

# Probability Theory on Finite Sets

Michael Betancourt

January 2023

## Table of contents

<b>1</b>	<b>Finite Sets</b>	<b>1</b>
1.1	Subsets . . . . .	2
1.2	Subset Operations . . . . .	4
<b>2</b>	<b>Measure and Probability Over Elements</b>	<b>7</b>
<b>3</b>	<b>Measure and Probability Over Subsets</b>	<b>13</b>
<b>4</b>	<b>Acknowledgements</b>	<b>18</b>
	<b>License</b>	<b>19</b>

Despite its infamous reputation the foundations of probability theory are quite straightforward. Much of the mathematical difficulty arises only when we implement probability theory on elaborate sets like the real numbers, and most of the philosophical difficulty arises only when trying to assign an interpretation to the mathematics. In this chapter I try to avoid this baggage by introducing the basics of abstract probability theory on the simplest, nontrivial set: a collection of a finite number of elements.

## 1 Finite Sets

A **finite set** contains a finite number of distinct elements,

$$X = \{x_1, \dots, x_N\}.$$

The numerical index here allows us to differentiate the between the  $N$  individual elements but it does not necessarily imply that the elements have any particular ordering to them. To avoid any assumption of ordering I'll use the five element set (Figure 1)

$$X = \{\square, \clubsuit, \diamond, \heartsuit, \spadesuit\}$$

for demonstration.

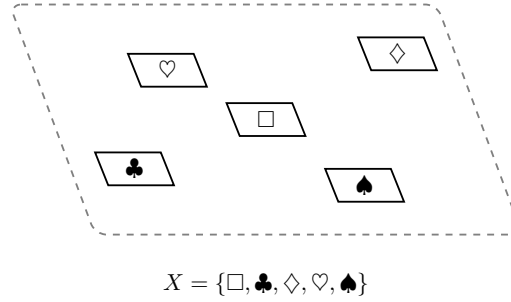


Figure 1: A finite set contains a finite number of elements. This particular set contains five elements.

In practical applications of probability theory the abstract elements  $x_n$  will model meaningful objects, but in this chapter I will avoid any particular interpretation and instead focus on the mathematical concepts. That said when  $X$  is intended to capture *all* of the objects of interest in a given application I will refer to it as the **ambient set**.

Once we've defined an ambient set we have various ways of organizing its individual elements and manipulating those organizations.

## 1.1 Subsets

A **subset** of  $X$  is any collection of elements in  $X$ . To avoid any ambiguity I will exclusively use lowercase roman letters  $x$  to denote a variable element in the ambient set  $X$  and lowercase sans serif letters  $x$  to denote a variable subset.

For example  $x = \{\square, \diamond, \heartsuit\}$  is a subset of  $X = \{\square, \clubsuit, \diamond, \heartsuit, \spadesuit\}$  (Figure 2). Importantly there is no notion of multiplicity in the concept of a subset, just membership: a subset can include an element  $x_n$  but it cannot include it multiple times.

If  $x$  is a subset of the ambient set  $X$  then we write  $x \subset X$ . When  $x$  might contain all of the elements of  $X$ , in which case  $x = X$ , then we write  $x \subseteq X$ .

Subsets are *recursive* objects. Selecting any elements from a subset yields a new subset. For example the collection  $\{\diamond, \heartsuit\}$  is a subset of both the previously introduced subset,

$$\{\diamond, \heartsuit\} \subset \{\square, \diamond, \heartsuit\},$$

and the full ambient set,

$$\{\diamond, \heartsuit\} \subset X.$$

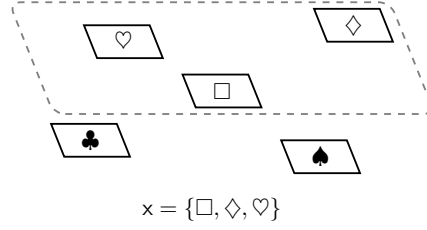


Figure 2: A subset  $x \subset X$  is any collection of elements from the ambient set  $X$ . Here  $x = \{\square, \diamond, \heartsuit\}$  contains just three of the five elements in  $X = \{\square, \clubsuit, \diamond, \heartsuit, \spadesuit\}$ .

