A generating function approach to the k-riffle shuffled, no-feedback, card guessing game

Tipaluck Krityakierne[1,3]   Thotsaporn Aek Thanatipanonda[2]

[1]Department of Mathematics, Faculty of Science, Mahidol University, Bangkok, Thailand
[2]Science Division, Mahidol University International College, Nakhon Pathom, Thailand
[3]Centre of Excellence in Mathematics, MHESI, Bangkok 10400, Thailand
Email: tipaluck.kri@mahidol.edu, thotsaporn.tha@mahidol.edu

Abstract

In this work, we introduce a new framework for analyzing a no-feedback, k-shuffled card guessing game. We show that the random variable representing the number on the i-th card follows a mixture distribution, allowing us to decompose its probability mass function into simpler, more manageable components. Building on this, we derive a closed-form expression for the probability generating function of each component, which is represented as a product of polynomials with a well-defined structure.

Keywords: card guessing; no feedback; discrete probability; generating function.

1 Preliminaries

Over the past few decades, the Card Guessing Game has been the subject of numerous studies, focusing on aspects such as optimal strategies, worst-case scenarios, and the distribution of correct guesses. One well-known feature of the game is the use of "riffle shuffles," commonly seen in casinos and poker games, where the deck is shuffled multiple times to achieve randomness. The famous result of Bayer and Diaconis [1] demonstrated that after just seven riffle shuffles, the positions of all 52 cards are nearly equally likely to be randomized.

The game comes in several variants, with or without feedback from the dealer. In the complete feedback version, similar to blackjack, the dealer reveals each card after a guess, allowing the player to adjust their strategy. In the no-feedback version, the focus of this work, the player must guess all the cards in advance, with the option to repeat guesses across different positions to maximize the probability of a correct guess.

In summary, the game considered in this work proceeds as follows:

Objective:
`Earn as many points as possible by correctly guessing the numbers on the cards.`

Rule:
- The game begins with a deck of n cards, initially ordered as $1, 2, 3, \ldots, n$.
- The dealer performs k-time riffle shuffles on the deck.
- The player then makes a series of guesses for the number on each card.
- No feedback is given after each guess.
- The game continues until the entire deck has been guessed.

Related work and our contribution

In 1998, Ciucu [2] published "No-feedback card guessing for dovetail shuffles," presenting the optimal guessing strategy for n cards (n even) after k riffle shuffles, when $k > 2\log_2(n)$. The analysis was based on the eigenvalues and eigenvectors of the transition matrix. Using the optimal strategy derived therein, the moments (and all higher moments) of the one-shuffle no-feedback card guessing game were established in [3].

In contrast, this work introduces a novel generating function framework for the no-feedback card guessing game, applicable to any number of cards n and any number of shuffles k. Our contributions are as follows:

- For any values of n and k, we show that the probability mass function (p.m.f.) of the random variable J_i, representing the card number at the i-th position in the deck, can be decomposed into more tractable p.m.f. components (Propositions 1 and 2).
- Building on this, we derive a simple closed-form expression for the probability generating function (p.g.f.) of each component in the mixture model, which can be interpreted as the product of polynomials that exhibit a clear and systematic pattern (Theorem 3). Finally, we provide the explicit formula for the p.g.f. of the random variable J_i in Corollary 5.

Example of a single riffle shuffle ($k = 1$)

Consider a single riffle shuffle with a deck of n cards. Split the deck into two piles (which may have zero cards in one pile), and interleave the cards in any possible way. *Can you determine all the possible permutations generated?*

For instance, when $n = 1$, $\{1\}$ is the only possible permutation, but with multiplicity 2, generated from the piles $(\{1\}, \{\})$ or $(\{\}, \{1\})$. In general, there are 2^n possibilities, among which $n+1$ are the identity permutation (having one increasing subsequence), and the others are of two increasing subsequences (with multiplicity 1). We give a few examples in Figure 1. For each card position, the color indicates the most likely number to show up in that position.

