \varepsilon_S \pi_S \gamma^{i-1} - \varepsilon_S \pi_F \gamma^{j-1}), \end{aligned}$$

using (2.1) and (2.2). Performing the four summations we obtain

$$\begin{aligned} &2 \sum_{1 \leq i < j \leq n} P_\nu(Y_i = S, Y_j = S) \\ &= 2\pi_S \left\{ \pi_S \frac{n(n-1)}{2} + \pi_F \gamma \left( \frac{n}{1-\gamma} - \frac{1-\gamma^n}{(1-\gamma)^2} \right) - \varepsilon_S \pi_S \left( \frac{n}{1-\gamma} - \frac{1-\gamma^n}{(1-\gamma)^2} \right) \right. \\ &\quad \left. - \varepsilon_S \pi_F \left( \frac{-n\gamma^n}{1-\gamma} + \frac{\gamma(1-\gamma^n)}{(1-\gamma)^2} \right) \right\} \\ &= \pi_S \left\{ n(n-1)\pi_S + 2n \frac{\pi_F \gamma - \varepsilon_S \pi_S}{1-\gamma} + 2n\gamma^n \frac{\varepsilon_S \pi_F}{1-\gamma} + 2(1-\gamma^n) \frac{\varepsilon_S \pi_S - \pi_F \gamma(1+\varepsilon_S)}{(1-\gamma)^2} \right\}, \end{aligned}$$

which, combined with (2.3), completes the proof.

### 3. The conditional variance of the Markov binomial distribution

Here we are interested in the variance of  $K_n$  given the state of the chain at time  $n$ . Let  $K_n^\tau$  be the random variable  $K_n$  conditioned on  $Y_n = \tau \in \{S, F\}$ . For completeness, we will first give the corresponding means  $E_v[K_n^S]$  and  $E_v[K_n^F]$ , which were also given in [8], [9], and [15]. Using (2.1) and (2.2), we obtain

$$\begin{aligned}
 E_v[K_n^S] &= E_v[K_n \mid Y_n = S] \\
 &= \sum_{k=1}^n P_v(Y_k = S \mid Y_n = S) \\
 &= \sum_{k=1}^n \frac{P_v(Y_n = S \mid Y_k = S) P_v(Y_k = S)}{P_v(Y_n = S)} \\
 &= \frac{\sum_{k=1}^n P^{n-k}(S, S) P_v(Y_k = S)}{\pi_S(1 - \varepsilon_S \gamma^{n-1})} \\
 &= \frac{\sum_{k=1}^n (\pi_S + \pi_F \gamma^{n-k}) \pi_S(1 - \varepsilon_S \gamma^{k-1})}{\pi_S(1 - \varepsilon_S \gamma^{n-1})} \\
 &= n \frac{\pi_S - \varepsilon_S \pi_F \gamma^{n-1}}{1 - \varepsilon_S \gamma^{n-1}} + \frac{(\pi_F - \varepsilon_S \pi_S)(1 - \gamma^n)}{(1 - \gamma)(1 - \varepsilon_S \gamma^{n-1})}, \tag{3.1}
 \end{aligned}$$

and, similarly,

$$E_v[K_n^F] = n \frac{\pi_S - \varepsilon_F \pi_F \gamma^{n-1}}{1 - \varepsilon_F \gamma^{n-1}} + \frac{(\varepsilon_F \pi_F - \pi_S)(1 - \gamma^n)}{(1 - \gamma)(1 - \varepsilon_F \gamma^{n-1})}. \tag{3.2}$$

**Proposition 3.1.** *The variances of  $K_n^\tau$ , a  $\text{Bin}(n, a, b, v)$  distributed random variable  $K_n$  conditioned on  $Y_n = \tau \in \{S, F\}$ , are given by*

$$\begin{aligned}
 \text{var}_v(K_n^S) &= n^2 \frac{\pi_S^2 - \varepsilon_S \pi_F^2 \gamma^{n-1}}{1 - \varepsilon_S \gamma^{n-1}} - \left( n \frac{\pi_S - \varepsilon_S \pi_F \gamma^{n-1}}{1 - \varepsilon_S \gamma^{n-1}} + \frac{(\pi_F - \varepsilon_S \pi_S)(1 - \gamma^n)}{(1 - \gamma)(1 - \varepsilon_S \gamma^{n-1})} \right)^2 \\
 &\quad - n \left( \frac{\pi_F \pi_S (1 + 3\varepsilon_S \gamma^{n-1})}{1 - \varepsilon_S \gamma^{n-1}} + 2 \frac{\varepsilon_S \pi_S^2 + \pi_F^2 \gamma^n - 2\pi_F \pi_S (1 + \varepsilon_S \gamma^{n-1})}{(1 - \gamma)(1 - \varepsilon_S \gamma^{n-1})} \right) \\
 &\quad + (1 - \gamma^n) \left( \frac{\pi_F \pi_S (4 + \varepsilon_S) - (\pi_F + \varepsilon_S \pi_S^2)}{(1 - \gamma)(1 - \varepsilon_S \gamma^{n-1})} + 2 \frac{\varepsilon_S \pi_S^2 + \pi_F^2 - 2\pi_F \pi_S (1 + \varepsilon_S)}{(1 - \gamma)^2 (1 - \varepsilon_S \gamma^{n-1})} \right),
 \end{aligned}$$

and

$$\begin{aligned}
 \text{var}_v(K_n^F) &= n^2 \frac{\pi_S^2 - \varepsilon_F \pi_F^2 \gamma^{n-1}}{1 - \varepsilon_F \gamma^{n-1}} - \left( n \frac{\pi_S - \varepsilon_F \pi_F \gamma^{n-1}}{1 - \varepsilon_F \gamma^{n-1}} + \frac{(\varepsilon_F \pi_F - \pi_S)(1 - \gamma^n)}{(1 - \gamma)(1 - \varepsilon_F \gamma^{n-1})} \right)^2 \\
 &\quad - n \left( \frac{\pi_F \pi_S (1 + (2 + \varepsilon_F) \gamma^{n-1})}{1 - \varepsilon_F \gamma^{n-1}} + 2 \frac{\pi_S^2 + \varepsilon_F \pi_F^2 \gamma^n - \pi_F \pi_S (1 + \varepsilon_F)(1 + \gamma^{n-1})}{(1 - \gamma)(1 - \varepsilon_F \gamma^{n-1})} \right) \\
 &\quad + (1 - \gamma^n) \left( \frac{\pi_F \pi_S (4 + \varepsilon_F) - (\pi_S + \varepsilon_F \pi_F^2)}{(1 - \gamma)(1 - \varepsilon_F \gamma^{n-1})} + 2 \frac{\pi_S^2 + \varepsilon_F \pi_F^2 - 2\pi_F \pi_S (1 + \varepsilon_F)}{(1 - \gamma)^2 (1 - \varepsilon_F \gamma^{n-1})} \right).
 \end{aligned}$$

