SET handout

A little introduction

First we want to think like mathematicians and generalize the game. We'll use n for the number of parameters, or attributes of the cards and v for the number of possible values of each of the parameters. So, for example, $n = 4$ and $v = 3$ for the commercially available game. We mostly keep v at 3, but consider the situation when $v \neq 3$ also. We'll use $\mathbb{F}_v^n$ to indicate the game with n attributes and v values.

There are two key facts about the game that allow us to answer many of the questions below.

Axiom 1. *There is exactly one card for each possible combination.*

Proposition 2. *Any two cards determine a SET.*

Proof. Given two cards, if they have the same value for one of the parameters, then the third card must have that same value for that parameter. Similarly, if the two cards have different values for one of the parameters, then the third card must have the third, unused value, for that parameter. Thus the value, in each parameter, for the third card is determined by the first two. □

Some Good Questions

1. Where did the game come from?

2. How many cards are in the game $\mathbb{F}_v^n$?

3. How many SETs are there that contain a given card in $\mathbb{F}_3^n$?

4. How many SETs are there in the game $\mathbb{F}_3^n$?

5. What is the maximum number of cards without a SET in the game $\mathbb{F}_3^n$?

6. What is the probability that any 3 cards form a SET in the game $\mathbb{F}_3^n$?

7. In the game $\mathbb{F}_3^4$ what is the probability of getting a SET with all parameters being all different? 3 parameters all different? 2 parameters all different? 1 parameter all different?

8. Prove that in the $\mathbb{F}_3^n$ game the following are equivalent:

    - Three cards form a SET.
    - The points in $\mathbb{F}_3^n$ corresponding to the cards sum to 0 mod 3.
    - The points in $\mathbb{F}_3^n$ corresponding to the cards are colinear.

9. Is the previous statement true for $\mathbb{F}_v^n$ when $v \neq 3$? What are the implications for the game $\mathbb{F}_v^n$ when $v \neq 3$?

10. What is the probability that given 12 (15, 18) cards that there is no SET in the game $\mathbb{F}_3^4$?

11. What is maximum number of SETs in 9 or 12 cards in the game $\mathbb{F}_3^4$?

12. A superset is a collection of 4 cards in the game $\mathbb{F}_3^4$ without a SET such that there is 1 card that can be added to the collection to form 2 SETs. What is the probability that 4 cards without a SET are a superset? Is it possible that 4 cards without a SET can be a superset in 2 different ways, e.g. there are 2 different cards that can be added that both form 2 sets when added to the collection?

13. What is the number of cards that can be left at the end of the $\mathbb{F}_3^4$ game?

14. For all of the questions about the $\mathbb{F}_3^4$ game can you generalize to $\mathbb{F}_3^n$?

Some Answers

1. To the best of my knowledge, it was created by Martha J. Falco in 1974. Martha was studying genetic traits of German shepherds and looking for patterns on cards. After enlisting her family in this endeavor, they wanted to make it fun and turned it into a game.

2. There are v cards in each of n categories, which gives v^n cards in the deck, since there is exactly one card for each possible combination (this turns out to be a key fact). So the commercially available game has 81 cards.

3. We do this first for $\mathbb{F}_3^4$ and then for $\mathbb{F}_3^n$. Given a particular card, any of the 80 other cards can be combined with it to make a SET. Then the third card is fixed. However, we over counted the second 2 cards by a factor of 2 (we can rearrange 2 cards in 2 ways), so the total number of SETs including a given card is $80/2 = 40$. For $\mathbb{F}_3^n$ the same argument works and we just have $(3^n - 1)/2$ SETs.

4. We can count this in any of 3 ways. We do each first for $\mathbb{F}_3^4$ and then for $\mathbb{F}_3^n$.

   (a) Using the previous answer, we know that for a given card it is contained in 40 SETs. There are 81 cards that could be that given card, giving $81 * 40$, but we've over counted by a factor of 3 (one way to think about this is that there are 3 places this card could have appeared at this point, and we have already taken care of the different ways to arrange the other two cards), so the total number is $81 * 40/3 = 1080$. So for $\mathbb{F}_3^n$ the number is $3^n(3^n - 1)/6$, which will be the same as in part (c).

