Asymptotics for Lacunary Sums of Binomial Coefficients and a Card Problem with Ranks

Tamás Lengyel
Mathematics Department
Occidental College
1600 Campus Road
Los Angeles, CA 90041
USA
lengyel@oxy.edu

Abstract

We find asymptotics for lacunary sums of binomial coefficients. As an application we determine the exact and approximate probability that the first card has the uniquely highest rank among the top l cards of a standard card deck.

1 Introduction

In card games, we often pick the first i cards from the top of a standard deck and check the ranks of these cards as they show up. The kings have the highest rank, 13, while the aces have the lowest, 1. What are the chances, p_i, that the first card has the uniquely highest rank among these i cards? One expects this number to be slightly less than $1/i$ for small values of i. We will obtain some asymptotics for this probability in the generalized case in which there are m denominations (of face values 1, 2, $\ldots$, m) in each of k suits, i.e., $m = 13$ and $k = 4$ for the standard deck. We prove

Theorem 1. *Let i, k and m be positive integers and p_i as defined above. Then*

$$p_i \approx \frac{1}{i} - \frac{1}{2m}\left(1 - \frac{1}{k}\right), \tag{1}$$

for $2 \le i \le m^{1/4}$ as $m \to \infty$.

We can rephrase Theorem 1 and obtain Theorem 2 which is of general interest since it concerns the k-sected version of the so-called upper summation [3, p. 174] (sometimes also referred to as the hockey stick pattern, though apparently, there is some confusion regarding this naming convention).

Theorem 2. *Let i, k and m be positive integers. We have that*

$$\sum_{l=1}^{m} \binom{k(l-1)}{i-1} \sim \binom{km}{i} \frac{1}{k} \left(1 - \frac{i}{2m}\left(1 - \frac{1}{k}\right)\right) \tag{2}$$

for $2 \leq i \leq m^{1/4}$ as $m \to \infty$.

Some particular lacunary sums of this type have combinatorial interpretations (cf. Table 1), and they appear in Sloane's On-Line Encyclopedia of Integer Sequences [5].

k	i	A-number	Comments
2	3	A002412	hexagonal pyramidal numbers, or greengrocer's numbers
2	4	A112742	second derivative of the n-th Chebyshev polynomial (of the first kind) evaluated at $x = 1$
3	4	A116689	partial sums of dodecahedral numbers

Table 1: Lacunary sums of binomial coefficients

We can extend Theorem 2 for other k-sections using simple identities for binomial coefficients.

Theorem 3. *Let i, k and m be positive integers and j be an integer so that $0 \leq j < \min\{i, k\}$. We have that*

$$\sum_{l=1}^{m} \binom{k(l-1)+j}{i-1} \sim \binom{km+j}{i} \frac{1}{k} \left(1 - \frac{i}{2(m + \frac{j}{k})}\left(1 - \frac{1}{k}\right)\right) \tag{3}$$

for $2 \leq i \leq m^{1/4}$ as $m \to \infty$.

The approximation (2) (and subsequently, (3)) can be improved by involving more terms in the last factor as it is indicated immediately after (14) for approximating p_i. Note that the sum in (2) has a closed form by standard binomial summation if $k = 1$, and it corresponds to the upper summation identity. In this case, equality holds in (2) for all positive integers i and m. The same applies to (3).

2   Preliminaries

The simplest version of the card selection problem is proposed in [4, Example 2.3.3] with the standard deck and $i = 2$: "Two cards are drawn from a poker deck without replacement.

What is the probability that the second is higher in rank than the first?" We use the notation

$$
\begin{aligned}
A_i &= \text{the rank of first card exceeds that of the other } i-1 \text{ cards, so} \\
A_2 &= \text{the rank of first card exceeds that of the second card,} \\
B &= \text{the rank of second card exceeds that of the first card,} \\
C &= \text{the first and second cards have the same rank.}
\end{aligned}
$$

Clearly, $P(A_2) = P(B)$, by symmetry, and $P(C) = \frac{3}{51}$ which yields $P(A_2) = \frac{24}{51} = \frac{1}{2} - \frac{3}{102} \approx 0.4706$, while the approximation (1) yields $p_2 \approx \frac{1}{2} - \frac{1}{26}\left(1 - \frac{1}{4}\right) \approx 0.4712$.