*Proof.* Since the calculation of  $\text{var}_v(K_n^F)$  is similar to  $\text{var}_v(K_n^S)$ , we only deal with  $\text{var}_v(K_n^S)$ . Note that  $\text{var}_v(K_n^S) = E_v[(K_n^S)^2] - (E_v[K_n^S])^2$ . Using (3.1), it suffices to calculate

$$\begin{aligned} E_v[(K_n^S)^2] &= E_v\left[\left(\sum_{k=1}^n \mathbf{1}_{\{Y_k=S\}}\right)^2 \mid Y_n = S\right] \\ &= \sum_{k=1}^n P_v(Y_k = S \mid Y_n = S) + 2 \sum_{1 \leq i < j \leq n} P_v(Y_i = S, Y_j = S \mid Y_n = S) \\ &= E_S[K_n] + 2 \sum_{1 \leq i < j \leq n} P_v(Y_i = S, Y_j = S \mid Y_n = S). \end{aligned}$$

It follows from (2.1) and (2.2) that

$$\begin{aligned} &P_v(Y_i = S, Y_j = S \mid Y_n = S) \\ &= \frac{P_v(Y_i = S) P_v(Y_j = S, Y_n = S \mid Y_i = S)}{P_v(Y_n = S)} \\ &= \frac{P_v(Y_i = S) P^{j-i}(S, S) P^{n-j}(S, S)}{P_v(Y_n = S)} \\ &= \frac{(1 - \varepsilon_S \gamma^{i-1})(\pi_S + \pi_F \gamma^{j-i})(\pi_S + \pi_F \gamma^{n-j})}{1 - \varepsilon_S \gamma^{n-1}} \\ &= \frac{\pi_F^2 \gamma^{n-i} - \pi_S^2 \varepsilon_S \gamma^{i-1}}{1 - \varepsilon_S \gamma^{n-1}} + \frac{\pi_S^2 - \pi_F^2 \varepsilon_S \gamma^{n-1}}{1 - \varepsilon_S \gamma^{n-1}} + \frac{\pi_F \pi_S (\gamma^{n-j} - \varepsilon_S \gamma^{j-1})}{1 - \varepsilon_S \gamma^{n-1}} \\ &\quad + \frac{\pi_F \pi_S (\gamma^{j-i} - \varepsilon_S \gamma^{n-1-(j-i)})}{1 - \varepsilon_S \gamma^{n-1}}. \end{aligned}$$

Performing the eight summations in the above equation we obtain

$$\begin{aligned} &2 \sum_{1 \leq i < j \leq n} P_v(Y_i = S, Y_j = S \mid Y_n = S) \\ &= \left(2(1 - \gamma^n) \frac{\pi_S^2 \varepsilon_S + \pi_F^2 \gamma}{(1 - \varepsilon_S \gamma^{n-1})(1 - \gamma)^2} - 2n \frac{\pi_S^2 \varepsilon_S + \pi_F^2 \gamma^n}{(1 - \varepsilon_S \gamma^{n-1})(1 - \gamma)}\right) \\ &\quad + \frac{n(n-1)(\pi_S^2 - \pi_F^2 \varepsilon_S \gamma^{n-1})}{1 - \varepsilon_S \gamma^{n-1}} + \frac{2\pi_S \pi_F}{1 - \varepsilon_S \gamma^{n-1}} \left(n \frac{1 + \varepsilon_S \gamma^n}{1 - \gamma} - (1 - \gamma^n) \frac{1 + \varepsilon_S \gamma}{(1 - \gamma)^2}\right) \\ &\quad + \frac{2\pi_S \pi_F}{1 - \varepsilon_S \gamma^{n-1}} \left(n \frac{\gamma + \varepsilon_S \gamma^{n-1}}{1 - \gamma} - (1 - \gamma^n) \frac{\gamma + \varepsilon_S}{(1 - \gamma)^2}\right) \\ &= n(n-1) \frac{\pi_S^2 - \pi_F^2 \varepsilon_S \gamma^{n-1}}{1 - \varepsilon_S \gamma^{n-1}} + 2n \frac{\pi_S \pi_F (1 + \gamma)(1 + \varepsilon_S \gamma^{n-1}) - (\pi_S^2 \varepsilon_S + \pi_F^2 \gamma^n)}{(1 - \varepsilon_S \gamma^{n-1})(1 - \gamma)} \\ &\quad + 2(1 - \gamma^n) \frac{\pi_S^2 \varepsilon_S + \pi_F^2 \gamma - \pi_S \pi_F (1 + \gamma)(1 + \varepsilon_S)}{(1 - \varepsilon_S \gamma^{n-1})(1 - \gamma)^2}, \end{aligned}$$

which, combined with (3.1), yields the expression for  $\text{var}_v(K_n^S)$ .

For the special initial distributions (0, 1) and (1, 0), we have the excentricities  $\varepsilon_S((0, 1)) = 1 = \varepsilon_F((1, 0))$ . Substituting them into (3.1), (3.2), and Proposition 3.1 we obtain

$$E_F[K_n^S] = E_S[K_n^F], \quad \text{var}_F(K_n^S) = \text{var}_S(K_n^F),$$

where

$$E_F := E_{(0,1)}, \quad E_S := E_{(1,0)}, \quad \text{var}_F := \text{var}_{(0,1)}, \quad \text{var}_S := \text{var}_{(1,0)}.$$

More generally, we have the following.

