Discrete Mathematics Letters
www.dmlett.com

Discrete Math. Lett. **12** (2023) 110–117
DOI: 10.47443/dml.2023.119

Research Article

Moments of the one-shuffle no-feedback card guessing game

Tipaluck Krityakierne[1,3], Poohrich Siriputcharoen[1], Thotsaporn Aek Thanatipanonda[2,*], Chaloemkiat Yapolha[1]

[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, Commission on Higher Education, Bangkok, Thailand*

(Received: 6 June 2023. Received in revised form: 22 July 2023. Accepted: 25 July 2023. Published online: 7 August 2023.)

Abstract

We consider card guessing with no feedback, a variant of the game previously studied by Ciucu in 1998. In this study, we derive an exact, closed-form formula for the asymptotic (in the number of cards, n) expected number of correct guesses, as well as higher moments, for a one-time riffle shuffle game under the optimal strategy. The problem is tackled using two different approaches: one approach utilizes a fast generating function based on a recurrence relation to obtain numerical moments, while the other is the symbolic approach employing the method of indicators for finding expected counts. The results obtained contribute to the existing literature on card guessing with no feedback.

Keywords: card guessing; higher-order moments; generating function; no feedback; combinatorics; method of indicators.

2020 Mathematics Subject Classification: 05A15.

1. Preliminary Discussion

In a typical card guessing game, the game starts with a deck of n cards ordered from 1 to n. After giving the deck riffle shuffles, a player guesses the card one at a time until all cards in the deck have been guessed. The objective of the game is to maximize the number of correct guesses. Depending on the rules, the player may or may not receive feedback. In the complete feedback game, the player sees the card after each guess and can adjust the strategy based on the cards that the player has already seen. The no-feedback game is equivalent to the player guessing all the cards beforehand. For a brief review of this topic, refer to e.g. [2–5, 8].

For a game with complete feedback, the expected number of correct guesses after one riffle shuffle was recently established in [9]. Subsequently, Krityakierne and Thanatipanonda [7] further improved upon the expression, and developed a unified framework for systematically calculating higher-order moments. In contrast, no progress has been made in the field of no-feedback games over the past several decades. Hence, we take this opportunity to present some of our findings in this area. Specifically, by building upon Ciucu's optimal guessing strategy for the one-time shuffle no-feedback game [3], we derive the exact closed-form formula (in terms of n) for the expected number of correct guesses and all higher moments. To achieve this goal, we first compute a straightforward generating function for each value of n that directly enumerates and counts the number of correct guesses for all possible permutations after a single shuffle. We then introduce a recurrence relation that efficiently computes generating functions and obtains numerical moments in Section 2. Furthermore, we develop a three-level combinatorial approach for computing the moments symbolically in Section 3. We provide a specific Maple command in the relevant sections of the paper. All Maple programs used to generate the results in this paper can be found at https://thotsaporn.com/Card3.html.

Card guessing in nutshell

A deck of n cards can be viewed as a permutation on $\{1, 2, ..., n\}$. Starting with the identity permutation, we give a formal definition of the Gilbert-Shannon-Reeds (GSR) model for riffle shuffles. Despite not fully reflecting the reality of someone attempting to perform a good riffle shuffle, the GSR model possesses several merits due to its simplicity. It effectively captures the way an amateur would perform riffle shuffles and provides an opportunity for analysis [1].

Definition 1.1 (GSR model [6])**.** *Gilbert-Shannon-Reeds model for riffle shuffles is performed by splitting the identity permutation* $[1, 2, \ldots, n]$ *into two sequences/piles (possible to have 0 cards in one of the piles) in such a way that the probability*

*Corresponding author (thotsaporn.tha@mahidol.edu).

of cutting the first t cards is $\binom{n}{t}/2^n$. Subsequently, interleave the two piles back into a single deck in a manner such that all the $\binom{n}{t}$ interleaving possibilities are equally likely to happen with a probability of $\frac{1}{2^n}$.

In this work, we will stick to the following optimal guessing strategy for a one-time shuffled deck.

Proposition 1.1 (Optimal guessing strategy [3])**.** *For a single-shuffled deck, the optimal guessing strategy $\mathcal{G}^*$ is to guess the top half of the deck with sequence*

$$1, 2, 2, 3, 3, 4, 4, \ldots$$

and guess the bottom half in the reverse manner, i.e.

$$\ldots, n-3, n-3, n-2, n-2, n-1, n-1, n,$$

where the middle card is included in the top half of the deck when n is odd.