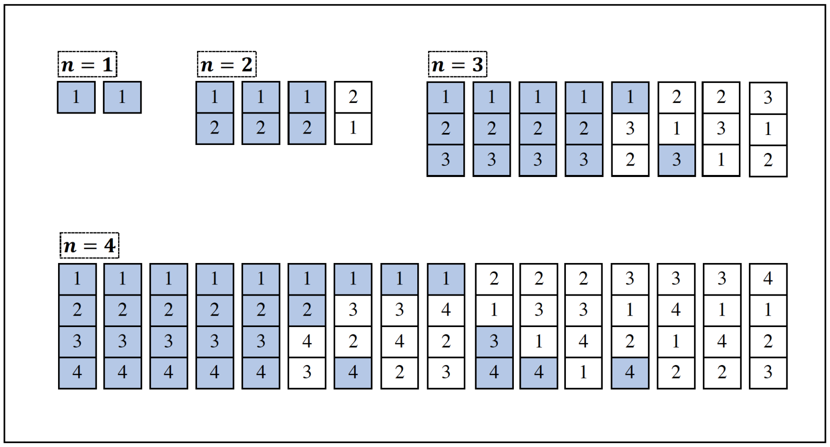

Figure 1: Examples of all possible permutations after 1-time riffle shuffling for $n = 1, 2, 3, 4$. For each card position, the color indicates the most likely number.

Frequency and probability matrices

We record the frequency of card j in position i in the frequency matrix $M = (m_{i,j})$. Row i represents the frequency at position i of the deck, and column j corresponds to card j. For example, $m_{1,3}$ gives the frequency of card 3 at the top position.

When $n = 4$, the frequency matrix is given by

$$M = \begin{bmatrix} \boxed{\mathbf{9}} & 3 & 3 & 1 \\ 4 & \boxed{\mathbf{6}} & 4 & 2 \\ 2 & 4 & \boxed{\mathbf{6}} & 4 \\ 1 & 3 & 3 & \boxed{\mathbf{9}} \end{bmatrix}.$$

Probability matrix for $k = 1$ shuffle, $P^{(1)}$

For a fixed n, we define a probability matrix for $k = 1$ shuffle, $P^{(1)} := M/2^n$. Let $a(i, j, n) := \frac{m_{i,j}}{2^n}$ be the (i, j) element of $P^{(1)}$. Then, $a(i, j, n)$ represents the probability of a card in position i being number j. It turns out that $a(i, j, n)$ has an exact formula:

$$a(i, j, n) = \frac{1}{2^i}\binom{i-1}{j-1} + \frac{1}{2^{n-i+1}}\binom{n-i}{j-i}. \tag{1}$$

This formula was mentioned and proved earlier in [2]. We notice that the first term is non-zero when $i \geq j$ and the second term is non-zero when $j \geq i$. The row sum and column sum are both equal to 1, i.e. $P^{(1)}$ is a *doubly stochastic matrix*. This is easy to show in one direction, but not so easy to show in the other direction:

$$\sum_{j=1}^{n} a(i,j,n) = \frac{1}{2^i}\sum_{j=1}^{i}\binom{i-1}{j-1} + \frac{1}{2^{n-i+1}}\sum_{j=i}^{n}\binom{n-i}{j-i} = \frac{1}{2}+\frac{1}{2} = 1,$$

and

$$\sum_{i=1}^{n} a(i,j,n) = \sum_{i=j}^{n}\frac{1}{2^i}\binom{i-1}{j-1} + \sum_{i=1}^{j}\frac{1}{2^{n-i+1}}\binom{n-i}{j-i} = 1.$$

Also, the symmetric property is easily observed through the matrix $P^{(1)}$:

$$a(i,j,n) = a(n+1-i, n+1-j, n). \tag{2}$$

Probability matrix for general k shuffles, $P^{(k)}$

Analogous to $a(i,j,n)$ for the case $k = 1$, we let $a_{i,j}^{(k)}$ denote the probability that the card number j will show up at position i after giving the deck k riffle shuffles (where we now make i,j subscripts, and suppress n from $a_{i,j}^{(k)}$ for simplicity). The probability matrix $P^{(k)}$ for k shuffles is subsequently defined by $\left(P^{(k)}\right)_{i,j} = a_{i,j}^{(k)}$.