By using conditional probabilities, we get that in general,

$$
\begin{aligned}
p_i = P(A_i) &= \sum_{l=1}^{m} P(A_i \mid D_l)P(D_l) = \sum_{l=1}^{m} \frac{\binom{k(l-1)}{i-1}}{\binom{km-1}{i-1}}\frac{k}{km} \\
&= \frac{k}{i}\frac{1}{\binom{km}{i}}\sum_{l=1}^{m}\binom{k(l-1)}{i-1},
\end{aligned}
\tag{4}
$$

with D_l being the event that the first card is of face value $l = 1, 2, \ldots, m$. Clearly, $p_1 = 1$ and $p_{k(m-1)+1} = 1/\binom{km}{k}$, since in the latter case, the first card and the last $k-1$ cards in the full deck of km cards are of the highest rank. Easy calculation shows that $p_2 = \frac{k(m-1)}{2(km-1)}$ but p_i looks less and less manageable as i gets large. Also, it is obvious that $p_i = 1/i$ for $k = 1$.

We applied symbolic summation techniques to rewrite (2), (3) and (4) but did not succeed in obtaining useful closed forms. Note, however, that we can derive the generating function of the last sum in (4) in the form of

$$
g(x) = \sum_{i=1}^{k(m-1)+1} p_i \frac{i}{k}\binom{km}{i}x^i = \sum_{i=1}^{k(m-1)+1}\sum_{l=1}^{m}\binom{k(l-1)}{i-1}x^i
$$

through the generating function in two variables

$$
\begin{aligned}
f(x,y) &= \sum_{i=1}^{k(m-1)+1}\sum_{l=1}^{m}\binom{k(l-1)}{i-1}x^i y^l = \sum_{l=1}^{m} y^l \sum_{i=1}^{k(l-1)+1}\binom{k(l-1)}{i-1}x^i \\
&= \sum_{l=1}^{m} y^l x(1+x)^{k(l-1)} = xy\frac{1-\left(y(1+x)^k\right)^m}{1-y(1+x)^k}.
\end{aligned}
$$

In fact, it follows that $g(x) = f(x,1) = x\frac{1-(1+x)^{km}}{1-(1+x)^k}$. However, this does not seem to help with the asymptotics.

3 Proof of Theorem 1

Clearly, $p_1 = 1$, and it is easy to see that $p_i = \frac{1}{i} - \frac{1}{2m}\left(1 - \frac{1}{k}\right) + \frac{(i-1)\left(k^2 - 6k + 5\right)}{12\,k^2 m^2} + O(\frac{1}{m^3})$ for $2 \le i \le 8$ as $m \to \infty$, in agreement with (1).

By standard asymptotical expansion of the binomial coefficients, we have that

$$
\binom{n}{k} = \frac{n^k}{k!}\left(1 - \frac{\binom{k}{2}}{n} + \frac{(3k-1)\binom{k}{3}}{4n^2} + O\left(\frac{k^6}{n^3}\right)\right),
\tag{5}
$$

which is useful when $k = O\left(n^{1/2-\epsilon}\right)$ with any small positive ϵ. We will use the inequality

$$\binom{n}{k} \leq \frac{n^k}{k!}, \tag{6}$$

which holds for every $k, 0 \leq k \leq n$, in the second part of the estimation (13). On the other hand,

$$\begin{aligned}\sum_{l=1}^{m}(l-1)^i &= \frac{1}{i+1}\sum_{l=0}^{i}\binom{i+1}{l}B_l m^{i+1-l} \\ &= \frac{1}{i+1}m^{i+1} - \frac{1}{2}m^i + \frac{i}{12}m^{i-1} + O\left(i^2 m^{i-2}\right)\end{aligned} \tag{7}$$

with $i = o\left(m\right)$ by using the Bernoulli numbers, B_l [3, Section 6.5]; for example, $B_0 = 1, B_1 = -\frac{1}{2}, B_2 = \frac{1}{6}, B_3 = 0, B_4 = -\frac{1}{30}, B_5 = 0$, and $B_6 = \frac{1}{42}$. We note that an alternative approach for approximating sums of powers is outlined in [1] based on an observation of Bernoulli which is related to the Faulhaber polynomials [2]

$$\sum_{l=1}^{m}(l-1)^i = \frac{m_1^{i+1}}{i+1}\left(1 - \frac{\binom{i+1}{2}}{12m_1^2} + \frac{7\binom{i+1}{4}}{240m_1^4} - \ldots\right) \quad \text{with } m_1 = m - \frac{1}{2}. \tag{8}$$