The symmetry in the above optimal guessing strategy $\mathcal{G}^*$ (i.e. mirror images along the position "half-deck") plays a crucial role in simplifying several calculations in this work. Note also that this optimal strategy is not unique. In fact, the player can choose any number from the set $\mathcal{S}_i$ for the card position i, where

$$\mathcal{S} = \{1\}, \{2\}, \{2\}, \{2,3\}, \{3\}, \{3,4\}, \{4\}, \{4,5\}, \{5\}, \ldots \qquad \text{Top half}$$

and (in the reverse manner)

$$\ldots, \{n-3, n-2\}, \{n-2\}, \{n-2, n-1\}, \{n-1\}, \{n-1\}, \{n\} \quad \text{Bottom half.}$$

The non-uniqueness of the optimal strategy for the one-time shuffled deck, which has not been previously mentioned elsewhere, can be verified through the position matrix M provided in [3, Lemma 2.1].

Example 1.1. *(4 cards, 1-time shuffle) Let us shuffle a deck of 4 cards once. Figure 1 gives all possible outcomes of 16 permutations. Since there are 5 identity permutations, in order to distinguish one permutation from another, a red arrow is used to indicate the position where the deck was split before interleaving.*

Permutation π	1	1	1	1	1	1	1	1	1	2	2	2	3	3	3	4
	2	2	2	2	2	2	3	3	4	1	3	3	1	4	1	1
	3	3	3	3	3	4	2	4	2	3	1	4	2	1	4	2
	4	4	4	4	4	3	4	2	3	4	4	1	4	2	2	3
#Correct	**4**	**4**	**4**	**4**	**4**	**2**	**2**	**1**	**1**	**2**	**1**	**0**	**1**	**0**	**0**	**0**

Figure 1: Sample space for an experiment with a deck of 4 cards after one shuffle. The red arrow marks the location where the deck was split before interleaving. The blue color indicates a correct guess under the optimal strategy $\mathcal{G}^* = [1, 2, 3, 4]$.

The fact that there are total number of 2^n shuffles for a single-shuffled deck is not a coincidence. It is attributed to the $\binom{n}{t}$ ways of interleaving the two piles, whose initial cut has t cards on the first pile. Consequently, the complete count of shuffles is given by $\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n} = 2^n$.

A quick refresher on the method of indicators (also known as the method of overlapping stages [10]) will be now given. Let X be a random variable representing the number of correct guesses (under $\mathcal{G}^*$) for a permutation π. Assume that $X = \sum_{i=1}^{n} X_i$, and X_i, taking only value 1 or 0, is the indicator of whether the ith guess is correct.

Example 1.2. *Using the previous example, we fist find $C[X]$, the sum of all possible numbers of correct guesses, and $C[X^2]$, the sum of all possible numbers of squared correct guesses:*

$$\begin{aligned}
C[X] &= C[(X_1 + X_2 + X_3 + X_4)] = 9 + 6 + 6 + 9 = 30.\\
C[X^2] &= C[(X_1 + X_2 + X_3 + X_4)^2]\\
&= C[X_1^2] + C[X_2^2] + C[X_3^2] + C[X_4^2]\\
&\quad + 2(C[X_1X_2] + C[X_1X_3] + C[X_1X_4] + C[X_2X_3] + C[X_2X_4] + C[X_3X_4])\\
&= (9 + 6 + 6 + 9) + 2(6 + 5 + 6 + 5 + 5 + 6) = 96.
\end{aligned}$$

Then, $E[X] = \frac{30}{2^4}$ *and* $E[X^2] = \frac{96}{2^4}$.

With the optimal strategy $\mathcal{G}^*$ in mind, the (counting) generating function can be built to keep track of the number of correct guesses for each of the 2^n permutations, after one riffle shuffle:

$$F_n(q) = \sum_{i=0}^{n} a_i q^i, \tag{1}$$

where a_i denotes the number of permutations with i correct guesses under the optimal strategy $\mathcal{G}^*$.

Example 1.3. *The generating function for the previous example is* $F_4(q) = 4+4q+3q^2+5q^4$. *Moreover,* $C[X] = F_4'(q)|_{q=1} = 30$ *and* $C[X^2] = \left[qF_4'(q)\right]'|_{q=1} = 96$.

</ >The Maple command for this empirical (slow) approach is **`GenSlow(n,q)`**. For example, try: **`GenSlow(12,q);`** This empirical method computes $F_n(q)$ for $n = 13$ in 14.3 seconds.