As mentioned in [2], since the transpose matrix $\left(P^{(1)}\right)^T$ is a probability transition matrix, it follows that $P^{(k)} = \left(P^{(1)}\right)^k$, the k-th power of the matrix, and so

$$a_{i,j}^{(k)} = \sum_{s_1=1}^{n} a_{i,s_1}^{(1)} a_{s_1,j}^{(k-1)} = \cdots = \sum_{s_1=1}^{n}\cdots\sum_{s_{k-1}=1}^{n} a_{i,s_1}^{(1)} a_{s_1,s_2}^{(1)} \ldots a_{s_{k-1},j}^{(1)}. \tag{3}$$

2 Main results

We let $i \in \{1,\ldots,n\}$ be fixed and focus exclusively on the i-th row of the matrix $P^{(k)}$. For this purpose, we define a random variable J_i representing the card number on the i-th position, whose p.m.f. is given by

$$P(J_i = j) = a_{i,j}^{(k)}, \quad j = 1,\ldots,n.$$

The left plot of Figure 2 illustrates the p.m.f. of J_{60} (for $j = 1,\ldots,n$) when $k = 3$ and $n = 500$. The plot suggested a structure that led us to hypothesize that $a_{i,j}^{(k)}$ is decomposable, implying J_i follows a mixture distribution. This insight motivated our approach, which we now describe.

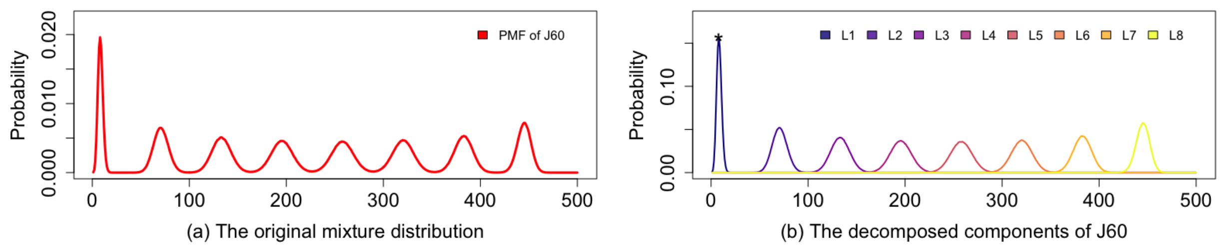

Figure 2: Example with $k = 3$, $n = 500$, $i = 60$: (a) the p.m.f. of the J_{60} : $\left(a^{(3)}_{60,j},\ j = 1, \dots, n\right)$; (b) the decomposed components p_L's of the mixture model J_{60}.

The expression for $a^{(k)}_{i,j}$ is decomposable

We start by decomposing $a^{(1)}_{i,j}$ $(k = 1)$:

$$a^{(1)}_{i,j} = \frac{1}{2}[b^{(1)}_{i,j} + b^{(2)}_{i,j}], \tag{4}$$

where

$$b^{(1)}_{i,j} := \frac{1}{2^{i-1}}\binom{i-1}{j-1}, \qquad b^{(2)}_{i,j} := \frac{1}{2^{n-i}}\binom{n-i}{j-i}. \tag{5}$$

In general, for $k \geq 1$, $a^{(k)}_{i,j}$ can be decomposed into the sum of 2^k products of a sequence of $b^{(1)}_{i,j}$ and $b^{(2)}_{i,j}$, as stated in the following proposition.