**Proposition 3.2.** *For any  $\text{Bin}(n, a, b, \nu)$  distributed random variable  $K_n$  and any positive integer  $m$ , the  $m$ th moment of  $K_n^S$  conditioned on  $Y_1 = F$  is equal to the  $m$ th moment of  $K_n^F$  conditioned on  $Y_1 = S$ , i.e. for  $m = 1, 2, \dots$ ,*

$$E_F[(K_n^S)^m] = E_S[(K_n^F)^m].$$

*Proof.* Note that, for  $m \leq n$ ,

$$\begin{aligned} K_n^m &= \left( \sum_{k=1}^n \mathbf{1}_{\{Y_k=S\}} \right)^m \\ &= C_1 \sum_{k=1}^n \mathbf{1}_{\{Y_k=S\}} + C_2 \sum_{i_1 < i_2} \mathbf{1}_{\{Y_{i_1}=S, Y_{i_2}=S\}} + \dots \\ &\quad + C_m \sum_{i_1 < i_2 < \dots < i_m} \mathbf{1}_{\{Y_{i_1}=S, Y_{i_2}=S, \dots, Y_{i_m}=S\}}, \end{aligned}$$

where the  $C_i$ s are constants related to  $n$  and  $m$ . This implies that, for  $\tau \in \{S, F\}$ ,

$$\begin{aligned} E_{\bar{\tau}}[(K_n^{\tau})^m] &= C_1 \sum_{k=1}^n P_{\bar{\tau}}(Y_k = S \mid Y_n = \tau) + C_2 \sum_{i_1 < i_2} P_{\bar{\tau}}(Y_{i_1} = S, Y_{i_2} = S \mid Y_n = \tau) + \dots \\ &\quad + C_m \sum_{i_1 < i_2 < \dots < i_m} P_{\bar{\tau}}(Y_{i_1} = S, Y_{i_2} = S, \dots, Y_{i_m} = S \mid Y_n = \tau), \end{aligned}$$

where  $\bar{S} = F, \bar{F} = S$  and  $P_F := P_{(0,1)}, P_S := P_{(1,0)}$ . Thus, we only need to show that, for  $1 \leq i_1 < \dots < i_k \leq n$ ,

$$P_F(Y_{i_1} = S, \dots, Y_{i_k} = S \mid Y_n = S) = P_S(Y_{n-i_k+1} = S, \dots, Y_{n-i_1+1} = S \mid Y_n = F). \tag{3.3}$$

It is easy to see that both sides of (3.3) equal 0 if  $i_1 = 1$ . Now suppose that  $i_1 \geq 2$ . Since  $\{Y_k, k \geq 1\}$  is a homogeneous time reversible Markov chain, we have

$$\begin{aligned} &P_F(Y_{i_1} = S, \dots, Y_{i_k} = S \mid Y_n = S) \\ &= \frac{P_F(Y_n = S \mid Y_{i_k} = S) P_F(Y_{i_k} = S \mid Y_{i_{k-1}} = S) \dots P_F(Y_{i_2} = S \mid Y_{i_1} = S) P_F(Y_{i_1} = S)}{P_F(Y_n = S)} \\ &= \frac{P^{n-i_k}(S, S) P^{i_k-i_{k-1}}(S, S) \dots P^{i_2-i_1}(S, S) P^{i_1-1}(F, S)}{P^{n-1}(F, S)} \\ &= \frac{P^{n-i_k}(S, S) P^{i_k-i_{k-1}}(S, S) \dots P^{i_2-i_1}(S, S) \pi_S P^{i_1-1}(S, F) / \pi_F}{\pi_S P^{n-1}(S, F) / \pi_F} \\ &= P_S(Y_{n-i_k+1} = S, Y_{n-i_{k-1}+1} = S, \dots, Y_{n-i_1+1} = S \mid Y_n = F), \end{aligned}$$

which yields (3.3). Thus, the proposition is established for  $m \leq n$ . In a similar way, we can show that the proposition holds for all  $m > n$ .

### 4. The probability mass function of the Markov binomial distribution

For any  $\text{Bin}(n, a, b, \nu)$  distributed random variable  $K_n$ , we will give sufficient and necessary conditions for the probability mass function of  $K_n$  to be unimodal, bimodal, or trimodal. These three kinds of shape are mentioned by Viveros *et al.* [15] without any further explanation.

Given  $n \geq 1$ , let  $f_n$  be the probability mass function of  $K_n$ , i.e.

$$f_n(j) = P_\nu(K_n = j).$$

In particular,  $f_n(j) = 0$  if  $j < 0$  or  $j > n$ . By an easy computation,

$$\begin{aligned} f_{n+2}(j + 1) &= P_\nu(K_{n+1} = j + 1, Y_{n+1} = F)P(F, F) \\ &\quad + P_\nu(K_{n+1} = j, Y_{n+1} = F)P(F, S) \\ &\quad + P_\nu(K_{n+1} = j + 1, Y_{n+1} = S)P(S, F) \\ &\quad + P_\nu(K_{n+1} = j, Y_{n+1} = S)P(S, S) \\ &= f_{n+1}(j + 1)P(F, F) + P_\nu(K_{n+1} = j + 1, Y_{n+1} = S)(P(S, F) - P(F, F)) \\ &\quad + f_{n+1}(j)P(S, S) + P_\nu(K_{n+1} = j, Y_{n+1} = F)(P(F, S) - P(S, S)) \\ &= P(F, F)f_{n+1}(j + 1) + P(S, S)f_{n+1}(j) + (P(S, F) - P(F, F))f_n(j), \end{aligned}$$

where the last equality holds since

$$P(S, F) + P(S, S) = P(F, F) + P(F, S) = 1.$$

Substituting (1.1) into the above recursion equation yields, for  $n \geq 1$ ,

$$f_{n+2}(j + 1) = (1 - b)f_{n+1}(j + 1) + (1 - a)f_{n+1}(j) - (1 - a - b)f_n(j) \tag{4.1}$$

with initial conditions

$$\begin{aligned} f_1(0) &= \nu_F, & f_1(1) &= \nu_S, \\ f_2(0) &= \nu_F(1 - b), & f_2(1) &= \nu_Fb + \nu_Sa, & f_2(2) &= \nu_S(1 - a). \end{aligned} \tag{4.2}$$

In [8], [9], and [15] (implicit) expressions for the probability mass function of  $K_n$  are given, but the closed form presented here is more helpful to study its shape.

