The Proportion of Comets in the Card Game SET

April 13, 2018

Dan Swenson, Black Hills State University

The game of SET

- The SET deck may be represented as $D = \mathbb{Z}_3^4$

The game of SET

- The SET deck may be represented as $D = \mathbb{Z}_3^4$
- Four dimensions: Number, Color, Shape, Shading

The game of SET

- The SET deck may be represented as $D = \mathbb{Z}_3^4$
- Four dimensions: Number, Color, Shape, Shading
- A set is a collection of 3 cards, whose sum is the zero vector

The game of SET

- The SET deck may be represented as $D = \mathbb{Z}_3^4$
- Four dimensions: Number, Color, Shape, Shading
- A set is a collection of 3 cards, whose sum is the zero vector
- (Such a collection is a line in the four-dimensional affine space D)

The game of SET

- The SET deck may be represented as $D = \mathbb{Z}_3^4$
- Four dimensions: Number, Color, Shape, Shading
- A set is a collection of 3 cards, whose sum is the zero vector
- (Such a collection is a line in the four-dimensional affine space D)
- Two cards uniquely determine a line (set):

$$A + B + C = 0 \iff C = -A - B$$

The game of SET

- The SET deck may be represented as $D = \mathbb{Z}_3^4$
- Four dimensions: Number, Color, Shape, Shading
- A set is a collection of 3 cards, whose sum is the zero vector
- (Such a collection is a line in the four-dimensional affine space D)
- Two cards uniquely determine a line (set):

$$A + B + C = 0 \iff C = -A - B$$

Fact: If A and B are distinct, then A, B, and $(-A - B)$ are all distinct.

According to [1], a comet is a collection of 9 cards which sum to the zero vector.

According to [1], a comet is a collection of 9 cards which sum to the zero vector.

“It follows that for every eight cards, there is a unique ninth card that makes a comet. However, it may happen that the ninth card needed to make a comet is one of the eight cards already present.

According to [1], a comet is a collection of 9 cards which sum to the zero vector.

"It follows that for every eight cards, there is a unique ninth card that makes a comet. However, it may happen that the ninth card needed to make a comet is one of the eight cards already present.
"This leads to the interesting open question: What is the probability that nine cards drawn at random are a comet?"

According to [1], a comet is a collection of 9 cards which sum to the zero vector.

“It follows that for every eight cards, there is a unique ninth card that makes a comet. However, it may happen that the ninth card needed to make a comet is one of the eight cards already present.
“This leads to the interesting open question: What is the probability that nine cards drawn at random are a comet?”

The denominator is the number of 9-card subsets of the deck:

$$\binom{81}{9} = 260,887,834,350$$

According to [1], a comet is a collection of 9 cards which sum to the zero vector.

"It follows that for every eight cards, there is a unique ninth card that makes a comet. However, it may happen that the ninth card needed to make a comet is one of the eight cards already present.
"This leads to the interesting open question: What is the probability that nine cards drawn at random are a comet?"

The denominator is the number of 9-card subsets of the deck:

$$\binom{81}{9} = 260,887,834,350$$

Suffices to count the number of those which sum to 0.

$$\sum_{x \in S} x = 0$$

Some code

```
import itertools

all_cards = [p for p in itertools.product(range(3),
                                          repeat = 4)]

for k in xrange(10):
    print sum(all(sum(x[i] for x in p)%3 == 0
        for i in xrange(4))
        for p in itertools.combinations(all_cards, k))
```

OK, this is taking a while. . .

k	subsets of size k that sum to 0	running time (seconds)
0	1	0.01
1	1	0.01
2	40	0.02
3	1080	0.23
4	20540	3.52
5	316316	57.48
6	4007016	787.14

OK, this is taking a while. . .

k	subsets of size k that sum to 0	running time (seconds)
0	1	0.01
1	1	0.01
2	40	0.02
3	1080	0.23
4	20540	3.52
5	316316	57.48
6	4007016	787.14

(I gave up before I got to $k = 9$.)