Proposition 1 (Decomposition of the structure of the p.m.f. of J_i)**.** *The expression for $a^{(k)}_{i,j}$ in* (3) *can be decomposed as follows:*

$$a^{(k)}_{i,j} = \frac{1}{2^k}\sum_L p_L(j), \tag{6}$$

where L is a list index of length k, with entries that are either 1 or 2. Specifically, $p_L(j)$ is given by

$$p_L(j) := \sum_{s_1=1}^{n} \cdots \sum_{s_{k-1}=1}^{n} b^{(L[1])}_{i,s_1} b^{(L[2])}_{s_1,s_2} \dots b^{(L[k])}_{s_{k-1},j}, \tag{7}$$

where $L[l]$ refers to the l-th element of the list L, and each term $b^{(L[l])}_{i,j}$ is defined as $b^{(1)}_{i,j}$ or $b^{(2)}_{i,j}$ in (5)*, depending on whether $L[l]$ is equal to 1 or 2.*

Proof. Straightforward from the decomposed structure (4) for $a^{(1)}_{i,j}$, and the relation (3) for $a^{(k)}_{i,j}$.

□

For a fixed value of k, there are a total of 2^k such list indices, which represent all possible ways to define a list L.

For example, for $k = 2$,

$$a_{i,j}^{(2)} = \frac{1}{4}\left[p_{[1,1]}(j) + p_{[2,1]}(j) + p_{[1,2]}(j) + p_{[2,2]}(j)\right],$$

where

$$\begin{aligned}
p_{[1,1]}(j) &= \sum_{s=1}^{n} b_{i,s}^{(1)} b_{s,j}^{(1)} = \sum_{s=1}^{n} \frac{1}{2^{i-1}}\binom{i-1}{s-1}\frac{1}{2^{s-1}}\binom{s-1}{j-1},\\
p_{[2,1]}(j) &= \sum_{s=1}^{n} b_{i,s}^{(2)} b_{s,j}^{(1)} = \sum_{s=1}^{n} \frac{1}{2^{n-i}}\binom{n-i}{s-i}\frac{1}{2^{s-1}}\binom{s-1}{j-1},\\
p_{[1,2]}(j) &= \sum_{s=1}^{n} b_{i,s}^{(1)} b_{s,j}^{(2)} = \sum_{s=1}^{n} \frac{1}{2^{i-1}}\binom{i-1}{s-1}\frac{1}{2^{n-s}}\binom{n-s}{j-s},\\
p_{[2,2]}(j) &= \sum_{s=1}^{n} b_{i,s}^{(2)} b_{s,j}^{(2)} = \sum_{s=1}^{n} \frac{1}{2^{n-i}}\binom{n-i}{s-i}\frac{1}{2^{n-s}}\binom{n-s}{j-s}.
\end{aligned}$$

Another example for $k = 3$, $L = [1, 1, 2]$ leads to

$$p_L(j) = \sum_{s_1=1}^{n}\sum_{s_2=1}^{n} b_{i,s_1}^{(1)} b_{s_1,s_2}^{(1)} b_{s_2,j}^{(2)} = \sum_{s_1=1}^{n}\sum_{s_2=1}^{n} \frac{1}{2^{i-1}}\binom{i-1}{s_1-1}\frac{1}{2^{s_1-1}}\binom{s_1-1}{s_2-1}\frac{1}{2^{n-s_2}}\binom{n-s_2}{j-s_2}.$$

The right plot in Figure 2 shows the eight components of the decomposed p.m.f. of J_{60} for $k = 3$ and $n = 500$.

Generating function for each decomposed term p_L

Let us fix the list index L (one of the 2^k possible indices) and consider the decomposed term p_L. The expression for p_L in (7) initially involves a $(k-1)$-nested binomial sum, which cannot be simplified directly. The key idea of our approach is to find the generating function for p_L, which leads to a k-nested sum that can be simplified.