   (b) Alternatively we could use the fact that $\binom{81}{2}$ is the number of ways of choosing 2 cards without regard to order. Then the third card is fixed, but 3 places it could occur (just as in part a). Thus the number of SETs is $\binom{81}{2}/3 = 1080$. So for $\mathbb{F}_3^n$ it is $\binom{3^n}{2}/3$.

(c) Finally, there are 81 choices for the first card and 80 choices for the second and then the third is fixed. But we over counted by the number of ways to rearrange 3 objects, which is $3! = 6$, so the number of SETs is $\frac{81*80*1}{6}$. So for $\mathbb{F}_3^n$ it is $3^n(3^n - 1)/6$.

5. Here are the answers for those dimensions for which this is known and information about the proofs.

   - $n = 1$ is 2: This is easy, the 3 cards in this game form a SET and 2 don't.
   - $n = 2$ is 4: This requires some work, but is not too hard. I know of at least 2 proofs. A sketch of one possible argument is to show a collection of 4 cards with no SET and then suppose there was such with 5 and use the pigeon hole principle to argue a contradiction. An alternate is to prove that all possible collections of 4 cards with no SET are the same up to affine transformation and then do an exhaustive search through the possible 5 point collections. A good reference on this topic is [DM].
   - $n = 3$ is 9: [DM] has two proofs of this. Both exhibit a collection of 9 and then argue that there cannot be 10. The first uses the pigeon hole principle and the second counts the same thing twice (classic combinatorics).
   - $n = 4$ is 20: [P] is the original source from 1971 and is in Italian. More recently [DM] give a more accessible, but still quite challenging proof that is like their second proof for $n = 3$.
   - $n = 5$ is 43: The only source I know of is [EFS]. This is a research level paper. The idea is to get narrow enough bounds on the possible numbers and then complete an exhaustive computer search.

6. For $\mathbb{F}_3^4$ this probability is $1/79$. This can be counted in two ways. Given any 2 cards there is only 1 of the remaining 79 cards that completes this set. Alternatively, we know the total number of SETs in the deck — 1080 — and the total number of ways to draw 3 cards from the deck without regard to order is $\binom{81}{3}$ and $1080/85320 = 1/79$. So for the general game $\mathbb{F}_3^n$ this probability is $1/(3^n - 2)$.

7. First we note that what is given below is just ONE way to do this count. In fact it is now followed by a SECOND way to count it. Also, we note that both of these arguments give an entirely different way (beyond the 3 above) of computing the total number of SETs in the deck. Since these are all the possible types of SETs the sum should be the total number of SETs — notice that $108+324+432+216 = 1080$!

   - Only one parameter is all different: There are 4 choices for that one parameter and then 3 possible values for the other 3 parameters, giving $4 * 3^3 = 108$ SETs of this type in the Deck. Thus the probability is $108/1080 = .1$.
   - Two parameters all different: There are $\binom{4}{2}$ ways to choose the two parameters that are all different. For these two parameters, it is easiest to consider an example. First, suppose these two parameters are color and number, we can put red with 1, green with 1 or purple with 1 — one of the 3 colors must go

with 1. So there are 3 choices of which color goes with 1 and then we have only 2 two choices of the color that can be assigned to 2 and finally the color that goes with 3 is fixed. So we have 6 choices for the combination of these two parameters. We still have 3 choices for values of each of the other two parameters. Thus the total number of SETs of this type is $\binom{4}{2}(3\cdot 2)(3^2) = 324$. Hence the probability is $324/1080 = .3$

- Three parameters all different: There are $\binom{4}{3}$ ways to choose the three parameters that are all different. For these three parameters, the possibilities are counted much like those for the previous case. First, for two of the parameters there are $3 \cdot 2$ combinations we can make (same argument as above) and now we can think of this combination along with the third parameter as 2 "parameters" that can again, be combined in $3 \cdot 2$ ways. There are 3 choices for the value of the fourth parameter, giving us a total of is $\binom{4}{3}(3\cdot 2)^2(3) = 432$ total SETs of this type. Hence the probability is $432/1080 = .4$

- Four parameters all different: Following our argument above, we could note that there are now $\binom{4}{4} = 1$ ways to choose these 4 parameters (this helps if you want to generalize more) and the same argument for 3 argues that there are $(3\cdot 2)^3$ ways of combining the possible values of parameters for these. So the total number of SETs of this type is $((3\cdot 2)^3 = 216$. So the probability is $216/1080 = .2$.