More formally a subset  $x'$  is a subset of another subset  $x$  if and only if all of the elements in  $x'$  are also in  $x$ .

Regardless of how many elements a finite set  $X$  contains three special subsets are always well-defined. The **empty set**  $\emptyset = \{\}$  contains no elements at all. On the other hand the entire set itself can be considered a subset containing all of the elements. A subset containing a single element is denoted  $\{x_n\}$  and referred to as an **atomic set**.

There are

$$\binom{N}{n} = \frac{N!}{n!(N-n)!}$$

ways to select  $n$  elements from a finite set of  $N$  total elements, and hence  $\binom{N}{n}$  total subsets of size  $n$ . For example there is only one subset that contains no elements,

$$\binom{N}{0} = \frac{N!}{0!(N-0)!} = \frac{N!}{N!} = 1,$$

which is just the empty set. Similarly there is only one subset that contains all of the elements,

$$\binom{N}{N} = \frac{N!}{N!(N-N)!} = \frac{N!}{N!} = 1,$$

which is just the full set itself. On the other hand there are

$$\binom{N}{1} = \frac{N!}{1!(N-1)!} = N$$

distinct atomic sets that contain a single element, one for each element in  $X$ .

Counting all of the subsets of all possible sizes gives

$$\sum_{n=0}^N \binom{N}{n} = 2^N$$

possible subsets that we can construct from a finite set with  $N$  elements. In other words the collection of all subsets is itself a finite set with  $2^N$  elements. We refer to this set as the **power set** of  $X$  and denote it  $2^X$ .

When the ambient set, and all hence all of the possible elements that any subset might contain, is fixed the prefix “sub” is sometimes dropped, with all subsets simply referred to as sets. For example collections of elements might be denoted “sets” when one takes the ambient set for granted and “subsets” when one wants to explicitly acknowledge the context of the ambient set. Precise terminology, however, can vary strongly between disciplines and even individuals.

## 1.2 Subset Operations

We can always construct subsets element by element, but we can also construct them by manipulating existing subsets.

For example given a subset  $x \subset X$  we can construct its **complement** by collecting all of the elements in  $X$  that are not already in  $x$ . The atomic set  $x = \{\diamond\}$  contains the lone element  $\diamond$  and its complement contains the remaining elements (Figure 3)

$$x^c = \{\square, \clubsuit, \heartsuit, \spadesuit\}.$$

By construction the complement of the empty set is the entire set,  $\emptyset^c = X$ , and the complement of the full set is the empty set,  $X^c = \emptyset$ .

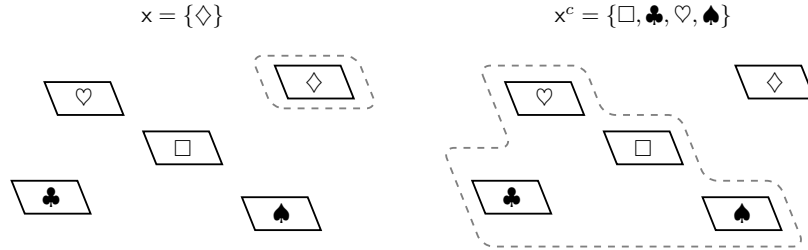


Figure 3: The complement of a subset  $x$  is the subset  $x^c$  consisting of all elements in the ambient set that are not in  $x$ .

More formally the construction of complementary subsets defines a *unary* function that takes in a subset as input and outputs the complementary subset,

$$\begin{array}{l} .^c : 2^X \rightarrow 2^X \\ x \mapsto x^c . \end{array}$$

The top line of this notation denotes the collection of all possible inputs and outputs, here both the power set, while the bottom line denotes the action on a particular input, here a subset mapped into its complement. For example applying the complement operator to the subset  $x = \{\clubsuit, \spadesuit\}$  gives

$$x^c = \{\clubsuit, \spadesuit\}^c = \{\square, \diamond, \heartsuit\}.$$

We can also construct subsets from more than one subsets. Consider, for example, two subsets  $x_1 = \{\square, \heartsuit\}$  and  $x_2 = \{\square, \spadesuit\}$  (Figure 4). The collection of all elements that are contained in *either* subset is itself a subset,