2. Fast computation methods of generating functions

2.1. Generating functions through a recurrence relation

2.1.1. Our first recurrence relation, G

Computing the generating function directly using (1) requires an exhaustive list of every possible permutation before we can count the number of correct guesses one by one, which becomes very slow whenever n gets large. We now discuss how to use a recurrence relation to speed up computations.

Because of the symmetry of the optimal guessing strategy, we will first concentrate on the top half of the deck and derive a recurrence relation for it. Let $h = \left\lceil \frac{n}{2} \right\rceil$ be the length of the top half. (Hence, $n-h$ is the length of the bottom half). We assume the cut creates two sequences as $A = [1, 2, \ldots, s-1]$ and $B = [s, s+1, \ldots, n]$ for some $1 \le s \le n+1$. Let $a_1, a_2 \ge 0$ be the lengths of the portions of A and B in the **top half** after interleaving, i.e. $a_1 + a_2 = h$.

The following example illustrates how to set up a recurrence relation that counts the total number of correct guesses in the top half as a function of a_1, a_2 and s. Let us denote that number by $G(a_1, a_2, s, q)$.

Example 2.1. *Assume* $h = 5$ *with* $a_1 = 3, a_2 = 2$ *and* $s = 7$. *These numbers imply that the cut creates two sequences as* $A = [1, 2, 3, 4, 5, 6]$ *and* $B = [7, 8, 9, 10]$ *(assuming n is even). Moreover, after a riffle shuffle, the top-half deck consists of the numbers* $1, 2, 3, 7, 8$, *with either 3 or 8 being the last card of the top half. Now, imagine the very last card of the top half: it can either be a correct or an incorrect guess. The correct guess for the* 5*th card under* $\mathcal{G}^*$ *is 3, which is a number in the sequence* A. *Hence, in this case, the number of correct guesses is one more than the number of correct guesses in the rest of the top half if the top half ends with the number from* A, *or remains the same otherwise. That is,*

$$G(3, 2, 7, q) = q^1 G(2, 2, 7, q) + q^0 G(3, 1, 7, q).$$

We state this example formally in the next proposition.

Proposition 2.1. *Let* $c = \left\lfloor \frac{a_1 + a_2}{2} \right\rfloor + 1$. *Then,* c *is the optimal guess at the position* $(a_1 + a_2)$*th for a 1-time shuffled deck under* $\mathcal{G}^*$, *and*

$$G(a_1, a_2, s, q) = q^{\delta(c, a_1)} G(a_1 - 1, a_2, s, q) + q^{\delta(c, s+a_2-1)} G(a_1, a_2 - 1, s, q), \tag{2}$$

where $G(0, 0, s, q) = 1$. $\delta(c, d)$ *is the Kronecker delta function, taking value* 1 *if* $c = d$, *and* 0 *otherwise.*

In order to use the same function G defined above with the bottom half, we need to relabel the entries in the bottom half in the following way: First, we mirror each entry of the bottom half about n, i.e. each entry e is relabelled to $n + 1 - e$. Then, the entries are reversed, the first entry becoming the last, and the last entry becoming the first. For example, consider the sequence $[4, 11, 5, 12, 6, 13, 14]$ in the bottom half with $n = 14$. This sequence will be rewritten as $[1, 2, 9, 3, 10, 4, 11]$. Moreover, the original optimal guessing strategy for the bottom half $[11, 11, 12, 12, 13, 13, 14]$ will be relabelled into $[1, 2, 2, 3, 3, 4, 4]$. Then, the calculation could be done the same way as the top half.

To get the full generating function, i.e. the number of correct guesses of the whole deck, we multiply the generating function of the top half with that of the bottom half (after relabelling):

$$G(a_1, a_2, a_1 + b_1 + 1, q) \cdot G(b_2, b_1, b_2 + a_2 + 1, q),$$

where $b_1, b_2 \ge 0$ are the lengths of the portions of A and B in the **bottom half** such that $b_1 + b_2 = n - h$.

We properly restate this in the next proposition by reparameterizing it in terms of h. With the reparameterization, the top half will depend only on a_1, and the bottom will depend only on b_2. We thus drop the subscripts 1 and 2 hereafter.

Proposition 2.2. *Shuffle an n-card deck once. Denote by $h = \left\lceil \frac{n}{2} \right\rceil$ the length of the top half. Let a and b be the length of the portion of A in the top half, and the length of the portion of B in the bottom half, respectively. Then,*

$$F(a,b,h,q) := G(a,h-a,a+(n-h-b)+1,q) \cdot G(b,n-h-b,b+(h-a)+1,q) \tag{3}$$

is the generating function, which counts the total number of correct guesses from all permutations with parameters a and b.