**Proposition 4.1.** *The probability mass function  $f_n$  of a  $\text{Bin}(n, a, b, \nu)$  distributed random variable  $K_n$  can be written as*

$$f_n(j) = \begin{cases} \nu_F(1 - b)^{n-1}, & j = 0, \\ (1 - b)^{n-j}(1 - a)^{j-1} \sum_{k=0}^{j-1} \binom{j-1}{k} \delta^k c_{j-1,k}(n), & 1 \leq j \leq n - 1, \\ \nu_S(1 - a)^{n-1}, & j = n, \\ 0, & \text{otherwise,} \end{cases}$$

where  $\delta = ab/((1 - a)(1 - b))$  and

$$c_{j,k}(n) = \nu_S \binom{n - 2 - j}{k - 1} + \frac{\nu_Sa + \nu_Fb}{1 - b} \binom{n - 2 - j}{k} + \frac{\nu_Fab}{(1 - b)^2} \binom{n - 2 - j}{k + 1}.$$

*Proof.* It is easy to see that the recursion equation (4.1) with initial conditions (4.2) has a unique solution. We only need to check that  $f_n$  presented in the proposition satisfies (4.1) and (4.2), and that the summation of  $f_n(j)$  from  $j = 0$  to  $n$  equals 1. It is easy to see that (4.1) holds for  $j < 0$  and  $j > n$ . Equation (4.1) holds for  $j = 0$  since

$$\begin{aligned} & (1 - b)f_{n+1}(1) + (1 - a)f_{n+1}(0) - (1 - a - b)f_n(0) \\ &= (1 - b)^{n+1}c_{0,0}(n + 1) + v_F(1 - a)(1 - b)^n - v_F(1 - a - b)(1 - b)^{n-1} \\ &= (1 - b)^{n-1}((1 - b)(v_Sa + v_Fb) + (n - 1)v_Fab + v_F(1 - a)(1 - b) \\ &\quad - v_F(1 - a - b)) \\ &= (1 - b)^{n-1}((1 - b)(v_Sa + v_Fb) + nv_Fab) \\ &= f_{n+2}(1). \end{aligned}$$

Similarly, (4.1) holds for  $j = n$ .

Suppose now that  $1 \leq j \leq n - 1$ . From simple properties of the binomial coefficients in  $c_{j,k}(n)$ , it follows that

$$c_{j-1,k}(n) = c_{j,k}(n + 1) = c_{j+1,k}(n + 2)$$

and

$$c_{j,k}(n + 2) = c_{j+1,k}(n + 2) + c_{j+1,k-1}(n + 2). \tag{4.3}$$

We write  $c_{j,k} := c_{j,k}(n + 2)$  for short. Thus,

$$\begin{aligned} & (1 - b)f_{n+1}(j + 1) + (1 - a)f_{n+1}(j) - (1 - a - b)f_n(j) \\ &= (1 - b)^{n+1-j}(1 - a)^j \sum_{k=0}^j \binom{j}{k} \delta^k c_{j+1,k} + (1 - b)^{n+1-j}(1 - a)^j \sum_{k=0}^j \binom{j-1}{k} \delta^k c_{j,k} \\ &\quad - (1 - \delta)(1 - b)^{n+1-j}(1 - a)^j \sum_{k=0}^j \binom{j-1}{k} \delta^k c_{j+1,k} \\ &= (1 - b)^{n+1-j}(1 - a)^j \left[ \sum_{k=0}^j \binom{j-1}{k} \delta^k c_{j+1,k} + \sum_{k=0}^j \binom{j-1}{k-1} \delta^k c_{j+1,k} \right. \\ &\quad \left. + \sum_{k=0}^j \binom{j-1}{k} \delta^k c_{j,k} - \sum_{k=0}^j \binom{j-1}{k} \delta^k c_{j+1,k} \right. \\ &\quad \left. + \sum_{k=0}^j \binom{j-1}{k-1} \delta^k c_{j+1,k-1} \right] \\ &= (1 - b)^{n+1-j}(1 - a)^j \left[ \sum_{k=0}^j \binom{j-1}{k-1} \delta^k (c_{j+1,k} + c_{j+1,k-1}) + \sum_{k=0}^j \binom{j-1}{k} \delta^k c_{j,k} \right] \\ &= (1 - b)^{n+1-j}(1 - a)^j \left[ \sum_{k=0}^j \binom{j-1}{k-1} \delta^k c_{j,k} + \sum_{k=0}^j \binom{j-1}{k} \delta^k c_{j,k} \right] \\ &= (1 - b)^{n+1-j}(1 - a)^j \sum_{k=0}^j \binom{j}{k} \delta^k c_{j,k} \\ &= f_{n+2}(j + 1). \end{aligned}$$



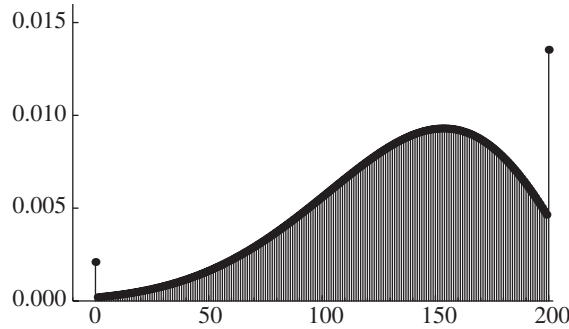


FIGURE 1: Probability mass function  $f_{200}$  of  $K_{200}$  with  $a = 0.01, b = 0.03,$  and  $\nu = (0.1, 0.9).$

Now we show by induction that  $\sum_{j=0}^n f_n(j) = 1$  for each  $n \geq 1$ . For  $n = 1, 2,$  we have  $f_1(0) + f_1(1) = \nu_F + \nu_S = 1$  and

$$f_2(0) + f_2(1) + f_2(2) = \nu_F(1 - b) + \nu_F b + \nu_S a + \nu_S(1 - a) = 1.$$

Suppose that  $f_n$  and  $f_{n+1}$  are probability mass functions. Then, by (4.1),

$$\begin{aligned} \sum_{j=0}^{n+2} f_{n+2}(j) &= (1 - a) \sum_{j=0}^{n+2} f_{n+1}(j) + (1 - b) \sum_{j=0}^{n+2} f_{n+1}(j - 1) \\ &\quad - (1 - a - b) \sum_{j=0}^{n+2} f_n(j - 1) \\ &= (1 - a) + (1 - b) - (1 - a - b) \\ &= 1. \end{aligned}$$

This completes the proof.

**Example 4.1.** Let  $n = 200, a = 0.01, b = 0.03,$  and  $\nu = (0.1, 0.9).$  By Proposition 4.1 we obtain the probability mass function of  $K_{200}$  shown in Figure 1. Apparently,  $f_{200}$  is trimodal.