In particular, we define the probability generating function of p_L as

$$\begin{aligned}
F_L(x) &= \sum_{j=1}^{n} p_L(j)\cdot x^j\\
&= \sum_{j=1}^{n}\sum_{s_1=1}^{n}\cdots\sum_{s_{k-1}=1}^{n} b_{i,s_1}^{(L[1])} b_{s_1,s_2}^{(L[2])}\dots b_{s_{k-1},j}^{(L[k])}\cdot x^j. \qquad (8)
\end{aligned}$$

Proposition 2. *The function $F_L(x)$ is a valid probability generating function. This condition guarantees that each p_L in (6) forms a valid probability mass function, and hence, J_i indeed follows a mixture distribution.*

Proof. This follows directly from the fact that $F_L(1) = 1$ and that $p_L(j)$ is clearly non-negative. $\square$

As an overview, after reordering the summation (8) (moving the sum over j to the innermost position) and evaluating from the inside out (starting with the sum over j and then over $s_{k-1}, \ldots, s_1$), we obtain a simple and tractable generating function for (8) for all $k \geq 1$. We now present this result.

Let us look at some examples of $F_L(x)$:

- for $k = 0$,

$$F_{[\,]}(x) = x^i.$$

  This is a trivial case. When cards are not shuffled, the probability that the card in position i is i is 1.

- for $k = 1$,

$$F_{[1]}(x) = (0x+1)^{n-i}\left(\frac{x}{2}+\frac{1}{2}\right)^{i-1} x, \qquad F_{[2]}(x) = \left(\frac{x}{2}+\frac{1}{2}\right)^{n-i}(x+0)^{i-1}\, x.$$

- for $k = 2$,

$$F_{[1,1]}(x) = (0x+1)^{n-i}\left(\frac{x}{4}+\frac{3}{4}\right)^{i-1} x, \qquad F_{[2,1]}(x) = \left(\frac{x}{4}+\frac{3}{4}\right)^{n-i}\left(\frac{x}{2}+\frac{1}{2}\right)^{i-1} x,$$

$$F_{[1,2]}(x) = \left(\frac{x}{2}+\frac{1}{2}\right)^{n-i}\left(\frac{3x}{4}+\frac{1}{4}\right)^{i-1} x, \qquad F_{[2,2]}(x) = \left(\frac{3x}{4}+\frac{1}{4}\right)^{n-i}(x+0)^{i-1}\, x.$$

$F_L(x)$ has a nice simple pattern!

It is utterly important that $F_L(x)$ has a simple pattern from which the underlying structure of the problem can be uncovered. To keep track of the pattern, we define a functional notation:

$$g(a,b) = [ax+(1-a)]^{n-i} \cdot [bx+(1-b)]^{i-1} \cdot x. \tag{9}$$

With this notation, the function $F_L(x)$ for $k = 0, 1$ can be expressed as

- for $k = 0$,

$$F_{[\,]}(x) = g(0,1).$$

- for $k = 1$,

$$F_{[1]}(x) = g\left(0, \frac{1}{2}\right), \quad F_{[2]}(x) = g\left(\frac{1}{2}, 1\right).$$