Summing over all possible values of parameters a and b, we obtain a fast approach for computing the generating function for the number of correct guesses of all permutations after one shuffle.

Corollary 2.1. *Let $F_n(q)$ be the generating function for the number of correct guesses of all permutations after one shuffle. Then,*

$$F_n(q) = \sum_{a=0}^{h} \sum_{b=0}^{n-h} G(a,h-a,a+(n-h-b)+1,q) \cdot G(b,n-h-b,b+(h-a)+1,q). \tag{4}$$

The fact that the generating function defined in Corollary 2.1 is a multiplication of two recursive functions allows us to compute the generating function $F_n(q)$ for as large as $n = 1000$. Figure 2 displays histograms for the distribution of the number of correct guesses for several values of n. It is clear that these distributions are not normal.

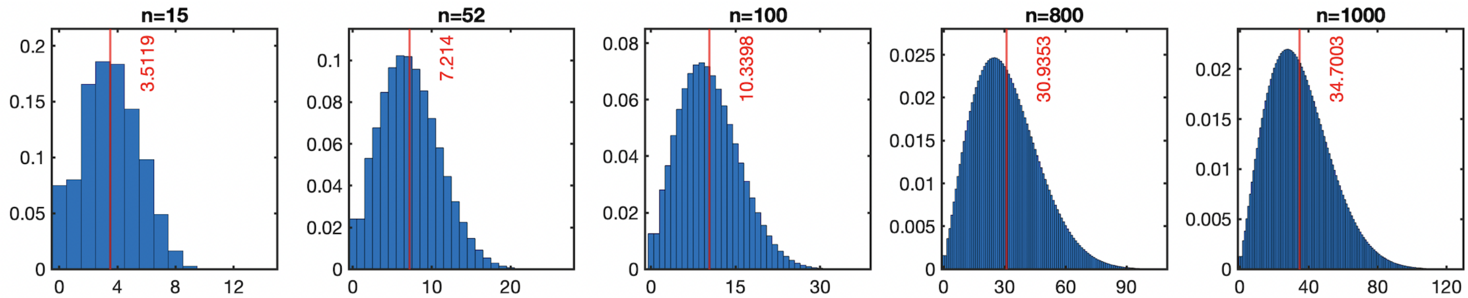

Figure 2: Probability histograms of the number of correct guesses when n varies. The red vertical line indicates the corresponding expected value $E[X]$.

2.1.2. Simplified recurrence relation, *H*

We build a new recurrence relation H, which will also be used to compute $F_n(q)$. Based on the observation that the only way for correct guesses in the top half can come from B is if either of the following two cases occurs in the top half:

$$[1|2,3,4,\ldots,h] \text{ or } [\varnothing|1,2,3,4,\ldots,h],$$

we set up H assuming that none do. We then correct for these cases in the calculation of $F_n(q)$ below in Proposition 2.3. (Here, the vertical bar is used to separate the numbers coming from A and B. The blue and red colors indicate the correct guesses from A and B, respectively.)

Thus, our new recurrence (for the top half) will keep track of only those correct guesses from A. A simplified recurrence, denoted by H, is set as

$$H(a_1,a_2,q) = q^{\delta(c,a_1)} H(a_1-1,a_2,q) + H(a_1,a_2-1,q),$$

with the same base case $H(0,0,q) = 1$.

As we do not have to take into account correct guesses from B, we dropped the variable s. This makes the computations very fast.

Proposition 2.3. *Let $F_n(q)$ be the generating function for the number of correct guesses of all permutations after one shuffle. Then,*

$$\begin{aligned} F_n(q) &= -2q^2 - 2q^3 + 4q^4 + \sum_{a=0}^{h} \sum_{b=0}^{n-h} H(a,h-a,q) \cdot H(b,n-h-b,q) \\ &= -2q^2 - 2q^3 + 4q^4 + \sum_{a=0}^{h} H(a,h-a,q) \cdot \sum_{b=0}^{n-h} H(b,n-h-b,q). \end{aligned} \tag{5}$$