THE ON-LINE ENCYCLOPEDIA OF INTEGER SEQUENCES®

founded in 1964 by N. J. A. Sloane

40, 1080, 20540, 316316 Search Hints

(Greetings from The On-Line Encyclopedia of Integer Sequences!)

Search: **seq:40,1080,20540,316316**

Sorry, but the terms do not match anything in the table.

If your sequence is of general interest, please submit it using the form provided and it will (probably) be added to the OEIS! Include a brief description and if possible enough terms to fill 3 lines on the screen. We need at least 4 terms.

Search completed in 0.004 seconds

- Idea: if you just wanted an estimate, you could try taking 9-card subsets at random (Monte Carlo estimation)

- Idea: if you just wanted an estimate, you could try taking 9-card subsets at random (Monte Carlo estimation)
- These simulations suggest that the proportion might be around 0.01234? Maybe 0.012345? Is the next digit a 6?

- Idea: if you just wanted an estimate, you could try taking 9-card subsets at random (Monte Carlo estimation)
- These simulations suggest that the proportion might be around 0.01234? Maybe 0.012345? Is the next digit a 6?
- What number would this be?

- Idea: if you just wanted an estimate, you could try taking 9-card subsets at random (Monte Carlo estimation)
- These simulations suggest that the proportion might be around 0.01234? Maybe 0.012345? Is the next digit a 6?
- What number would this be?

$0.01234567\ldots$

- Idea: if you just wanted an estimate, you could try taking 9-card subsets at random (Monte Carlo estimation)
- These simulations suggest that the proportion might be around 0.01234? Maybe 0.012345? Is the next digit a 6?
- What number would this be?

$$0.01234567\ldots = \frac{1}{100} + \frac{2}{1000} + \frac{3}{10000} + \ldots$$

- Idea: if you just wanted an estimate, you could try taking 9-card subsets at random (Monte Carlo estimation)
- These simulations suggest that the proportion might be around 0.01234? Maybe 0.012345? Is the next digit a 6?
- What number would this be?

$$0.01234567\ldots = \frac{1}{100} + \frac{2}{1000} + \frac{3}{10000} + \ldots = \frac{1}{100} \sum_{k=0} k \left(\frac{1}{10}\right)^{k-1}$$

- Idea: if you just wanted an estimate, you could try taking 9-card subsets at random (Monte Carlo estimation)
- These simulations suggest that the proportion might be around 0.01234? Maybe 0.012345? Is the next digit a 6?
- What number would this be?

$$0.01234567\ldots = \frac{1}{100}+\frac{2}{1000}+\frac{3}{10000}+\ldots = \frac{1}{100}\sum_{k=0} k\left(\frac{1}{10}\right)^{k-1}$$

…(use Taylor series or something) …

- Idea: if you just wanted an estimate, you could try taking 9-card subsets at random (Monte Carlo estimation)
- These simulations suggest that the proportion might be around 0.01234? Maybe 0.012345? Is the next digit a 6?
- What number would this be?

$$0.01234567\ldots = \frac{1}{100} + \frac{2}{1000} + \frac{3}{10000} + \ldots = \frac{1}{100} \sum_{k=0} k \left(\frac{1}{10}\right)^{k-1}$$

...(use Taylor series or something) ...

$$= \frac{1}{81}$$

Heuristically, the answer should be close to $\frac{1}{81}$, but maybe not equal.

Heuristically, the answer should be close to $\frac{1}{81}$, but maybe not equal.

If we sampled the cards with replacement then the ninth card would have probability $\frac{1}{81}$ of being the one that we want.

Heuristically, the answer should be close to $\frac{1}{81}$, but maybe not equal.

If we sampled the cards with replacement then the ninth card would have probability $\frac{1}{81}$ of being the one that we want.

"However, [since we draw cards without replacement] it may happen that the ninth card needed to make a comet is one of the eight cards already present."

Heuristically, the answer should be close to $\frac{1}{81}$, but maybe not equal.

If we sampled the cards with replacement then the ninth card would have probability $\frac{1}{81}$ of being the one that we want.

"However, [since we draw cards without replacement] it may happen that the ninth card needed to make a comet is one of the eight cards already present."