For an application of this technique see [6]. Note that while in the expansion (7) the fourth term vanishes since $B_3 = 0$, in the alternative approach the second term vanishes already. This fact might help in getting the asymptotic evaluations below a little faster, although by somewhat more complicated expansions. This approach can be easily tested by some software with symbolic calculation capabilities. We achieved this by using the *Mathematica* function `FullSimplify` with the assumption that the exponents in the terms of the form $(m - \frac{1}{2})^i$ are positive integers. In fact, to get the approximation (1), we need only

$$\frac{1}{m^i}\left(1 + \frac{\binom{i}{2}}{km}\right)\left(\frac{(m-\frac{1}{2})^i}{i} - \frac{\binom{i-1}{2}}{k}\frac{(m-\frac{1}{2})^{i-1}}{i-1}\right),$$

while our main approach requires an extra term from the expansion (7) (cf. (11))

$$\frac{1}{m^i}\left(1 + \frac{\binom{i}{2}}{km}\right)\left(m^i\left(\frac{1}{i} - \frac{1}{2m}\right) - \frac{\binom{i-1}{2}}{k}m^{i-1}\frac{1}{i-1}\right)$$

to obtain (14). In order to determine the next term (15) of the approximation, we need 4 and 6 terms of the expansions (8) and (7), respectively. We note, however, that in both approaches, every other term vanishes after the initial terms.

The good news is that the dominant terms in the sum (4) have large values of l in the binomial coefficients for which we can use the asymptotics given in (5). This fact will make the approximation surprisingly effective even for fairly large values of i.

Let i be any integer between 9 and the sufficiently large $m^{1/4}$, and $a(i) = \lceil \frac{i^{8/3}}{k} \rceil + 1$. This choice guarantees

$$\left(\frac{a(i)}{m}\right)^i = O\left(\left(\frac{1}{km^{1/3}}\right)^i\right) = O\left(\frac{1}{m^3}\right). \tag{9}$$

4

Note that as $m \to \infty$, formulas (7) and (9) yield the asymptotics

$$\sum_{l=1}^{a(i)-1} (l-1)^{i-1} = a(i)^i \left(\frac{1}{i} - \frac{1}{2a(i)} + \frac{i-1}{12a(i)^2} + O\left(\frac{i^2}{a(i)^3}\right) \right) \tag{10}$$

and

$$
\begin{aligned}
\sum_{l=a(i)}^{m} (l-1)^{i-1} &= \sum_{l=1}^{m}(l-1)^{i-1} - \sum_{l=1}^{a(i)-1}(l-1)^{i-1} \\
&= m^i \left(\frac{1}{i} - \frac{1}{2m} + \frac{i-1}{12m^2} + O\left(\frac{i^2}{m^3}\right) \right) + m^i \left(\frac{a(i)}{m}\right)^i O\left(\frac{1}{i}\right) \\
&= m^i \left(\frac{1}{i} - \frac{1}{2m} + \frac{i-1}{12m^2} + O\left(\frac{i^2}{m^3}\right) \right) + m^i O\left(\frac{1}{im^3}\right).
\end{aligned}
\tag{11}
$$

The identity (5) implies that

$$
\begin{aligned}
\binom{km}{i} &= \frac{(km)^i}{i!}\left(1 - \frac{\binom{i}{2}}{km} + \frac{(3i-1)\binom{i}{3}}{4(km)^2} + O\left(\frac{i^6}{(km)^3}\right) \right), \\
\frac{1}{\binom{km}{i}} &= \frac{i!}{(km)^i}\left(1 + \frac{\binom{i}{2}}{km} + \frac{(3i-2)\binom{i+1}{3}}{4(km)^2} + O\left(\frac{i^6}{(km)^3}\right) \right).
\end{aligned}
\tag{12}
$$