A finite sequence of real numbers  $\{x_i\}_{i=0}^n$  is said to be *unimodal* if there exists an index  $0 \leq n^* \leq n,$  called a *mode* of the sequence, such that  $x_0 \leq x_1 \leq \dots \leq x_{n^*}$  and  $x_{n^*} \geq x_{n^*+1} \geq \dots \geq x_n.$  In particular, we call the sequence  $\{x_i\}_{i=0}^n$  *strictly unimodal* if all modes  $n^*$  satisfy  $0 < n^* < n.$  From the definition, it is easy to see that a monotonic sequence is unimodal.

A nonnegative sequence  $\{x_i\}_{i=0}^n$  is called *log-concave* (or *strictly log-concave*) if  $x_{i-1}x_{i+1} \leq x_i^2$  (or  $x_{i-1}x_{i+1} < x_i^2$ ) for all  $1 \leq i \leq n - 1.$  It is well known that the sequence  $\{x_i\}_{i=0}^n$  is log-concave if and only if  $x_{i_1-1}x_{i_2+1} \leq x_{i_1}x_{i_2}$  for all  $1 \leq i_1 \leq i_2 \leq n - 1.$  Moreover, log-concavity implies unimodality.

The definitions of unimodality and log-concavity can be extended naturally to infinite sequences.

**Proposition 4.2.** Let  $a + b \geq 1,$  and let  $f_n$  be the probability mass function of a Bin( $n, a, b, \nu$ ) distributed random variable  $K_n.$  Then the sequence  $\{f_n(j)\}_{j=0}^n$  is log-concave, and, hence, unimodal. Moreover, the mode  $n^*$  satisfies  $\lfloor E_\nu[K_n] \rfloor \leq n^* \leq \lceil E_\nu[K_n] \rceil.$

*Proof.* Let  $G_n$  be the generating function of  $K_n$ , i.e. for all real  $s$ ,

$$G_n(s) = E_v[s^{K_n}] = \sum_{j=0}^n f_n(j)s^j.$$

Without loss of generality, we suppose that  $0 < v_S < 1$ . Then (by Proposition 4.1)  $G_n$  has positive coefficients. It follows from the recursion equation (4.1) that

$$G_{n+2}(s) = ((1 - a)s + (1 - b))G_{n+1}(s) - (1 - a - b)sG_n(s).$$

Since  $a + b \geq 1$ , we find, by Corollary 2.4 of [10], that, for each  $n \geq 1$ , all 0s of  $G_n$  are real. Thus, the sequence  $\{f_n(j)\}_{j=0}^n$  is log-concave and, hence, unimodal with mode  $n^*$  between  $\lfloor E_v[K_n] \rfloor$  and  $\lceil E_v[K_n] \rceil$ .

We remark that it follows from Proposition 4.2 and Theorem 1 of [14] that, when  $a + b \geq 1$ , the distribution of  $K_n$  is a Bernoulli convolution, i.e. the distribution of a random variable  $\sum_{i=1}^n X_i$  with  $X_i$  independent Bernoulli random variables having parameters  $p_i \in \{1/(1 - s) : G_n(s) = 0\}$ .

When  $a + b < 1$ , Figure 1 suggests that the probability mass function  $\{f_n(j)\}_{j=0}^n$  is not unimodal. However, Figure 1 also suggests that  $\{f_n(j)\}_{j=1}^{n-1}$  is unimodal. We will indeed show in Proposition 4.3 that the sequence  $\{f_n(j)\}_{j=1}^{n-1}$  is log-concave, implying unimodality. In order to prove Proposition 4.3, it is helpful to use the following lemma which can be derived directly from Lemma 2.2 and Proposition 2.4 of [17]. To be more self-contained, we give a proof that uses simple properties of binomial coefficients and log-concave sequences.

**Lemma 4.1.** *For any positive integer  $j$  and a nonnegative log-concave sequence  $\{x_k\}_k$ , let  $d_{j,k} := \binom{j}{k}x_k$ . Then, for any  $0 \leq 2\ell \leq m \leq 2j$ ,*

$$\sum_{k=\ell}^{\lfloor m/2 \rfloor} D_{j,k}(m) \geq 0,$$

where, for  $k < m/2$ ,

$$D_{j,k}(m) = 2d_{j,k}d_{j,m-k} - d_{j-1,k}d_{j+1,m-k} - d_{j+1,k}d_{j-1,m-k},$$

and, for even  $m$  and  $k = m/2$ ,

$$D_{j,k}(m) = d_{j,k}^2 - d_{j-1,k}d_{j+1,k}.$$

*Proof.* Note that, for  $k < m/2$ ,

$$\begin{aligned} D_{j,k}(m) &= \left[ 2\binom{j}{k}\binom{j}{m-k} - \binom{j-1}{k}\binom{j+1}{m-k} - \binom{j+1}{k}\binom{j-1}{m-k} \right] x_k x_{m-k} \\ &= \left[ \binom{j-1}{k-1}\binom{j}{m-k} - \binom{j}{k-1}\binom{j-1}{m-k} \right] x_k x_{m-k} \\ &\quad - \left[ \binom{j-1}{k}\binom{j}{m-k-1} - \binom{j}{k}\binom{j-1}{m-k-1} \right] x_k x_{m-k}. \end{aligned}$$

For brevity, we only show the lemma for odd  $m$ . Let  $m = 2s + 1$ . Then, for  $0 \leq \ell \leq s < j$ ,

$$\begin{aligned} \sum_{k=\ell}^s D_{j,k}(m) &= \sum_{k=\ell-1}^{s-1} \left[ \binom{j-1}{k} \binom{j}{m-k-1} - \binom{j}{k} \binom{j-1}{m-k-1} \right] x_{k+1} x_{m-k-1} \\ &\quad - \sum_{k=\ell}^s \left[ \binom{j-1}{k} \binom{j}{m-k-1} - \binom{j}{k} \binom{j-1}{m-k-1} \right] x_k x_{m-k} \\ &= \sum_{k=\ell}^{s-1} \left[ \binom{j-1}{k} \binom{j}{m-k-1} - \binom{j}{k} \binom{j-1}{m-k-1} \right] (x_{k+1} x_{m-k-1} - x_k x_{m-k}) \\ &\quad + \left[ \binom{j-1}{\ell-1} \binom{j}{m-\ell} - \binom{j}{\ell-1} \binom{j-1}{m-\ell} \right] x_\ell x_{m-\ell} \\ &\geq 0, \end{aligned}$$

where the last inequality holds since

$$\binom{j-1}{k} \binom{j}{m-k-1} \geq \binom{j}{k} \binom{j-1}{m-k-1} \quad \text{for } k \leq s-1$$

and the sequence  $\{x_k\}_k$  is log-concave. This completes the proof.