$$\{\square, \heartsuit, \spadesuit\} \subset X,$$

as is the collection of all elements that are contained only in *both* subsets,

$$\{\square\} \subset X.$$

These derived subsets are referred to as the **union**,

$$x_1 \cup x_2 = \{\square, \heartsuit\} \cup \{\square, \spadesuit\} = \{\square, \heartsuit, \spadesuit\},$$

and **intersection**,

$$x_1 \cap x_2 = \{\square, \heartsuit\} \cap \{\square, \spadesuit\} = \{\square\},$$

respectively (Figure 5). Note that the union and intersection are both symmetric in the sense that either order of the input subsets results in the same output subset,

$$x_1 \cup x_2 = x_2 \cup x_1$$

and

$$x_1 \cap x_2 = x_2 \cap x_1.$$

Two subsets are **disjoint** if they don't share any elements; in this case their intersection is the empty set,

$$x_1 \cap x_2 = \emptyset.$$

The union and intersection of a subset with itself returns that subset,

$$x \cup x = x \cap x = x.$$

Because the empty set does not contain any elements its union with any subset returns back that subset,

$$x \cup \emptyset = \emptyset \cup x = x,$$

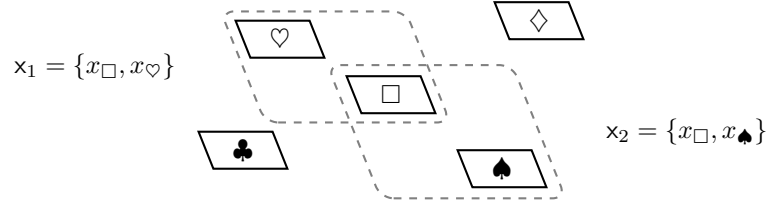


Figure 4: We can manipulate two subsets into a new subset in various ways.

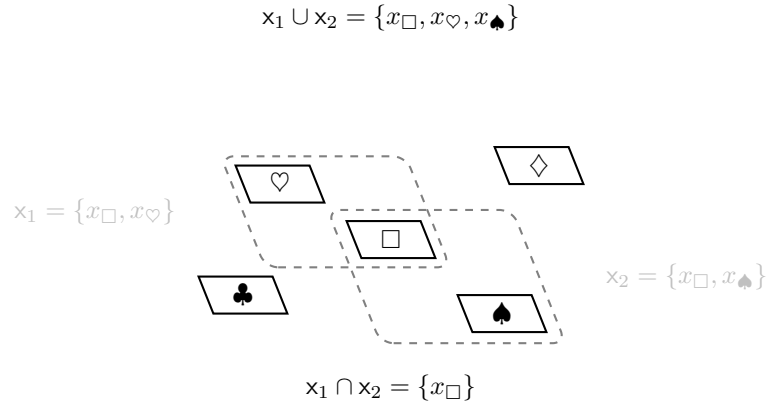


Figure 5: The union of two subsets,  $x_1 \cup x_2$ , is a subset containing all of the elements in either input subset. On the other hand the intersection of two subsets,  $x_1 \cap x_2$ , is a subset containing just the elements that occur in both input subsets.

and its intersection with any subset returns the empty set,

$$x \cap \emptyset = \emptyset \cap x = \emptyset.$$

Similarly the union of any subset with the full set returns the full set,

$$x \cup X = X \cup x = X,$$

and the intersection of any subset with the full set returns back the subset,

$$x \cap X = X \cap x = x.$$

Because they require two inputs the union and intersection define *binary* functions that consume *two* input subsets and return a single output subset,

$$\begin{aligned} \cdot \cup \cdot : 2^X \times 2^X &\rightarrow 2^X \\ x_1, x_2 &\mapsto x_1 \cup x_2 \end{aligned}$$

and

$$\begin{aligned} \cdot \cap \cdot : 2^X \times 2^X &\rightarrow 2^X \\ x_1, x_2 &\mapsto x_1 \cap x_2. \end{aligned}$$

Here  $2^X \times 2^X$  denotes the collection of all *pairs* of subsets.