According to (4), (5), (6), (9), (10) and (12), and with $A = \frac{k}{i}\frac{1}{\binom{km}{i}}$, we get that

$$
\begin{aligned}
p_i &= A \sum_{l=1}^{m} \binom{k(l-1)}{i-1} \\
&= A \left(\sum_{l=a(i)}^{m} \binom{k(l-1)}{i-1} + \frac{k^{i-1}}{i!} O(a(i)^i) \right) \\
&= \frac{k}{i}\frac{i!}{(km)^i}\left(1 + \frac{\binom{i}{2}}{km} + \frac{(3i-2)\binom{i+1}{3}}{4(km)^2} + O\left(\frac{i^6}{(km)^3}\right) \right) \\
&\quad \times \sum_{l=a(i)}^{m} \frac{k^{i-1}(l-1)^{i-1}}{(i-1)!}\left(1 - \frac{\binom{i-1}{2}}{k(l-1)} + \frac{(3i-4)\binom{i-1}{3}}{4(k(l-1))^2} + O\left(\frac{i^6}{(l-1)^3}\right) \right) + O\left(\frac{1}{im^3}\right),
\end{aligned}
\tag{13}
$$

for $a(i) \le l$ implies that $i \le (k(l-1))^{3/8}$, and thus, we can use (5) in the last part of (13). Now by (11), we derive that

$$
\begin{aligned}
p_i &= \frac{k}{i}\frac{i!}{(km)^i}\left(1 + \frac{\binom{i}{2}}{km} + \frac{(3i-2)\binom{i+1}{3}}{4(km)^2} + O\left(\frac{i^6}{(km)^3}\right) \right) \frac{k^{i-1}}{(i-1)!} \\
&\quad \times \Big(\left(\frac{1}{i}m^i - \frac{1}{2}m^{i-1} + \frac{i-1}{12}m^{i-2} \right) - \frac{\binom{i-1}{2}}{k}\left(\frac{1}{i-1}m^{i-1} - \frac{1}{2}m^{i-2} + \frac{i-2}{12}m^{i-3} \right) \\
&\quad + \frac{(3i-4)\binom{i-1}{3}}{4k^2}\left(\frac{1}{i-2}m^{i-2} - \frac{1}{2}m^{i-3} + \frac{i-3}{12}m^{i-4} \right) \Big) + O\left(\frac{i^5}{m^3}\right) + O\left(\frac{1}{im^3}\right)
\end{aligned}
$$

which can be rewritten as

$$
\begin{aligned}
&\left(1 + \frac{\binom{i}{2}}{km} + \frac{(3i-2)\binom{i+1}{3}}{4(km)^2} \right)\left(\left(\frac{1}{i} - \frac{1}{2m} + \frac{i-1}{12m^2}\right) - \frac{\binom{i-1}{2}}{k}\left(\frac{1}{(i-1)m} - \frac{1}{2m^2} \right) \right. \\
&\left. + \frac{(3i-4)\binom{i-1}{3}}{4k^2}\left(\frac{1}{(i-2)m^2} - \frac{1}{2m^3} \right) \right) + O\left(\frac{i^5}{m^3}\right) \\
&= \frac{1}{i} - \frac{1}{2m}\left(1 - \frac{1}{k} \right) + O\left(\frac{i}{m^2}\right).
\end{aligned}
\tag{14}
$$

$\square$

The error term in (14) can be further determined to yield

$$\frac{(i-1)\,(k^2 - 6\,k + 5)}{12\,k^2 m^2} + O(\frac{1}{m^3}) \tag{15}$$

as $m \to \infty$ and i is kept fixed. With more involved calculations we find that the next error term is $\frac{(i-1)\,(k-1)(3(k-3)\,i-4\,k+8)}{24\,k^3 m^3} + O(\frac{1}{m^4})$ which explains why the previous approximation becomes even better if $k = 3$ and m is sufficiently large. Note that if $k = 1$ then the second term of (1) and the above mentioned two terms vanish as one might expect since $p_i = 1/i$ in this case.

Numerical calculation and plotting p_i against i confirms that the approximation (1) is valid for a much larger set of values i. (Of course, this observation is of no surprise as in the above proof we dealt with the bounds on i quite generously. We extend the range of values of i in the next section.) For example, in the case of the standard deck, we get that $p_i \approx \frac{1}{i} - 0.029$, in agreement with the fact that second term in (1) does not depend on i, and it seems to be a fairly precise over-estimation for any $2 \le i \le 22$. In general, however, numerical experimentation also suggests that the approximation (1) needs correcting terms when i grows.