Inspired by the proof of Theorem 3.10 of [17], we use Lemma 4.1 to show the log-concavity of an important class of sequences.

**Lemma 4.2.** *Let  $\delta > 0$ , and let  $\{c_{j,k}\}_{j,k \in \mathbb{Z}}$  be a nonnegative double sequence satisfying*

$$c_{j,k} = c_{j+1,k} + c_{j+1,k-1},$$

and  $c_{j,k} = 0$  for all  $j \in \mathbb{Z}$  and  $k \leq -2$ . Then the sequence

$$\left\{ \sum_{k=0}^j \binom{j}{k} \delta^k c_{j,k} \right\}_{j \geq 0}$$

is log-concave.

*Proof.* We fix  $j \geq 1$ . Let  $d_{j,k} := \binom{j}{k} \delta^k$ . We have to show that  $z_j^2 \geq z_{j-1} z_{j+1}$ , where

$$z_j := \sum_{k=0}^j \binom{j}{k} \delta^k c_{j,k} = \sum_{k=0}^j d_{j,k} c_{j,k}.$$

We use the shorthand notation  $v_k := c_{j+1,k}$ . Since  $c_{j,k} = c_{j+1,k} + c_{j+1,k-1}$ , this yields

$$\begin{aligned} z_{j+1} &= \sum_{k=0}^{j+1} d_{j+1,k} v_k, & z_j &= \sum_{k=0}^j d_{j,k} (v_k + v_{k-1}), \\ z_{j-1} &= \sum_{k=0}^{j-1} d_{j-1,k} (v_k + 2v_{k-1} + v_{k-2}). \end{aligned}$$

Note that  $v_k = c_{j+1,k} = 0$  for all  $j$  and  $k \leq -2$  and  $d_{j,k} = 0$  for  $k < 0$  or  $k > j$ , by the definition of  $\binom{j}{k}$ . Rewrite

$$z_{j+1} = \sum_{k=0}^{j+2} d_{j+1,k-1} v_{k-1}, \quad z_j = \sum_{k=0}^{j+2} (d_{j,k-1} + d_{j,k}) v_{k-1},$$

$$z_{j-1} = \sum_{k=0}^{j+2} (d_{j-1,k-1} + 2d_{j-1,k} + d_{j-1,k+1}) v_{k-1}.$$

Then  $z_j^2 - z_{j-1}z_{j+1}$  can be rewritten in a quadratic form of  $j+3$  variables  $v_{-1}, v_0, v_1, \dots, v_{j+1}$ :

$$z_j^2 - z_{j-1}z_{j+1} = \sum_{m=0}^{2(j+2)} \sum_{k=0}^{\lfloor m/2 \rfloor} e_{j,k}(m) v_{k-1} v_{m-k-1},$$

where

$$e_{j,k}(m) = 2(d_{j,k-1} + d_{j,k})(d_{j,m-k-1} + d_{j,m-k}) - (d_{j-1,k-1} + 2d_{j-1,k} + d_{j-1,k+1})d_{j+1,m-k-1} - d_{j+1,k-1}(d_{j-1,m-k-1} + 2d_{j-1,m-k} + d_{j-1,m-k+1}).$$

Since the  $v_k$ s are all nonnegative, it suffices to show that  $\sum_{k=0}^{\lfloor m/2 \rfloor} e_{j,k}(m) \geq 0$  for all  $0 \leq m \leq 2(j+2)$ . Rewrite

$$e_{j,k}(m) = P_k + 2Q_k + R_k,$$

where

$$P_k = 2d_{j,k-1}d_{j,m-k-1} - d_{j-1,k-1}d_{j+1,m-k-1} - d_{j+1,k-1}d_{j-1,m-k-1},$$

$$Q_k = d_{j,k-1}d_{j,m-k} + d_{j,k}d_{j,m-k-1} - d_{j-1,k}d_{j+1,m-k-1} - d_{j+1,k-1}d_{j-1,m-k},$$

$$R_k = 2d_{j,k}d_{j,m-k} - d_{j-1,k+1}d_{j+1,m-k-1} - d_{j+1,k-1}d_{j-1,m-k+1}.$$

Then we only need to show that

$$\sum_{k=0}^{\lfloor m/2 \rfloor} P_k \geq 0, \quad \sum_{k=0}^{\lfloor m/2 \rfloor} Q_k \geq 0, \quad \sum_{k=0}^{\lfloor m/2 \rfloor} R_k \geq 0.$$

For brevity, we show this only for the case when  $m$  is odd. For even  $m$ , the proof is very similar, but somewhat longer. Let  $m = 2s + 1$ . It follows from Lemma 4.1 that

$$\sum_{k=0}^s P_k = \sum_{k=0}^s D_{j,k-1}(m-2) = \sum_{k=0}^{s-1} D_{j,k}(m-2) \geq 0,$$

where the second equality holds since  $D_{j,k}(m-2) = 0$  for  $k < 0$ . Recalling from Lemma 4.1 that  $D_{j,s}(m-1) = d_{j,s}^2 - d_{j-1,s}d_{j+1,s}$ , we also have

$$\sum_{k=0}^s Q_k = \sum_{k=-1}^{s-1} (d_{j,k}d_{j,m-k-1} - d_{j+1,k}d_{j-1,m-k-1}) + \sum_{k=0}^s (d_{j,k}d_{j,m-k-1} - d_{j-1,k}d_{j+1,m-k-1})$$

$$= \sum_{k=0}^s D_{j,k}(m-1)$$

$$\geq 0.$$

Moreover,

$$\begin{aligned} \sum_{k=0}^s R_k &= 2 \sum_{k=0}^s d_{j,k}d_{j,m-k} - \sum_{k=1}^{s+1} d_{j-1,k}d_{j+1,m-k} - \sum_{k=-1}^{s-1} d_{j+1,k}d_{j-1,m-k} \\ &= \sum_{k=0}^s D_{j,k}(m) + d_{j-1,0}d_{j+1,m} \\ &\geq \sum_{k=0}^s D_{j,k}(m) \\ &\geq 0. \end{aligned}$$

This completes the proof.