Suffices to count the 8-card subsets which already do contain the "required" card:

$$\left(-\sum_{x \in S} x\right) \in S$$

OK, let's say S is a collection of 8 cards, and $d = \left(-\sum_{x \in S} x\right) \in S$. In fact, let's say d is the 8th card:

OK, let's say S is a collection of 8 cards, and $d = \left(-\sum_{x \in S} x\right) \in S$. In fact, let's say d is the 8th card:

$$\sum_{x \in S} x = x_1 + x_2 + \ldots + x_7 + x_8$$

$$-x_8 = x_1 + x_2 + \ldots + x_7 + x_8$$

$$-2x_8 = x_1 + x_2 + \ldots + x_7$$

$$x_8 = x_1 + x_2 + \ldots + x_7$$

OK, let's say S is a collection of 8 cards, and $d = \left(-\sum_{x \in S} x\right) \in S$. In fact, let's say d is the 8th card:

$$\sum_{x \in S} x = x_1 + x_2 + \ldots + x_7 + x_8$$

$$-x_8 = x_1 + x_2 + \ldots + x_7 + x_8$$

$$-2x_8 = x_1 + x_2 + \ldots + x_7$$

$$x_8 = x_1 + x_2 + \ldots + x_7$$

So, the first 7 cards must not have contained their sum.

OK, let's say S is a collection of 8 cards, and $d = \left(-\sum_{x \in S} x\right) \in S$. In fact, let's say d is the 8th card:

$$\sum_{x \in S} x = x_1 + x_2 + \ldots + x_7 + x_8$$

$$-x_8 = x_1 + x_2 + \ldots + x_7 + x_8$$

$$-2x_8 = x_1 + x_2 + \ldots + x_7$$

$$x_8 = x_1 + x_2 + \ldots + x_7$$

So, the first 7 cards must not have contained their sum. Suffices to count the 7-card subsets which do contain their sum:

$$\left(\sum_{x \in S} x\right) \in S$$

OK, let's say S is a collection of 7 cards, and $d = \left(\sum_{x \in S} x\right) \in S$. In fact, let's say d is the 7th card:

OK, let's say S is a collection of 7 cards, and $d = \left(\sum_{x \in S} x\right) \in S$. In fact, let's say d is the 7th card:

$$\sum_{x \in S} x = x_1 + x_2 + \ldots + x_6 + x_7$$

$$x_7 = x_1 + x_2 + \ldots + x_6 + x_7$$

$$0 = x_1 + x_2 + \ldots + x_6$$

OK, let's say S is a collection of 7 cards, and $d = \left(\sum_{x \in S} x\right) \in S$. In fact, let's say d is the 7th card:

$$\sum_{x \in S} x = x_1 + x_2 + \ldots + x_6 + x_7$$

$$x_7 = x_1 + x_2 + \ldots + x_6 + x_7$$

$$0 = x_1 + x_2 + \ldots + x_6$$

So, the first 6 cards must have summed to 0.

OK, let's say S is a collection of 7 cards, and $d = \left(\sum_{x \in S} x\right) \in S$. In fact, let's say d is the 7th card:

$$\sum_{x \in S} x = x_1 + x_2 + \ldots + x_6 + x_7$$

$$x_7 = x_1 + x_2 + \ldots + x_6 + x_7$$

$$0 = x_1 + x_2 + \ldots + x_6$$

So, the first 6 cards must have summed to 0.

To count the 9-card subsets summing to 0, it suffices to count the 6-card subsets summing to 0. (!!)

OK, let's say S is a collection of 7 cards, and $d = \left(\sum_{x \in S} x\right) \in S$. In fact, let's say d is the 7th card:

$$\sum_{x \in S} x = x_1 + x_2 + \ldots + x_6 + x_7$$

$$x_7 = x_1 + x_2 + \ldots + x_6 + x_7$$

$$0 = x_1 + x_2 + \ldots + x_6$$

So, the first 6 cards must have summed to 0.

To count the 9-card subsets summing to 0, it suffices to count the 6-card subsets summing to 0. (!!)