## 2 Measure and Probability Over Elements

Measure theory, and its special case of probability theory, is often burdened with intricate if not mysterious interpretations. From a mathematical perspective, however, measure theory simply concerns the consistent **allocation** of some abstract quantity across the ambient set.

Consider a *reservoir* of some positive, continuous, and *conserved* quantity,  $M \in [0, \infty]$  (Figure 6). Because  $M$  is conserved any amount  $m_n$  that is allocated to the element  $x_n \in X$  has to be *depleted* from the reservoir, leaving less to be allocated to the remaining elements.

We have to be careful if the total content of the reservoir  $M$  is infinite. In this case we can allocate an infinite amount from the reservoir while still having an infinite quantity left. At the same time allocating an infinite amount can depleting the reservoir completely or even leave any finite quantity. Infinity is, at the very least, mathematically awkward.

To make the mathematics as useful as possible we will avoid endowing  $M$  with any particular meaning for the time being. Instead interpretations will arise only when we use these allocations to *model* meaningful systems.

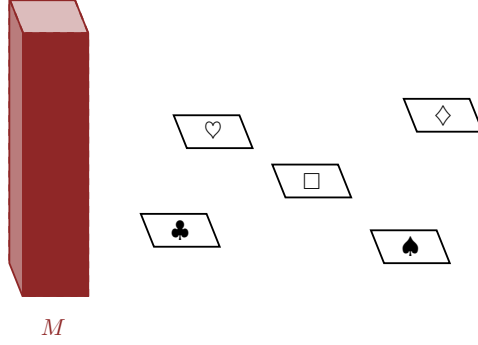


Figure 6: Measure theory concerns the allocation of some continuous and positive quantity  $M$  over the individual elements of the ambient set.

An *exhaustive* allocation of  $M$  across the ambient set ensures that the reservoir is completely emptied. In another words all of  $M$  has to be allocated to the elements  $x_n \in X$ .

For example consider our demonstrative ambient set  $X = \{\square, \clubsuit, \diamond, \heartsuit, \spadesuit\}$ . If we allocate  $m_{\square}$  to  $\square$  then that leaves  $M - m_{\square}$  remaining to allocate to the other four elements (Figure 7a). Allocating  $m_{\clubsuit}$  to  $\clubsuit$  depletes the reservoir a bit more (Figure 7b). At the end we have to allocate all

$$M - m_{\square} - m_{\clubsuit} - m_{\diamond} - m_{\heartsuit}$$

that remains to  $\spadesuit$  to completely empty the reservoir (Figure 7e).

A **measure** is any consistent allocation of the quantity  $M$  to the elements of an ambient set. Mathematically any measure over a finite set can be characterized by  $N$  numbers (Figure 8)

$$\mu = \{m_1, \dots, m_N\}$$

that satisfy

$$0 \leq m_n$$

and

$$\sum_{n=1}^N m_n = M.$$

For example any measure over the five-element set  $X = \{\square, \clubsuit, \diamond, \heartsuit, \spadesuit\}$  is specified by any five positive numbers  $\{m_{\square}, m_{\clubsuit}, m_{\diamond}, m_{\heartsuit}, m_{\spadesuit}\}$  satisfying

$$m_{\square} + m_{\clubsuit} + m_{\diamond} + m_{\heartsuit} + m_{\spadesuit} = M.$$

The larger  $m_n$  the more of  $M$  is allocated to the element  $x_n$ . Following this terminology we will also refer to  $M$  as the **total measure** and  $m_n$  as the **measure allocated to  $x_n$** .