The modification part (the first three terms of (5)) is due to the four sequences which are miscalculated:

$$[\overline{1|2,3,4,\ldots,h},\underline{h+1,\ldots,n-1,n}],\ \ [\overline{\varnothing|1,2,3,4,\ldots,h},\underline{h+1,\ldots,n-1,n}],$$

$$[\overline{1,2,\ldots,h},\underline{h+1,\ldots,n-1|n}],\ \ [\overline{1,2,\ldots,h},\underline{h+1,\ldots,n-1,n|\varnothing}].$$

Each of the four half-permutations leads to the same full-deck identity permutation, $[1,2,3,\ldots,n]$, each contributing 4 number of correct guesses (i.e. q^4). The overline and underline indicate the top and bottom halves. The blue entries indicate those correct guesses already counted prior to the modification, which are numbers from A for the top half and B for the bottom half. For example, since the object $[1|2,3,4,\ldots,h]$ is equivalent to $[1|2,3,4,\ldots,h,h+1,\ldots,n-1,n]$, and those blue correct guesses have been counted already in the double series, we thus subtract q^3 from the final generating function.

</ >The Maple command for obtaining generating functions through our first recurrence relation (G) and the simplified one (H) are **GenFast(n,q)** and **GenFastest(n,q)**, respectively. For example, try: **GenFast(120,q);** and **GenFastest(120,q);** We can generate $F_n(q)$ for $n = 1000$ via **GenFastest(n,q)** within 11.4 seconds. With the same amount of time, **GenFast(n,q)** can only compute $F_n(q)$ for $n = 200$. It is also important to note that the complexity of the method in this section is only polynomial time.

2.2. Numerical moments from generating functions

For a fixed value of $r \geq 0$, we can use any versions of the generating function $F_n(q)$, (1), (4), or (5), to get a numeric rth moment for the number of correct guesses:

$$C[X^r] = D^{(r)}F_n(q)|_{q=1}, \tag{6}$$

where the operator $Dt(q) := qt'(q)$ and $D^{(r)}$ means we repeatedly apply the operator D, r times.

Not only is the simplified recurrence relation (5) very computationally efficient, it also leads us to the following useful results. In particular, ignoring B (A), the last term of the relation (5) boils down to the generating function for the number of correct guesses from A (B) in the top half (bottom half). The next corollary connects this fact with the operator $D^{(r)}$.

Corollary 2.2. *Let Y_A be the number of correct guesses in the top half from A and Z_B the number of correct guesses in the bottom half from B, respectively. Then,*

$$C[Y_A^r] = 2^{n-h} \cdot D^{(r)}F_A(q)|_{q=1} \tag{7}$$

$$C[Z_B^r] = 2^{h} \cdot D^{(r)}F_B(q)|_{q=1}, \tag{8}$$

where $F_A(q) = \sum_{a=0}^{h} H(a,h-a,q)$ and $F_B(q) = \sum_{b=0}^{n-h} H(b,n-h-b,q)$.

Lastly, the simplified recurrence relation (5) also implies the independence between the number of correct guesses from A in the top and those from B in the bottom halves of the deck. We state this in the next corollary.

Corollary 2.3. *Y_A and Z_B are independent of each other.*

3. A combinatorial approach to the rth moment

Again, we let X be a random variable representing the number of correct guesses for the entire deck after a single shuffle. Our ultimate goal is to derive a closed-form formula for $E[X^r]$, the rth moment of X, as a function of n based on the optimal guessing strategy $\mathcal{G}^*$. While the relation (5) allows us to obtain numerical moments (by differentiating both sides of the relation), calculating the closed-form expression becomes significantly more complex using this approach. In this section, we introduce a systematic framework for calculating the rth moment from the combinatorial perspective, which enables us to derive higher moments symbolically.

3.1. Three-level procedure for finding the rth moment

As before, assume that the cut creates two sequences A and B, and let a be the length of the portion of A in the top half, and b the length of the portion of B in the bottom half. Recall that $h = \lceil n/2 \rceil$ denotes the length of the top half.

We now present the three-level procedure, a combinatorial approach for calculating the rth moment about the origin (raw moment).

High Level: Let Y_A be the number of correct guesses in the top half from A and Z_B the number of correct guesses in the bottom half from B, respectively. Then,

$$E[X^r] = E[(Y_A + Z_B)^r] + \frac{e(r)}{2^n}, \tag{9}$$

where

$$E[(Y_A + Z_B)^r] = \sum_{i=0}^{r} \binom{r}{i} E[Y_A^i] \cdot E[Z_B^{r-i}], \tag{10}$$

and $e(r)$ is given by (11).

