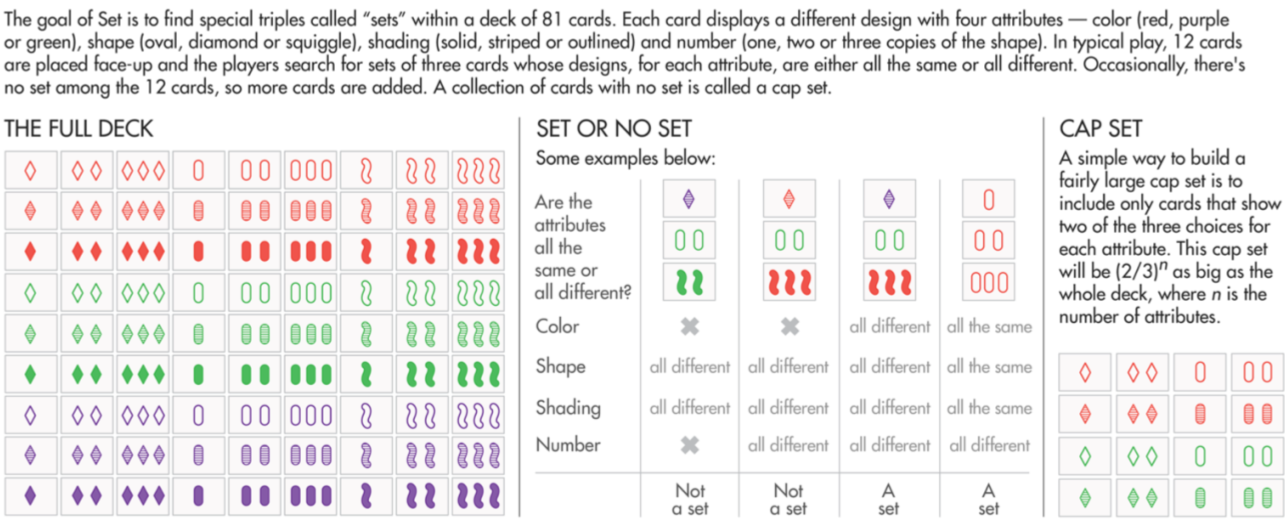
Exploring Attribute Distributions for Insights into the Capset Problem

Abstract

The Capset Problem, a long-standing mathematical puzzle, remains unsolved. This manuscript takes an unconventional approach by examining the attribute distributions of capsets, aiming to uncover patterns. Our goal is to determine the size of capsets with 'n' attributes. We propose that capsets follow specific distribution patterns, which, if identified, could lead to a breakthrough in determining their size.

The Game 'Set'



The Capset Problem

The Capset Problem is a fundamental and unsolved mathematical challenge situated at the intersection of combinatorics and additive number theory. This problem centers on the concept of sets of numbers within a finite mathematical space, commonly considered within the framework of finite fields.

In affine geometry, a cap set is a subset of Z_3^n with no three elements in a line. The cap set problem is the problem of finding the size of the largest possible cap set, as a function of n.

Consider a finite field, frequently chosen to be the field of integers modulo a prime number F_p, where p is a prime number. Within this finite field, the objective is to ascertain the largest possible set such that no triplet of numbers within the set forms an arithmetic progression.

The first few capset sizes are 1, 2, 4, 9, 20, 45, 112

The Capset Problem is an important mathematical puzzle with implications in fields like combinatorics, number theory, algebra, coding, and error correction. Solving it would deepen our understanding of fundamental math concepts and their practical uses. However, even though many mathematicians have tried, the Capset Problem remains unsolved. It's a significant and long-standing challenge. Researchers are still looking for new ways to crack this mathematical mystery.

My Work

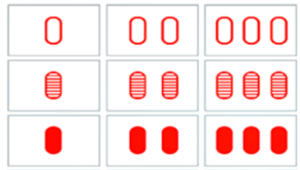
Terminology

Attributes : In the context of this problem, the term "attribute" signifies a dimension. For example, in a 2-dimensional space, two attributes are required to denote points. This nomenclature is adopted from the "Set" game, where cards are represented in a 4-dimensional space using four attributes: color, shape, number, and fill.

Design or Category : The design or category refers to the various values an attribute can assume. In this particular problem, each attribute can adopt one of three distinct values. Consequently, for every attribute, there are precisely three categories present.

Two attributes

In the scope of our analysis, we consider two attributes, namely "number" and "fill." Given that each attribute has three categories, there are a total of $3^2 = 9$ possible cards within this attribute space.



It is worth noting that exhaustive mathematical examination has determined that the largest possible cardinality of a capset within this 2-dimensional attribute space is 4.

In the context of a capset comprising 4 cards with only 2 attributes, focusing on the "number" attribute, each card can assume a value of 1, 2, or 3. It's important to note that there can be a maximum of 3 cards sharing the same number, and there must be at least 2 distinct categories or numbers present within this capset.

When discussing "distribution" in this context, we refer to how a specific attribute is represented or broken down among the available categories. It describes both the quantity of cards belonging to a particular category and the number of categories present within the collection of cards.

Indeed, to clarify, we are not concerned with the permutations within the values of the categories when considering these distributions. For example, [1,1,2] can represent 1 card with the number 1, 1 card with the number 2, and 2 cards with the number 3. It can also represent 2 cards with the number 1, 1 card with the number 2, and 1 card with the number 3, or 2 cards with the number 2, 1 card with the number 1, and 1 card with the number 3. In essence, the specific numerical values do not matter; what's important is that the distribution or breakdown of categories remains the same.