4   Extending the range

The proof presented in the previous section can be improved to work for a larger set of values of i; thus, we can extend the validity of Theorems 1-3. In fact, let i be any integer between $9/(1 - \epsilon)$ and the sufficiently large $m^{1/3}$, and $a(i) = \lceil \frac{i^{2+\epsilon}}{k} \rceil + 1$ with any small positive ϵ. We note that $a(i) \le l$ implies that $i \le (k(l-1))^{\frac{1}{2+\epsilon}}$; thus, we can use the approximation (5) for binomial coefficients. We also need $9/(1 - \epsilon) \le k$ to guarantee (9). After streamlining the steps outlined in (10)-(14), we obtain a complete proof. First, we observe that

$$\binom{km}{i} = \frac{(km)^i}{i!}\left(1 - \frac{\binom{i}{2}}{km} + O\left(\frac{i^4}{(km)^2}\right)\right),$$

$$\frac{1}{\binom{km}{i}} = \frac{i!}{(km)^i}\left(1 + \frac{\binom{i}{2}}{km} + O\left(\frac{i^4}{m^2}\right)\right).$$

It follows that

$$\begin{aligned}
p_i &= A \sum_{l=1}^{m} \binom{k(l-1)}{i-1} \\
&= A\left(\sum_{l=a(i)}^{m} \binom{k(l-1)}{i-1} + \frac{k^{i-1}}{i!}O(a(i)^i)\right) \\
&= \frac{k}{i}\frac{i!}{(km)^i}\left(1 + \frac{\binom{i}{2}}{km} + O\left(\frac{i^4}{m^2}\right)\right) \\
&\quad \times \sum_{l=a(i)}^{m} \frac{k^{i-1}(l-1)^{i-1}}{(i-1)!}\left(1 - \frac{\binom{i-1}{2}}{k(l-1)} + O\left(\frac{i^4}{(l-1)^2}\right)\right) + O\left(\frac{1}{im^3}\right),
\end{aligned}$$

therefore,

$$\begin{aligned}
p_i &= \frac{k}{i}\frac{i!}{(km)^i}\left(1 + \frac{\binom{i}{2}}{km} + O\left(\frac{i^4}{m^2}\right)\right)\frac{k^{i-1}}{(i-1)!} \\
&\quad \times\left(\left(\frac{1}{i}m^i - \frac{1}{2}m^{i-1} + \frac{i-1}{12}m^{i-2}\right) - \frac{\binom{i-1}{2}}{k}\left(\frac{1}{i-1}m^{i-1} - \frac{1}{2}m^{i-2}\right)\right. \\
&\quad \left. + O(i^4)\frac{1}{i-2}m^{i-2}\right) + O\left(\frac{i^5}{m^3}\right) + O\left(\frac{1}{im^3}\right).
\end{aligned}$$

6

This can be rewritten as

$$\left(1 + \frac{\binom{i}{2}}{km}\right)\left(\left(\frac{1}{i} - \frac{1}{2m}\right) - \frac{\binom{i-1}{2}}{k}\frac{1}{(i-1)m}\right) + O\left(\frac{i^3}{m^2}\right)$$
$$= \frac{1}{i} - \frac{1}{2m}\left(1 - \frac{1}{k}\right) + O\left(\frac{i^3}{m^2}\right).$$

$\square$

By including more correcting terms, one might be able to extend the range for i to $m^{1/2-\epsilon}$ with any positive ϵ.

Acknowledgments

References

[1] B. L. Burrows and R. F. Talbot, Sums of powers of integers, *Amer. Math. Monthly* **91** (1984), 394–403.

[2] A. W. F. Edwards, A quick route to sums of powers, *Amer. Math. Monthly* **93** (1986), 451–455.

[3] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics, A foundation for computer science*, Addison-Wesley, Reading, MA, 1994, 2nd edition.

[4] M. Larsen and M. Marx, *An Introduction to Mathematical Statistics and Its Applications*, Prentice Hall, Upper Saddle River, NJ, 2006, 4th edition.

[5] N. J. A. Sloane, *The On-Line Encyclopedia of Integer Sequences*, published electronically at http://www.research.att.com/~njas/sequences.

[6] B. Tsaban, Bernoulli numbers and the probability of a birthday surprise, *Discrete Appl. Math.* **127** (2003), 657–663.

(Concerned with sequences A002412, A112742, and A116689.)

Return to Journal of Integer Sequences home page.