Alternatively we can count this as follows. These all assume we are given a single card to start with. So the sample space is the collection of all SETs containing this card. There are 40 such SETs.

- One parameter all different: Given one card, there are $\binom{4}{1} = 4$ ways to choose the one parameter that is all different. The parameters that are all the same are fixed by the first card. The one parameter that is all different, there are 2 values for the second card and the third is fixed. However, there are two ways to rearrange the second two cards, leaving $4 * 2/2 = 4$ total SETs of the 40 of this kind. Thus this probability is $4/40 = 1/10 = .1$.

- Two parameters all different: Regardless of the first card there are $\binom{4}{2} = 6$ ways to choose which two parameters are different. For the two parameters that are all different, there are 2 choices of values for the second card for each parameter. The parameters that are all the same are fixed by the first card. The third card is fixed. Thus there are $6 * 4 = 24$ possible SETs, but we over counted the second two cards by a factor of the 2 ways to arrange 2 cards. Thus the number of SETs of this form, within the 40 that use the card given, is $24/2 = 12$. Hence the probability is $12/40 = .3/10 = .3$.

- Three parameters all different: Again there are $\binom{4}{3} = 4$ ways to choose the 3 parameters that are all different. The on that is the same is chosen by the card given. There are 2 choices for the value of each of the other 3 parameters, but of course we can rearrange the second two cards in two ways. So there are $4 * 2^3/2 = 16$ SETs of this type. The probability is $16/40 = 2/5 = .4$.

- Four parameters all different: Finally, following the same argument the number of SETs of this type is $\binom{4}{4} * 2^4/2 = 8$. So the probability is $8/40 = 1/5 = .2$.

As before note that $4 + 12 + 16 + 8 = 40$.

8. We prove this as two lemmas. I present it that way here as that is how we did it in class.

   - **Lemma 1**: We prove that three cards form a SET if and only if the points in $\mathbb{F}_3^n$ corresponding to the cards sum to 0 mod 3.

     Proof. ($\Rightarrow$): Suppose we have 3 cards that form a SET. Then the coordinates of the points corresponding to these 3 cards are either all the same or all different. If they are all the same value, say a, then $a + a + a = 3a \equiv 0 \mod 3$. If the coordinates are all different, then each of 0, 1, and 2, must occur as the values in that coordinate exactly once. Hence the sum is $0 + 1 + 2 = 3 \equiv 0 \mod 3$.

     ($\Leftarrow$): Suppose that we have 3 points in $\mathbb{F}_3^n$ whose coordinates sum to 0. We prove this by considering the cases and looking at the possible coordinate values when the values are not all the same or all different. If the coordinate values are not all the same or all different we must have one of the following scenarios.

     (a) 0,0,1 and $0 + 0 + 1 \equiv 1 \mod 3$.
     (b) 0,0,2 and $0 + 0 + 2 \equiv 2 \mod 3$.
     (c) 0,1,1 and $0 + 1 + 1 \equiv 2 \mod 3$.
     (d) 0,2,2 and $0 + 2 + 2 \equiv 1 \mod 3$.
     (e) 1,1,2 and $1 + 1 + 2 \equiv 1 \mod 3$.
     (f) 1,2,2 and $1 + 2 + 2 \equiv 2 \mod 3$.

     Hence, for the sum to be 0 mod 3, the values in a particular coordinate must be all the same or all different. □

   - **Lemma 2**: We prove that three points in $\mathbb{F}_3^n$ are colinear if and only if the coordinates of the points sum to 0 mod 3.