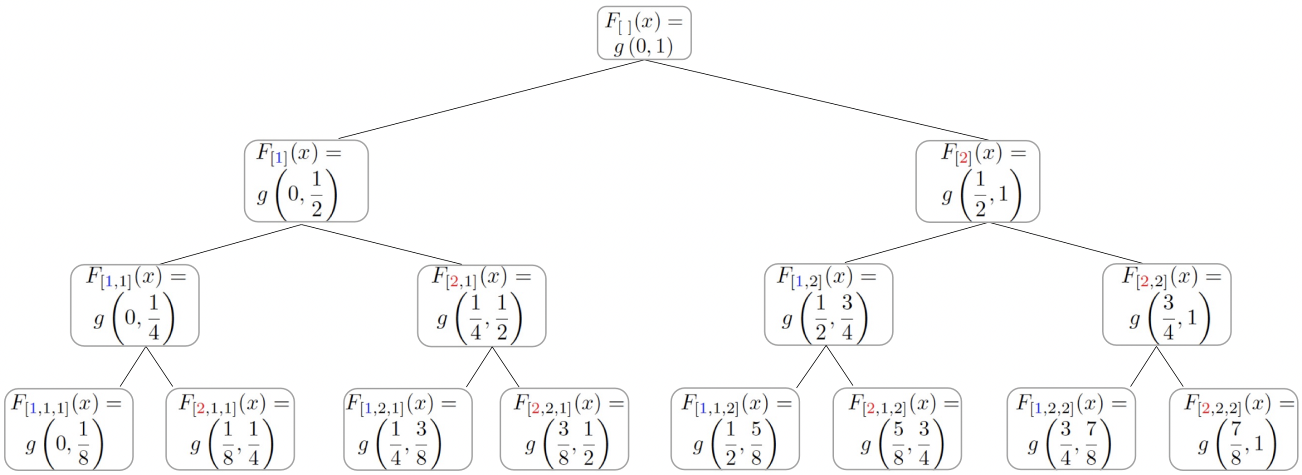

Figure 3: A bisection tree structure. Each node of depth k represents $F_{L_t}(x)$ for $t = 0, 1, \ldots, 2^k - 1$.

The order of the list indices L_t

For general $k \geq 1$, a bisection tree in Figure 3 illustrates the connections between the expressions $F_{L_t}(x)$ for k- and $(k+1)$-shuffles. At each subtree, the parent's argument (a, b) for $g(\cdot, \cdot)$ is bisected into left $(a, \frac{a+b}{2})$ and right $(\frac{a+b}{2}, b)$ child nodes. The parent's index L_t is expanded by adding 1 or 2 to form the child indices $[1, L_t]$ and $[2, L_t]$. This mapping links each index t to a unique list index L_t, defining a relationship between the generating functions for k- and $(k+1)$-shuffles. This relationship underpins the induction proof for Theorem 3 and leads to the closed-form expression for $F_{L_t}(x)$.

Theorem 3 (Generating function F_{L_t})**.** *For any given $k \geq 1$, and the list indices $L_0, L_1, \ldots, L_{2^k-1}$, labelled according to the generalization pattern shown in Figure 3, we have*

$$F_{L_t}(x) = g\left(\frac{t}{2^k}, \frac{t+1}{2^k}\right), \qquad \textit{for } t = 0, 1, \ldots, 2^k - 1,$$

where $g(a,b) = [ax + (1-a)]^{n-i} \cdot [bx + (1-b)]^{i-1} \cdot x$ for $a, b \in [0, 1]$.

From this point onward, let us fix $t \in \{0, 1, \ldots, 2^k - 1\}$, and write $L := L_t$ for simplicity. We use the following lemma to find the sum of the terms in $F_L(x)$ recursively.

Lemma 4. *Assume $f(i) = A^{n-i}B^{i-1}x$ where $A = ax + (1-a)$ and $B = bx + (1-b)$. Then,*

$$T_1(f) := \sum_{s=1}^{i} \frac{1}{2^{i-1}} \binom{i-1}{s-1} f(s) = A^{n-i} \left(\frac{A+B}{2}\right)^{i-1} x;$$

$$T_2(f) := \sum_{s=i}^{n} \frac{1}{2^{n-i}} \binom{n-i}{s-i} f(s) = \left(\frac{A+B}{2}\right)^{n-i} B^{i-1} x.$$

Proof. First, by the binomial theorem, it is easy to show that

$$\sum_{s=1}^{i} \frac{1}{2^{i-1}} \binom{i-1}{s-1} x^{s-1} = \left(\frac{x}{2} + \frac{1}{2}\right)^{i-1}$$

and

$$\sum_{s=i}^{n} \frac{1}{2^{n-i}} \binom{n-i}{s-i} x^{s-1} = \left(\frac{x}{2} + \frac{1}{2}\right)^{n-i} \cdot (x+0)^{i-1}.$$

For the general case,

$$\begin{aligned}\sum_{s=1}^{i} \frac{1}{2^{i-1}} \binom{i-1}{s-1} f(s) &= \sum_{s=1}^{i} \frac{1}{2^{i-1}} \binom{i-1}{s-1} A^{n-1} \left(\frac{B}{A}\right)^{s-1} x \\ &= A^{n-1} x \left(\frac{A+B}{2A}\right)^{i-1}, \quad \text{from the above identity} \\ &= A^{n-i} \left(\frac{A+B}{2}\right)^{i-1} x.\end{aligned}$$

This proves the first claim regarding $T_1(f)$. The second claim can be done similarly. $\square$

We now return to the proof of Theorem 3.