**Proposition 4.3.** *For any Bin( $n, a, b, v$ ) distributed random variable  $K_n$ , let  $f_n$  be its probability mass function. Then the sequence  $\{f_n(j)\}_{j=1}^{n-1}$  is log-concave.*

*Proof.* According to Proposition 4.1 we have, for  $1 \leq j \leq n - 1$ ,

$$f_n(j) = (1 - b)^{n-j}(1 - a)^{j-1} \sum_{k=0}^{j-1} \binom{j-1}{k} \delta^k c_{j-1,k},$$

where  $\delta > 0$ , and the double sequence  $\{c_{j,k}\}_{j,k \in \mathbb{Z}}$  satisfies the recursion equation  $c_{j,k} = c_{j+1,k} + c_{j+1,k-1}$  (cf. (4.3)), and  $c_{j,k} = 0$  for  $k \leq -2$ . It follows from Lemma 4.2 that the sequence  $\{f_n(j)\}_{j=1}^{n-1}$  is log-concave.

In fact, we can show by sharpening the proof of Lemma 4.2 that  $\{f_n(j)\}_{j=1}^{n-1}$  is strictly log-concave, i.e.  $f_n(j)^2 > f_n(j - 1)f_n(j + 1)$  for  $j = 2, \dots, n - 2$ . Proposition 4.3 implies that the shape of the probability mass function of  $K_n$ , which can be unimodal, bimodal, or trimodal, is determined by the following six values:  $f_n(0), f_n(1), f_n(2), f_n(n - 2), f_n(n - 1)$ , and  $f_n(n)$ .

**Theorem 4.1.** *For any Bin( $n, a, b, v$ ) distributed random variable  $K_n$ , let  $f_n$  be its probability mass function. Then  $f_n$  is unimodal, except that*

- $f_n$  is bimodal with one peak on the left if and only if  $f_n(0) > f_n(1) \leq f_n(2)$  and either  $f_n(n - 1) \geq f_n(n)$  or  $f_n(n - 2) < f_n(n - 1) < f_n(n)$ ;
- $f_n$  is bimodal with one peak on the right if and only if  $f_n(n - 2) \geq f_n(n - 1) < f_n(n)$  and either  $f_n(0) \leq f_n(1)$  or  $f_n(0) > f_n(1) > f_n(2)$ ;
- $f_n$  is trimodal if and only if  $f_n(0) > f_n(1) \leq f_n(2)$  and  $f_n(n - 2) \geq f_n(n - 1) < f_n(n)$ .

**Example 4.2.** We consider the special case  $v = \pi = (b/(a + b), a/(a + b))$  and  $n = 50$ . It follows from Proposition 4.1 that

$$f_n(0) = (1 - b)^{n-1} \frac{a}{a + b}, \quad f_n(1) = \frac{(1 - b)^{n-2} ab}{a + b} \left( 2 + (n - 2) \frac{a}{1 - b} \right),$$

and

$$\begin{aligned} f_n(2) &= \frac{(1 - b)^{n-3} ab}{a + b} \left\{ (1 - a) \left( 2 + (n - 3) \frac{a}{1 - b} \right) \right. \\ &\quad \left. + b \left( 1 + (n - 3) \frac{2a}{1 - b} + \left( \frac{a}{1 - b} \right)^2 \binom{n - 3}{2} \right) \right\}. \end{aligned}$$

In a similar way we obtain the formulae for  $f_n(n - 2)$ ,  $f_n(n - 1)$ , and  $f_n(n)$ . Figure 2 is obtained via Theorem 4.1. For some examples of probability mass functions in this class, see Figure 3.

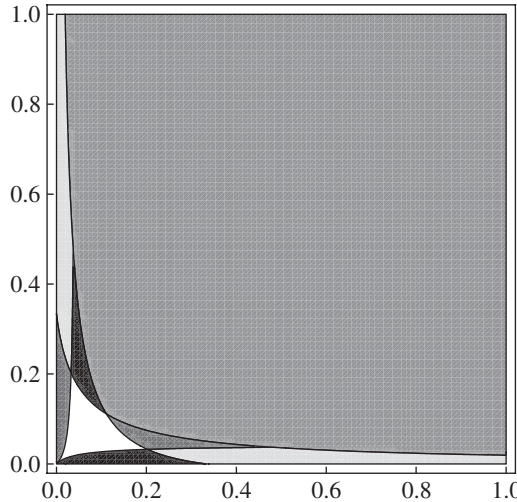


FIGURE 2: (i) When  $(a, b)$  is in the grey region,  $f_{50}$  is strictly unimodal; when  $(a, b)$  is in the lower light-grey region,  $f_{50}$  is decreasing; when  $(a, b)$  is in the upper light-grey region,  $f_{50}$  is increasing. (ii) When  $(a, b)$  is in the dark-grey region,  $f_{50}$  is bimodal with one peak on the left; when  $(a, b)$  is in the black region,  $f_{50}$  is bimodal with one peak on the right. (iii) When  $(a, b)$  is in the white region,  $f_{50}$  is trimodal.

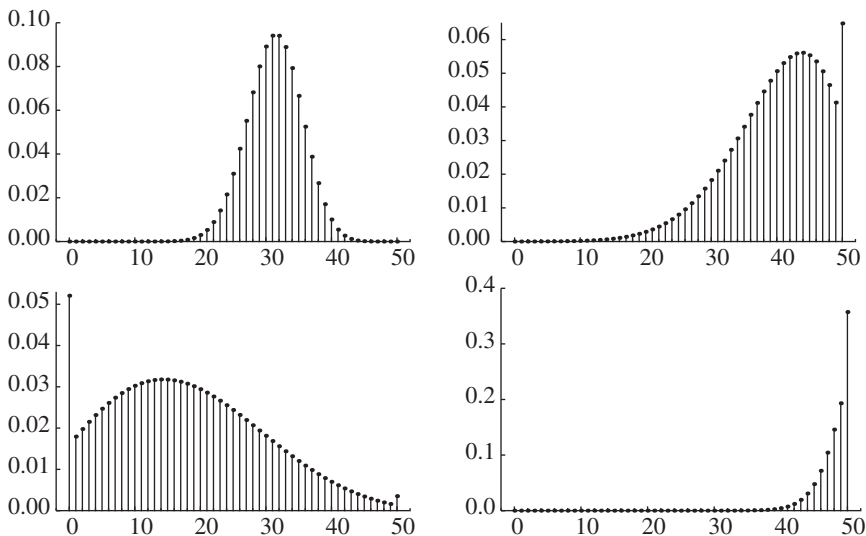


FIGURE 3: Probability mass function of  $K_{50}$  with  $\nu = \pi$ . In the upper-left graph  $a = 0.3$  and  $b = 0.5$ ; in the upper-right graph  $a = 0.05$  and  $b = 0.2$ ; in the lower-left graph  $a = 0.09$  and  $b = 0.05$ ; and in the lower-right graph  $a = 0.02$  and  $b = 0.5$ .