(Some bookkeeping remains: have to deal with over-counting, *etc.*)

The Recurrence Relations

Define

$$A_k = \left\{ S \subseteq D \middle| \left(\sum_{x \in S} x \right) \in S, \text{ and } |S| = k \right\},$$

$$B_k = \left\{ S \subseteq D \middle| \left(-\sum_{x \in S} x \right) \in S, \text{ and } |S| = k \right\},$$

$$C_k = \left\{ S \subseteq D \middle| \left(\sum_{x \in S} x \right) = 0, \text{ and } |S| = k \right\},$$

and let $a_k = |A_k|$, and $b_k = |B_k|$, and $c_k = |C_k|$.

The Recurrence Relations

Define

$$A_k = \left\{ S \subseteq D \middle| \left(\sum_{x \in S} x \right) \in S, \text{ and } |S| = k \right\},$$

$$B_k = \left\{ S \subseteq D \middle| \left(-\sum_{x \in S} x \right) \in S, \text{ and } |S| = k \right\},$$

$$C_k = \left\{ S \subseteq D \middle| \left(\sum_{x \in S} x \right) = 0, \text{ and } |S| = k \right\},$$

and let $a_k = |A_k|$, and $b_k = |B_k|$, and $c_k = |C_k|$.

(We originally wanted c_9.)

The Recurrence Relations

Theorem

$$a_{k+1} = (c_k)(81 - k)$$

$$b_{k+1} = \binom{81}{k} - a_k$$

$$c_{k+1} = \frac{\binom{81}{k} - b_k}{k+1}$$

□

Results

k	a_k	b_k	c_k
0	0	0	1
1	81	1	1
2	80	0	40
3	3160	3160	1080
4	84240	82160	20540
5	1581580	1579500	316316
6	24040016	24040016	4007016
7	300526200	300500200	42928600
8	3176716400	3176690400	397089550
9	28987537150	28987537150	3220840350

In particular,

$$\frac{c_9}{\binom{81}{9}} = \frac{3220840350}{260887834350} = \frac{550571}{44596211} \approx 0.01234569$$

Results

k	a_k	b_k	c_k
0	0	0	1
1	81	1	1
2	80	0	40
3	3160	3160	1080
4	84240	82160	20540
5	1581580	1579500	316316
6	24040016	24040016	4007016
7	300526200	300500200	42928600
8	3176716400	3176690400	397089550
9	28987537150	28987537150	3220840350

In particular,

$$\frac{1}{81} \neq \frac{c_9}{\binom{81}{9}} = \frac{3220840350}{260887834350} = \frac{550571}{44596211} \approx 0.01234569$$

Generalizations

Let V be a finite vector space in characteristic $p > 0$, and suppose you want to count the number of k-subsets of V which sum to 0.

We can do this in terms of $(k-1)$-subsets of V which satisfy some other condition.

Generalizations

Let V be a finite vector space in characteristic $p > 0$, and suppose you want to count the number of k-subsets of V which sum to 0.

We can do this in terms of $(k-1)$-subsets of V which satisfy some other condition. (…And so on.)

Generalizations

Let V be a finite vector space in characteristic $p > 0$, and suppose you want to count the number of k-subsets of V which sum to 0.

We can do this in terms of $(k-1)$-subsets of V which satisfy some other condition. (…And so on.)

$$_nA_k = \left\{ S \subseteq V \middle| n\left(\sum_{v\in S} v\right) \in S, \text{ and } |S| = k \right\}$$

Generalizations

Let V be a finite vector space in characteristic $p > 0$, and suppose you want to count the number of k-subsets of V which sum to 0.

We can do this in terms of $(k-1)$-subsets of V which satisfy some other condition. (...And so on.)

$$ {}_nA_k = \left\{ S \subseteq V \middle| n\left(\sum_{v \in S} v\right) \in S, \text{ and } |S| = k \right\} $$

$$ {}_\infty A_k = \left\{ S \subseteq V \middle| \left(\sum_{v \in S} v\right) = 0, \text{ and } |S| = k \right\}. $$

Generalizations

Let V be a finite vector space in characteristic $p > 0$, and suppose you want to count the number of k-subsets of V which sum to 0.