     Proof. ($\Rightarrow$): Lets assume we have three colinear points and prove that the sum of their coordinates must be 0 mod 3. Let $\mathbf{p}_1$, $\mathbf{p}_2$ and $\mathbf{p}_3$ be the three points (the p's are bold to remind us that these are points with coordinates, not just single numbers). Also, we remark that by expressions like $\mathbf{p}_1 + \mathbf{p}_2$ we mean to add the two points, coordinatewise — or to add just the coordinates of the two points.

     The parametric equation of the line passing through the first two points is $\mathbb{Z} = (\mathbf{p}_1 - \mathbf{p}_2)t + \mathbf{p}_2 \mod 3$. Since the 3 points are colinear, it must be the case that

$$\mathbf{p}_3 = (\mathbf{p}_1 - \mathbf{p}_2)2 + \mathbf{p}_2 \mod 3$$
$$\mathbf{p}_3 = 2\mathbf{p}_1 - \mathbf{p}_2 \mod 3$$
$$\mathbf{p}_3 - 2\mathbf{p}_1 + \mathbf{p}_2 = 0 \mod 3$$

Thus

$$0 = \mathbf{p}_3 - 2\mathbf{p}_1 + \mathbf{p}_2 = \mathbf{p}_3 + \mathbf{p}_1 + \mathbf{p}_2 \mod 3$$

so the sum of the coordinates of the 3 points is 0 mod 3.

($\Leftarrow$): We assume the sum of the three points is 0 and prove and prove that the points must be colinear. This proof is most easily done by doing the last proof backwards. First, we remark again, that the line through two of the points is given by $\mathbb{Z} = (\mathbf{p}_1 - \mathbf{p}_2)t + \mathbf{p}_2 \mod 3$ and that the third point on this line is $(\mathbf{p}_1 - \mathbf{p}_2)2 + \mathbf{p}_2 = 2\mathbf{p}_1 - \mathbf{p}_2 \mod 3$. So we just need to show that $\mathbf{p}_3$ is equal to this point.

Since the sum of the coordinates of the three points are 0 mod 3, we have

$$\begin{aligned} 0 =& \mathbf{p}_3 + \mathbf{p}_1 + \mathbf{p}_2 \mod 3 \\ =& \mathbf{p}_3 - 2\mathbf{p}_1 + \mathbf{p}_2 \mod 3 \\ 2\mathbf{p}_1 - \mathbf{p}_2 = \mathbf{p}_3 \mod 3. \end{aligned}$$

This last equation is the one needed to show that $\mathbf{p}_3$ is the unique third point on the line given by $\mathbf{p}_1$ and $\mathbf{p}_2$ so the three points are colinear. □

9. The previous statements are false with higher numbers of values. We gave examples in $\mathbb{F}_4^2$ of each possible implication failing (e.g. both a SET for which the points are not colinear and 3 colinear points that are not a SET). In $\mathbb{F}_5^2$ things were more subtle, and we noticed that it is true that if three cards form a SET then the corresponding points sum to 0 mod 5, but that all other implications are still false. Finally, in $\mathbb{F}_2^n$ any two cards form a SET and any two points are colinear, but it is possible to have both of these and not have the coordinates sum to 0 mod 2.

10. This appears to be a very difficult question in general. The directions published with the commercial game state that the odds of having a SET in 12 cards is approximately 33 : 1. This means that the probability of having a SET is 33 times that of not having a set. For this to be true, the probability of having a SET must be 33/34 and of not having a SET 1/34. However, they say "approximately." Given the difficult nature of this question and the lack of a published solution, it appears that the only well-known approach is computer simulation. An analytic solution would be nice to have.

11. An upper bound is easy to get for any number of cards. The best we could ever possibly do is to have the situation where for any two cards in the collection, the third card that makes them a SET is also in the collection. So for k cards, the absolute largest possible number of SETs is $k(k-1)/6$. For 9 cards this number is 12, and since 9 cards can correspond to $\mathbb{F}_3^2$, this number of SETs is obtainable — simply restrict to 2 parameters and you'll have 9 cards that contain 12 SETs. A similar argument can be given for the maximum number of SETs obtainable in 27 cards is $(27)(26)/6 = 117$. The case of the maximum number of SETs obtainable in 12 cards is much more subtle. The upper bound is $22 = (12)(11)/6$. However, the maximum that anyone has ever assembled is 14. It is also pretty easy to see

that we aren't likely to have all 22 SETs as this would require $\binom{12}{9}$ versions of the 2-dimension game to be present in the 12 cards as this seems difficult. It is an open question if it is possible to build more than 14 SETs into 12 cards, or to prove that 14 is actually maximal. (Note that this makes any number of cards that is a multiple of 3, but not a power of 3, an interesting number of cards for this question).