Proof of Theorem 3. We prove by induction on $k = |L|$. For the base case, we verify that $F_{[\,]}(x) = x^i$, is indeed $g(0,1)$. For the induction step, let $L' = [l_{k-1}, .., l_2, l_1]$ and $L = [l_k, l_{k-1}, ..., l_2, l_1] = [l_k, L']$. Assume $f := F_{L'}(x) = g\left(\frac{t}{2^{k-1}}, \frac{t+1}{2^{k-1}}\right)$. Then it follows from Lemma 4 that $F_L(x) = T_1(f) = g\left(\frac{2t}{2^k}, \frac{2t+1}{2^k}\right)$ if $l_k = 1$ and $F_L(x) = T_2(f) = g\left(\frac{2t+1}{2^k}, \frac{2(t+1)}{2^k}\right)$ if $l_k = 2$. $\square$

Open Problem 1: Find a combinatorial interpretation for the generating function $F_L(x)$ of this card guessing problem.

Corollary 5. *For fixed i, n, and k, let J_i be a random variable representing the card number in the i-th position of an n-card deck after k shuffles. Define its probability generating function as*

$$G_i(x) = \sum_{j=1}^{n} P(J_i = j) x^j.$$

Then, the p.g.f. of J_i is the sum of the product of polynomials given by

$$G_i(x) = \frac{1}{2^k} \sum_{t=0}^{2^k-1} \left(\frac{t}{2^k} x + 1 - \frac{t}{2^k}\right)^{n-i} \left(\frac{t+1}{2^k} x + 1 - \frac{t+1}{2^k}\right)^{i-1} x. \tag{10}$$

Proof.

$$G_i(x) = \sum_{j=1}^{n} a_{i,j}^{(k)} x^j = \frac{1}{2^k} \sum_{t=0}^{2^k-1} \sum_{j=1}^{n} p_{L_t}(j) x^j = \frac{1}{2^k} \sum_{t=0}^{2^k-1} F_{L_t}(x).$$

The second equality follows from (6), and the result is a consequence of Theorem 3. □

Observations and applications

- For fixed i, n, and k, the p.g.f. $G_i(x)$ of J_i in (10) simplifies the p.m.f. calculation $P(J_i = j) = a_{i,j}^{(k)}$, reducing the summands in (3) from $k-1$ to a single term over t.
- For a fixed card position i, the optimal strategy is to select j_i^* that maximizes the probability $P(J_i = j)$. For example, in Figure 2 with $n = 500$ and $k = 3$, the maximizer j_{60}^* corresponds to the mode of the p.m.f., or the degree j^* of x^j in the p.g.f. $G_{60}(x)$ with the largest coefficient. Here, $j_{60}^* = 8$, meaning we should guess "8" for the 60th card.
- Formula (10) also shows that, when k is fixed, $G_i(x)$ is a mixture model with 2^k components, each being the product of simple polynomials. This simple structure has the potential to analyze cases where k is fixed and $n \to \infty$, opening the door for future studies on asymptotic behavior of optimal strategies.

References

[1] Dave Bayer, Persi Diaconis. *Trailing the dovetail shuffle to its lair*, Ann. Appl. Probab. 2 (1992), 294-313.

[2] Mihai Ciucu. *No-feedback card guessing for dovetail shuffles.* Ann. Appl. Probab. 8(4) (1998), 1251-1269.

[3] Tipaluck Krityakierne, Poohrich Siriputcharoen, Thotsaporn Aek Thanatipanonda, Chaloemkiat Yapolha. *Moments of the one-shuffle no-feedback card guessing game.* Discrete Mathematics Letters, 12 (2023), 110-117.