We can do this in terms of $(k-1)$-subsets of V which satisfy some other condition. (...And so on.)

$$ {}_nA_k = \left\{ S \subseteq V \middle| n \left(\sum_{v \in S} v \right) \in S, \text{ and } |S| = k \right\} $$

$$ {}_\infty A_k = \left\{ S \subseteq V \middle| \left(\sum_{v \in S} v \right) = 0, \text{ and } |S| = k \right\}. $$

Let's write ${}_na_k = |{}_nA_k|$.

Theorem

For $k > 0$:

$$ {}_{\infty}a_k = \frac{\binom{|V|}{k-1} - {}_{(p-1)}a_{k-1}}{k}, $$

$$ {}_{1}a_k = {}_{\infty}a_{k-1}\left(|V| - (k-1)\right), $$

and for $n \notin \{1, \infty\}$,

$$ {}_{n}a_k = \binom{|V|}{k-1} - {}_{\left(\frac{n}{1-n}\right)}a_{k-1}. $$

Further, we have the initial values

$$ {}_{n}a_0 = \begin{cases} 0, & \textit{if } n \neq \infty \\ 1, & \textit{if } n = \infty \end{cases} $$

$\square$

Corollary

For $k > 0$ and $n \in \{1, \ldots, p-1, \infty\}$, we can write ${}_n a_k$ in terms of ${}_{(\frac{n}{1-n})} a_{k-1}$, where by $\frac{n}{1-n}$ we mean $n \cdot (1-n)^{-1} \pmod{p}$, and in particular we define $\frac{\infty}{1-\infty} = -1 = p-1$, and $\frac{1}{1-1} = \infty$. $\square$

This $\sigma(n) = \frac{n}{1-n}$ is a permutation on $\{1, \ldots, p-1, \infty\}$, and its order is p (it is a p-cycle.)

Corollary

For $k > 0$ and $n \in \{1, \ldots, p-1, \infty\}$, we can write ${}_n a_k$ in terms of ${}_{(\frac{n}{1-n})}a_{k-1}$, where by $\frac{n}{1-n}$ we mean $n \cdot (1-n)^{-1} \pmod{p}$, and in particular we define $\frac{\infty}{1-\infty} = -1 = p-1$, and $\frac{1}{1-1} = \infty$. $\square$

This $\sigma(n) = \frac{n}{1-n}$ is a permutation on $\{1, \ldots, p-1, \infty\}$, and its order is p (it is a p-cycle.)

To see this, notice that σ is a fractional linear transformation; it can be written as

$$\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$$

Corollary

For $k > 0$ and $n \in \{1, \ldots, p-1, \infty\}$, we can write ${}_n a_k$ in terms of ${}_{(\frac{n}{1-n})} a_{k-1}$, where by $\frac{n}{1-n}$ we mean $n \cdot (1-n)^{-1} \pmod{p}$, and in particular we define $\frac{\infty}{1-\infty} = -1 = p-1$, and $\frac{1}{1-1} = \infty$. $\square$

This $\sigma(n) = \frac{n}{1-n}$ is a permutation on $\{1, \ldots, p-1, \infty\}$, and its order is p (it is a p-cycle.)

To see this, notice that σ is a fractional linear transformation; it can be written as

$$\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$$

To calculate ${}_n a_k$, visit all the other ${}_m a_*$ sequences, and return to ${}_n a_*$ after exactly p steps; *i.e.*, at ${}_n a_{k-p}$.

Example

Suppose $V = \mathbb{Z}_7^3$*, and we wish to count the number of subsets of size* 18 *whose sum is* 0*. That is, we want* ${}_{\infty}a_{18}$*.*

Example

Suppose $V = \mathbb{Z}_7^3$, *and we wish to count the number of subsets of size* 18 *whose sum is* 0. *That is, we want* ${}_{\infty}a_{18}$.

We know $18 \equiv 4 \pmod 7$, so we will need to visit ${}_{\infty}a_4$ at some point. We calculate $\sigma^4(\infty) = 5$, so we want to start at ${}_5a_0$.