### 5. The conditional probability mass functions

For any  $\text{Bin}(n, a, b, \nu)$  distributed random variable  $K_n$ , let  $f_n^\tau$  be the probability mass function of  $K_n^\tau$  with  $\tau \in \{S, F\}$ , i.e.

$$f_n^\tau(j) = P_\nu(K_n^\tau = j) = P_\nu(K_n = j \mid Y_n = \tau).$$

In order to deal with  $f_n^\tau$ , it is simpler to deal with the partial probability mass functions

$$\hat{f}_n^\tau(j) = P_\nu(K_n = j, Y_n = \tau) = f_n^\tau(j) P_\nu(Y_n = \tau).$$

Since  $\hat{f}_n^F = f_n - \hat{f}_n^S$ , we only deal with  $\hat{f}_n^S$ . It is easy to obtain the recursion equation

$$\hat{f}_{n+2}^S(j + 1) = (1 - b)\hat{f}_{n+1}^S(j + 1) + (1 - a)\hat{f}_{n+1}^S(j) - (1 - a - b)\hat{f}_n^S(j),$$

with initial conditions

$$\begin{aligned} \hat{f}_1^S(0) &= 0, & \hat{f}_1^S(1) &= \nu_S, \\ \hat{f}_2^S(0) &= 0, & \hat{f}_2^S(1) &= \nu_F b, & \hat{f}_2^S(2) &= \nu_S(1 - a). \end{aligned}$$

Then we obtain the following proposition in a similar way as Proposition 4.1.

**Proposition 5.1.** *The partial probability mass function  $\hat{f}_n^S$  of a  $\text{Bin}(n, a, b, \nu)$  distributed random variable  $K_n$  can be written as*

$$\hat{f}_n^S(j) = \begin{cases} (1 - b)^{n-j} (1 - a)^{j-1} \sum_{k=0}^{j-1} \binom{j-1}{k} \delta^k c_{j-1,k}^S(n), & 1 \leq j \leq n - 1, \\ \nu_S (1 - a)^{n-1}, & j = n, \\ 0, & \text{otherwise,} \end{cases}$$

where  $\delta = ab/((1 - a)(1 - b))$  and

$$c_{j,k}^S(n) = \nu_S \binom{n - 2 - j}{k - 1} + \frac{\nu_F b}{1 - b} \binom{n - 2 - j}{k}.$$

From Lemma 4.2, it follows that the sequence  $\{\hat{f}_n^S(j)\}_{j=0}^{n-1}$  is log-concave, and, hence,  $\{f_n^S(j)\}_{j=0}^{n-1}$  is log-concave. Thus, in contrast to  $f_n$ ,  $f_n^S$  cannot have a trimodal shape. The unimodal or bimodal (with one peak on the right) shape of  $f_n^S$  depends on the values of  $f_n^S(j)$  for  $j = n - 2, n - 1, n$ .

Similarly, the shape of  $f_n^F$  can only be unimodal or bimodal (with one peak on the left) depending on the values of  $f_n^F(j)$  for  $j = 0, 1, 2$ .

### Acknowledgements

We would like to thank an anonymous referee for pointing out additional references and for the remark after the proof of Proposition 4.2. The second author was partially supported by the National Natural Science Foundation of China, under grant 10971069, and Shanghai Education Committee Project 11ZZ41.

## References

- [1] ČEKANA VIČIUS, V. AND ROOS, B. (2007). Binomial approximation to the Markov binomial distribution. *Acta Appl. Math.* **96**, 137–146.
- [2] ČEKANA VIČIUS, V. AND ROOS, B. (2009). Poisson type approximations for the Markov binomial distribution. *Stoch. Process. Appl.* **119**, 190–207.
- [3] DAUGMAN, J. (2003). The importance of being random: statistical principles of iris recognition. *Pattern Recognition* **36**, 279–291.
- [4] DEKING, M. AND HENSBERGEN, A. (2009). A problem with the assessment of an iris identification system. *SIAM Rev.* **51**, 417–422.
- [5] DEKING, M. AND KONG, D. (2011). A simple stochastic kinetic transport model. Preprint. Available at <http://arxiv.org/abs/1106.2912v1>.
- [6] DOBRUŠIN, R. L. (1953). Limit theorems for a Markov chain of two states. *Izvestiya Akad. Nauk SSSR. Ser. Mat.* **17**, 291–330 (in Russian). English translation: *Select. Transl. Math. Statist. Prob.* **1** (1961), pp. 97–134.
- [7] EDWARDS, A. W. F. (1960). The meaning of binomial distribution. *Nature* **186**, 1074.
- [8] GABRIEL, K. R. (1959). The distribution of the number of successes in a sequence of dependent trials. *Biometrika* **46**, 454–460.
- [9] HELGERT, H. J. (1970). On sums of random variables defined on a two-state Markov chain. *J. Appl. Prob.* **7**, 761–765.
- [10] LIU, L. L. AND WANG, Y. (2007). A unified approach to polynomial sequences with only real zeros. *Adv. Appl. Math.* **38**, 542–560.
- [11] MARKOV, A. A. (1924). *Probability Theory*, 4th edn. Moscow (in Russian).
- [12] MICHALAK, A. A. M. AND KITANIDIS, P. K. (2000). Macroscopic behavior and random-walk particle tracking of kinetically sorbing solutes. *Water Resources Res.* **36**, 2133–2146.
- [13] OMEY, E., SANTOS, J. AND VAN GULCK, S. (2008). A Markov-binomial distribution. *Appl. Anal. Discrete Math.* **2**, 38–50.
- [14] PITMAN, J. (1997). Probabilistic bounds on the coefficients of polynomials with only real zeros. *J. Combinatorial Theory A* **77**, 279–303.
- [15] VIVEROS, R., BALASUBRAMANIAN, K. AND BALAKRISHNAN, N. (1994). Binomial and negative binomial analogues under correlated Bernoulli trials. *Amer. Statistician* **48**, 243–247.
- [16] WANG, Y. H. (1981). On the limit of the Markov binomial distribution. *J. Appl. Prob.* **18**, 937–942.
- [17] WANG, Y. AND YEH, Y.-N. (2007). Log-concavity and LC-positivity. *J. Combinatorial Theory A* **114**, 195–210.