The last equality holds because $E[Y_A^i Z_B^{r-i}] = E[Y_A^i]E[Z_B^{r-i}]$ for the independent random variables Y_A and Z_B (see Corollary 2.3). Also, $e(r)$ is the four permutations that were ignored at first and can be calculated from (5). Let $f(q) := 4q^4 - 2(q^3 + q^2)$ and define the operator $Dt(q) := qt'(q)$. Then,

$$e(r) = D^{(r)} f(q)|_{q=1} = 4 \cdot 4^r - 2(3^r + 2^r). \tag{11}$$

Example 3.1. $e(1) = 6$ *and* $e(2) = 38$.

Due to the direct connection $E[Y_A^r] = \dfrac{C[Y_A^r]}{2^n}$, we will present the procedure for the Middle and Low Levels in terms of $C[Y_A^r]$.

Middle Level: Decompose Y_A into a sum of indicator random variables:

$$Y_A = Y_1 + Y_2 + \cdots + Y_h,$$

where $Y_i = 1$ if the ith position is guessed correctly and comes from A, and 0 otherwise.

The rth moment of Y_A can be found by

$$\begin{aligned} C[Y_A^r] &= C[(Y_1 + Y_2 + \cdots + Y_h)^r] \\ &= \sum_{i_1=1}^{h} \sum_{i_2=1}^{h} \cdots \sum_{i_r=1}^{h} C[Y_{i_1} Y_{i_2} \dots Y_{i_r}]. \end{aligned} \tag{12}$$

The formula for $C[Y_{i_1} Y_{i_2} \dots Y_{i_r}]$ when $1 \le i_1 < i_2 < \cdots < i_r \le h$ is given by (13) below.

Low Level, the Building Block $C[Y_{i_1} Y_{i_2} \dots Y_{i_r}]$

For $1 \le i_1 < i_2 < \cdots < i_r \le h$,

$$C[Y_{i_1} Y_{i_2} \dots Y_{i_r}] = \binom{i_1 - 1}{\lfloor i_1/2 \rfloor} \cdot \binom{i_2 - i_1 - 1}{\lfloor i_2/2 \rfloor - \lfloor i_1/2 \rfloor - 1} \cdots \binom{i_r - i_{r-1} - 1}{\lfloor i_r/2 \rfloor - \lfloor i_{r-1}/2 \rfloor - 1} \cdot 2^{n-i_r}. \tag{13}$$

Remark 3.1. *We only need the formula* $C[Y_{i_1} Y_{i_2} \dots Y_{i_r}]$ *for all distinct indices. If, for example,* $i_1 = i_2$*, then*

$$C[Y_{i_1}^2 Y_{i_3} Y_{i_4} \dots Y_{i_r}] = C[Y_{i_1} Y_{i_3} Y_{i_4} \dots Y_{i_r}],$$

and so the formula for $r - 1$ *distinct indices applies.*

Justification of the $C[Y_{i_1} Y_{i_2} \dots Y_{i_r}]$ formula

Starting with $r = 1$, $C[Y_i]$ corresponds to the number of ways the fixed position i coming from A is guessed correctly, and (13) is simplified to the following. For $1 \le i \le h$,

$$C[Y_i] = \binom{i-1}{\lfloor i/2 \rfloor} \cdot 2^{n-i}. \tag{14}$$

The formula can be justified using a combinatorial counting method as follows. For a fixed i, Y_i is the indicator variable indicating the correctness of the guess at the ith position, and coming from A with value $\lfloor i/2 \rfloor + 1$. The number of ways to put numbers before the ith position is $\binom{i-1}{\lfloor i/2 \rfloor}$. Each of the values that comes after the ith position can come from either A or B, providing 2^{n-i} possibilities.

When $r = 2$, $C[Y_iY_j]$ corresponds to the number of ways to obtain the correct guesses at two different positions i and j, both coming from A. In addition to the ith position, the jth position must have value $\lfloor j/2 \rfloor + 1$. Hence, the number of ways to put numbers between the ith and jth position is $\binom{j-i-1}{\lfloor j/2 \rfloor - \lfloor i/2 \rfloor - 1}$. The rest have 2^{n-j} possibilities. Thus, we obtain the following formula. For $1 \le i < j \le h$,

$$C[Y_iY_j] = \binom{i-1}{\lfloor i/2 \rfloor} \cdot \binom{j-i-1}{\lfloor j/2 \rfloor - \lfloor i/2 \rfloor - 1} \cdot 2^{n-j}. \tag{15}$$

This idea can be generalized to justify the formula for $C[Y_{i_1}Y_{i_2}\dots Y_{i_r}]$, for $r \ge 2$. We remark that $C[Y_{i_1}Y_{i_2}\dots Y_{i_r}]$ is calculated without the length of A or B in consideration.