Consequently, the values of the 4 cards can be arranged in the following distributions, including [1,1,2], [0,1,3], or [0,2,2], while considering these distributions as distinct from one another.

To elaborate further, since there are 2 attributes in consideration, each offering 3 possible options, we arrive at 6 distinct combinations of attribute distributions.

$\binom{3+2-1}{2} = 6$ (As choose 2 from 3 with repetition allowed but order is not important)
Example : [1,1,2], [0,2,2] – The numbers are distributed in the ratio [1,1,2] while the fills are distributed in the ratio [0,2,2]

The table below provides a breakdown of each combination, illustrated with one example for each its viability in creating a set of 4 cards, its potential as a capset, and associated rationales.

Distribution	Possible	Set	Capset	Reason
[1, 1, 2], [1, 1, 2]	Yes	 	Yes No	These configurations constitute capsets when both categories containing two cards share a common card. If this overlap is absent, the distribution does not form a capset.
[1, 1, 2], [0, 2, 2]	Yes		Yes	In situations where one of the two attributes exhibits a [2,2] distribution, it invariably forms a capset. This outcome arises from the presence of two cards belonging to different categories within that attribute. Consequently, it becomes impossible to construct sets comprising cards exclusively from the same category since none of the categories have all three cards present. Simultaneously, forming sets containing entirely distinct cards is unattainable due to the absence of the third category card.
[1, 1, 2], [0, 3, 1]	Yes		No	These configurations do not qualify as capsets because they inherently include a specific type of set. In such cases, one of the attributes exhibits a distribution of [3,1], signifying the presence of three cards of a single type within that attribute. Given that there are only two attributes in this context, it implies the existence of a set in which three cards of the same type are present.
[0, 2, 2], [0, 2, 2]	Yes		Yes	These configurations consistently represent capsets. This is because, as mentioned earlier, they involve two cards from two different categories, which prevents the formation of sets consisting entirely of unique or distinct elements.
[0, 2, 2], [0, 3, 1]	No			This configuration is not feasible. It's impossible to have only two categories in one attribute and have all three cards of one category in another attribute. If all three cards are present in the second attribute, then by necessity, all categories must be represented in the first.
[0, 3, 1], [0, 3, 1]	No			Same as above, this configuration is not possible.

The table presented above provides a concise summary, indicating that among the 6 combinations of attribute distributions considered, only 3 have the capability to form capsets.

To substantiate this assertion, we have meticulously generated all possible combinations of 4 cards from the 9 available cards, specifying whether they qualify as capsets or not, along with their associated distribution patterns.

Attached herewith is a file containing a comprehensive list of all $\binom{9}{4} = 126$ possible combinations of 4 cards from the initial 9. It is noteworthy that among these, 43 meet the criteria for being capsets. Intriguingly, all these capsets can be categorized into one of three distinct distribution patterns:

1. [1, 1, 2], [1, 1, 2]
2. [1, 1, 2], [0, 2, 2]
3. [0, 2, 2], [0, 2, 2]

This empirical evidence underscores the robustness of the claim that only specific distribution patterns can give rise to capsets within this attribute space.

Three attributes

Similarly, when we expand our analysis to 3 attributes, we encounter a total of $3^3 = 27$ possible cards within this attribute space.

In this context, the capset size is established as 9. There are $\binom{27}{9} = 4{,}686{,}825$ unique combinations of 9 cards to consider.

Breaking down the number 9 into 12 different distributions, as follows:
[0,0,9], [0,1,8], [0,2,7], [0,3,6], [0,4,5], [1,1,7], [1,2,6], [1,3,5], [1,4,4], [2,2,5], [2,3,4], [3,3,3]

$\binom{12+3-1}{3} = 364$ (As choose 3 from 12 with repetition allowed but order is not important)

Out of these $4{,}686{,}825$ combinations, there exist 2106 capsets.

Remarkably, among these 2106 capsets. It is worth noting that all of these capsets strictly conform to one of the 3 specific distribution patterns as outlined below:

1. [1,4,4], [1,4,4], [1,4,4]
2. [1,4,4], [1,1,4], [3,3,3]
3. [1,4,4], [3,3,3], [3,3,3]

It's important to note that while not all configurations of cards following these distributions constitute capsets, all capsets indeed adhere to one of these specific distribution patterns. This observation underscores the significance of these distributions in understanding the nature of capsets in this particular mathematical context.

Attached is a file containing a comprehensive record of all possible capsets involving 3 attributes, along with their associated attribute distributions.

Four attributes

Although we haven't been able to generate all capsets for 4 attributes due to the presence of 81 cards and a capset size of 20 (As there exist $\binom{81}{20}$ combinations of 20 cards), we have examined a subset.

It's worth noting that among the combinations we have generated, they all adhere to one of the following three distribution patterns.

1. [6,6,8], [6,6,8], [6,6,8], [6,6,8]
2. [6,6,8], [6,6,8], [6,6,8], [2,9,9]
3. [6,6,8], [6,6,8], [2,9,9], [2,9,9]

Future work

In the subsequent phases of our research, we intend to undertake the following key investigations:

1. Pattern Identification: Our primary focus will be on identifying and analyzing patterns within attribute distributions.

2. Secondary Conditions: We will systematically explore the secondary conditions under which combinations, despite exhibiting these specific distribution patterns, do not culminate in capsets.

3. In-Depth Understanding: Our research will delve into a deeper understanding of the fundamental principles that render these distribution patterns conducive to capset formation.

4. Methodology Development: Our ultimate goal is to contribute to the formulation of a systematic methodology for generating capsets in 'n' dimensions.

These pursuits collectively form the core of our forthcoming research efforts, aimed at advancing our comprehension of this captivating mathematical phenomenon.