12. Given three cards without a SET, there are 3/78 cards remaining that make these three cards a superset. One way to count this is to take and of the 3 possible pairs we can form with the 3 cards we already have. Then there is a unique third card that makes a SET with this pair. Consider that card with the remaining card we already have. There is a unique third card that makes them a SET, this is the card we need to make a superset. Since we have 3 pairs we can consider, there are 3 cards in the remaining deck that make this a superset.

    This argument assumes we already have 3 cards without at SET. Thus the probability that any 4 cards make a super set is $P($ superset $\mid$ 3 cards without a SET$)\cdot$ $P($getting 3 cards without a SET$) = (3/78)(78/79) = 3/79$.

    Our argument for the probability shows that 4 cards can be a superset in at most one way, since once we have 3 cards with no set each of the 3 possible cards we can add gives the set of 4 only one superset – the cards used are unique at that point.

13. It is possible to have 0, 6, 9, 12, 15, or 18 cards left at the end of the game, although, through experience and computer simulation we are mostly left with 6 and 9. It would be nice to have probabilities for not having a SET in 6, 9, 12, 15, and 18 cards giving some indication of the probabilities of these ending numbers (this is very open). What is interesting is that we cannot possibly have 3 cards left at the end of the game.

    We argue that we cannot have 3 cards left at the end of the game. Our argument uses Lemma 1 above. We argue first that if there are 3 cards left at the end of the game they either have all the same number of objects on the cards or all different. Then the arguments for the other 3 parameters are very similar. Note that there are 27 cards in the deck with 1 figure on the card, 27 with 2 and 27 with 3, so there are a total of $27\cdot1+27\cdot2+27\cdot3 = 162$ figures on the cards. By Lemma 1 any valid SET taken from the deck has the property that the sum of the number of figures on those cards is a multiple of 3. So, if we have 3 cards left at the end of the game, we've removed 26 valid SETs from the deck and $3k$ figures for some positive integer k. Thus the remaining 3 cards have $162 - 3k = 3(54 - k)$ figures. Since the sum of the number of figures on the remaining 3 cards is a multiple of 3, they must be all the same or all different, by Lemma 1.

    This is, in some sense the easiest one since we can count the total number of figures. To argue the other three we have two options. We can use the corresponding points and then we can give the same count and argument above for all the parameters. Or, we can argue that the number of cards with each value for a particular parameter must be equivalent mod 3 after each valid SET is removed from the deck. Thus, when there are 3 cards left, the possible values of each parameter are, all values are

1 or two are 0 and one is 3, meaning that within that parameter the 3 remaining cards are either all the same or all different and hence form a SET.

14. Of course we leave this one for homework.

In the references we include the three key mathematics papers as well as well as a fourth paper that has some interesting, largely open questions [Z] and a web page that seems to have lots of good information [M]. There are a number of other web pages out there, but these really seem to be the best for mathematics, not repeating too much and that address more than just the most basic questions.

References

[DM] B. L. Davis and D. Maclagan. The card game SET. *Intell.* 25(3):33–40, 2003.

[EFS] Y. Edel, S. Ferret, and L. Storme. The classificiation of the largest caps in $AG(5, 3)$. *Journal of Combinatorial Theory, Series A*, 99:95–110, 2002.

[M] T. Magliery. The "Unofficial" SET Home Page. Available at `http://magliery.com/Set/`.

[P] G. Pellegrino. Sul massimo ordine delle calotte in $S_{4,3}$. *Matematiche (Catania)*, 25:149--157 (1971), 1970.

[Z] M. Zabrocki. The Joy of SET. Available at www.math.yorku.ca/ zabrocki/set/joyofset.p