The first moment, $E[X]$

We are now ready to obtain a closed-form formula for the first moment. Starting from the Low Level formula $C[Y_i]$ and proceeding backwards, we provide detailed calculations for $\sum_{i=1}^{h} C[Y_i]$ of the Middle Level, and eventually for $E[X]$ of the High Level.

Middle Level: Equations (12) and (14) for $r = 1$ lead to

$$C[Y_A] = \sum_{i=1}^{h} C[Y_i] = 2^n \sum_{i=1}^{h} \binom{i-1}{\lfloor i/2 \rfloor} \frac{1}{2^i},$$

which can be evaluated in closed form. In order to get rid of the floor function, we consider the following two cases.

Case 1: h is even. Assume $h = 2L$. Then,

$$\begin{aligned} C[Y_A] &= 2^n \sum_{s=1}^{L} \binom{2s-1}{s} \frac{1}{4^s} + 2^n \sum_{s=0}^{L-1} \binom{2s}{s} \frac{1}{2 \cdot 4^s} = 2^n \left[\sum_{s=0}^{L-1} \binom{2s}{s} \frac{1}{4^s} + \frac{1}{2}\binom{2L}{L}\frac{1}{4^L} - \frac{1}{2} \right] \\ &= 2^n \left[\frac{2L + 1/2}{4^L} \binom{2L}{L} - \frac{1}{2} \right] = 2^n \left[\frac{h + 1/2}{4^L} \binom{2L}{L} - \frac{1}{2} \right]. \end{aligned}$$

Case 2: h is odd. Assume $h = 2L - 1$. Then,

$$\begin{aligned} C[Y_A] &= 2^n \sum_{s=1}^{L-1} \binom{2s-1}{s} \frac{1}{4^s} + 2^n \sum_{s=0}^{L-1} \binom{2s}{s} \frac{1}{2 \cdot 4^s} = 2^n \left[\sum_{s=0}^{L-1} \binom{2s}{s} \frac{1}{4^s} - \frac{1}{2} \right] \\ &= 2^n \left[\frac{2L}{4^L} \binom{2L}{L} - \frac{1}{2} \right] = 2^n \left[\frac{h+1}{4^L} \binom{2L}{L} - \frac{1}{2} \right]. \end{aligned}$$

In both of these cases, we have used the following identity (which can be routinely proved by induction once it is found or by Maple)

$$\sum_{s=0}^{L-1} \binom{2s}{s} \frac{1}{4^s} = \frac{2L}{4^L} \binom{2L}{L}.$$

High Level: By linearity of expectation and by (9) with $e(1) = 6$,

$$E[X] = E[Y_A] + E[Z_B] + \frac{6}{2^n}.$$

The formula of $E[X]$ therefore depends on the parity of h and $n - h$, which together are determined by n mod 4. We obtain the following theorem.

Theorem 3.1. *Let* $L = \lceil h/2 \rceil$ *and* $\alpha \in \{-1, 0, 1, 2\}$. *For* $n = 4L + \alpha$,

$$E[X] = \frac{n + 1 - \alpha/2}{4^L} \binom{2L}{L} - 1 + \frac{6}{2^n}. \tag{16}$$

This theorem provides an exact, closed-form formula for $E[X]$. The leading term of $E[X]$ was derived earlier in [3,8]:

$$E[X] = \frac{2\sqrt{n}}{\sqrt{\pi}} + \mathcal{O}(1).$$

3.2. The higher moments and the distribution of X

All higher moments, $E[X^r]$, can be obtained using the same procedure as that of the first moment. However, the reader may have noticed that the bottleneck lies in deriving a closed-form expression for the sum $\sum C[Y_{i_1}Y_{i_2}\dots Y_{i_r}]$ in (12), which becomes extremely tedious due to the presence of the floor function in (13).