Example

Suppose $V = \mathbb{Z}_7^3$*, and we wish to count the number of subsets of size* 18 *whose sum is* 0*. That is, we want* ${}_{\infty}a_{18}$*.*

We know $18 \equiv 4 \pmod 7$, so we will need to visit ${}_{\infty}a_4$ at some point. We calculate $\sigma^4(\infty) = 5$, so we want to start at ${}_5a_0$.

The initial conditions say that ${}_ma_0 = 0$ unless $m = \infty$.

Example

Suppose $V = \mathbb{Z}_7^3$*, and we wish to count the number of subsets of size* 18 *whose sum is* 0*. That is, we want* ${}_{\infty}a_{18}$*.*

We know $18 \equiv 4 \pmod 7$, so we will need to visit ${}_{\infty}a_4$ at some point. We calculate $\sigma^4(\infty) = 5$, so we want to start at ${}_5a_0$.

The initial conditions say that ${}_m a_0 = 0$ unless $m = \infty$.

Next, $\sigma^{-1}(n) = \frac{n}{1+n}$, so we can calculate ${}_{\sigma^{-1}(5)}a_1 = {}_2a_1$ in terms of ${}_5a_0$.

Example

Suppose $V = \mathbb{Z}_7^3$*, and we wish to count the number of subsets of size* 18 *whose sum is* 0*. That is, we want* ${}_\infty a_{18}$*.*

We know $18 \equiv 4 \pmod 7$, so we will need to visit ${}_\infty a_4$ at some point. We calculate $\sigma^4(\infty) = 5$, so we want to start at ${}_5 a_0$.

The initial conditions say that ${}_m a_0 = 0$ unless $m = \infty$.

Next, $\sigma^{-1}(n) = \frac{n}{1+n}$, so we can calculate ${}_{\sigma^{-1}(5)} a_1 = {}_2 a_1$ in terms of ${}_5 a_0$.

Next we calculate ${}_3 a_2$, and ${}_6 a_3$, and ${}_\infty a_4$, and keep going until we get to ${}_\infty a_{18}$.

Example

Suppose $V = \mathbb{Z}_7^3$, and we wish to count the number of subsets of size 18 whose sum is 0. That is, we want ${}_{\infty}a_{18}$.

We know $18 \equiv 4 \pmod 7$, so we will need to visit ${}_{\infty}a_4$ at some point. We calculate $\sigma^4(\infty) = 5$, so we want to start at ${}_5a_0$.

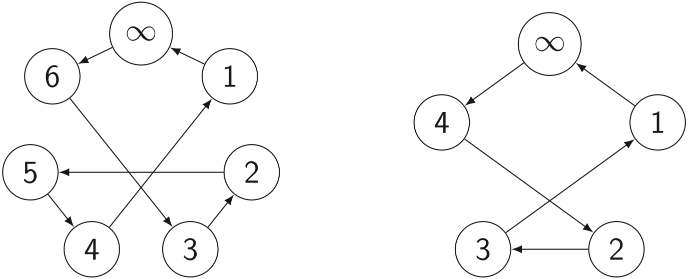
The initial conditions say that ${}_ma_0 = 0$ unless $m = \infty$.

Next, $\sigma^{-1}(n) = \frac{n}{1+n}$, so we can calculate ${}_{\sigma^{-1}(5)}a_1 = {}_2a_1$ in terms of ${}_5a_0$.

Next we calculate ${}_3a_2$, and ${}_6a_3$, and ${}_{\infty}a_4$, and keep going until we get to ${}_{\infty}a_{18}$.

This procedure allows us to calculate only the coefficients that we need.

Thank you!



(The permutation σ in characteristic 7 and characteristic 5)

References

[1] *Sets, Planets, and Comets*. Mark Baker, Jane Beltran, Jason Buell, Brian Conrey, Tom Davis, Brianna Donaldson, Jeanne Detorre-Ozeki, Leila Dibble, Tom Freeman, Robert Hammie, Julie Montgomery, Avery Pickford, and Justine Wong. *The College Mathematics Journal*, **Vol. 44**, No. 4 (September 2013), pp. 258-264