To assist with such tedious computations, we use Maple. The following is the list of symbolic moments for the case when $n = 4L$, obtained using the three-level procedure proposed.

$$E[X^0] = 1;$$

$$E[X^1] = \frac{(4L+1)}{4^L}\binom{2L}{L} - 1 + \frac{6}{2^{4L}};$$

$$E[X^2] = \frac{(4L+1)^2}{2\cdot 4^{2L}}\binom{2L}{L}^2 - \frac{6(2L+1)}{4^L}\binom{2L}{L} + 4L + \frac{11}{2} + \frac{38}{2^{4L}};$$

$$E[X^3] = -\frac{3(8L+5)(4L+1)}{2\cdot 4^{2L}}\binom{2L}{L}^2 + \frac{2(20L^2+48L+17)}{4^L}\binom{2L}{L} - 24L - \frac{53}{2} + \frac{186}{2^{4L}};$$

$$E[X^4] = \frac{(256L^3+1024L^2+736L+151)}{2\cdot 4^{2L}}\binom{2L}{L}^2 - \frac{8(50L^2+94L+33)}{4^L}\binom{2L}{L} + 48L^2 + 232L + \frac{377}{2} + \frac{830}{2^{4L}}.$$

</>The Maple command to find these moments is **`ForMoX(n1,r,L)`**, where **`n1`** denotes the α value in Theorem 3.1. For example, for the result of $E[X^4]$ above, try **`ForMoX(0,4,L)`**.

To study the distribution of X, the number of correct guesses after a single shuffle, we calculate the moments about the mean. The list for the case $n = 4L$ is provided below.

$$E[X-\mu] = 0;$$

$$E[(X-\mu)^2] = -\frac{(4L+1)^2}{2\cdot 4^{2L}}\binom{2L}{L}^2 - \frac{4(L+1)}{4^L}\binom{2L}{L} + 4L + \frac{9}{2} + o(1);$$

$$E[(X-\mu)^3] = \frac{(4L+1)^3}{2\cdot 4^{3L}}\binom{2L}{L}^3 + \frac{6(L+1)(4L+1)}{4^{2L}}\binom{2L}{L}^2 - \frac{(8L^2-6L-11/2)}{4^L}\binom{2L}{L} - 12(L+1) + o(1);$$

$$E[(X-\mu)^4] = -\frac{128L^3+320L^2+128L+1/2}{4^{2L}}\binom{2L}{L}^2 - \frac{48L^2+184L+112}{4^L}\binom{2L}{L} + 48L^2 + 160L + 225/2 + o(1).$$

</>The Maple command to find these moments is **`ForMoXMean(n1,r,L)`**. For example, for the result of $E[(X-\mu)^4]$ above, try **`ForMoXMean(0,4,L)`**.

Finally, the sequence (in r) of the rth standardized moments as L approaches infinity, i.e.,

$$\lim_{L\to\infty} \frac{E[(X-\mu)^r]}{E[(X-\mu)^2]^{r/2}},$$

is found to be

$$\left[0, 1, \frac{4-\pi}{(\pi-2)^{3/2}}, \frac{(3\pi-8)\pi}{(\pi-2)^2}, \dots\right].$$

This is a little disappointing as the sequence does not seem to fit any existing distributions (as far as we know). However, it confirms the non-normality of the distribution we observed earlier in the histograms of Figure 2.

References

[1] M. Aigner, G. Ziegler, *Proofs from the Book*, 6th Edition, Springer, Berlin, 2018.
[2] D. Bayer, P. Diaconis, Trailing the dovetail shuffle to its lair, *Ann. Appl. Probab.* **2** (1992) 294–313.
[3] M. Ciucu, No-feedback card guessing for dovetail shuffles, *Ann. Appl. Probab.* **8** (1998) 1251–1269.
[4] P. Diaconis, R. Graham, X. He, S. Spiro, Card guessing with partial feedback, *Combin. Probab. Comput.* **31** (2022) 1–20.
[5] P. Diaconis, R. Graham, S. Spiro, Guessing about guessing: Practical strategies for card guessing with feedback, *Amer. Math. Monthly.* **129** (2022) 607–622.
[6] E. Gilbert, *Theory of Shuffling*, Technical memorandum, Bell Laboratories, 1955.
[7] T. Krityakierne, T. A. Thanatipanonda, The Card Guessing Game: A generating function approach, *J. Symbolic Comput.* **115** (2023) 1–17.
[8] T. Krityakierne, T. A. Thanatipanonda, No feedback? No worries! The art of guessing the right card, *arXiv*:2205.08793 [math.CO], (2022).
[9] P. Liu, On card guessing game with one time riffle shuffle and complete feedback, *Discrete Appl. Math.* **288** (2021) 270–278.
[10] D. Zeilberger, Symbolic moment calculus I: Foundations and permutation pattern statistics, *Ann. Comb.* **8** (2004) 369–378.