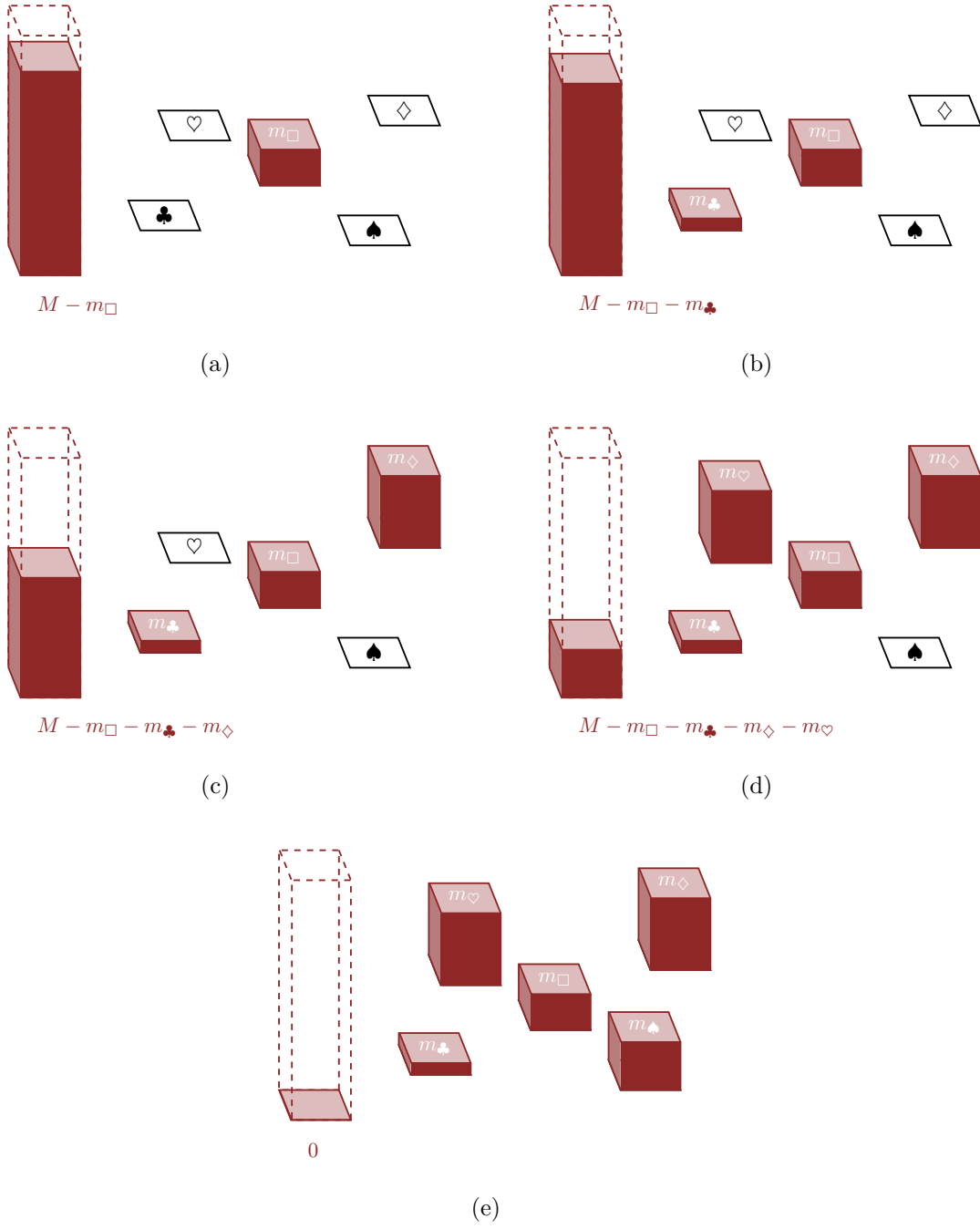
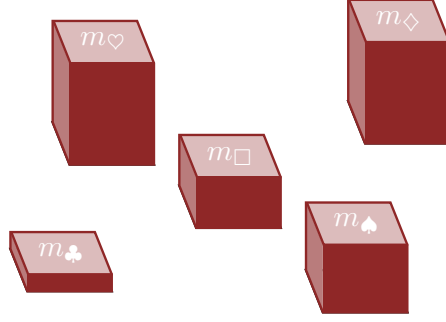


Figure 7: Because the total quantity  $M$  is conserved every allocation  $m_n$  to an element  $x_n \in X$  depletes the amount available for the allocation to the remaining elements. An exhaustive allocation leaves nothing left in the initial reservoir after each element has received its allocation.



$$m_{\square} + m_{\clubsuit} + m_{\diamond} + m_{\heartsuit} + m_{\spadesuit} = M$$

Figure 8: A measure  $\mu$  over the finite set  $X$  is any consistent allocation of  $M$  to the elements  $x_n \in X$ . Every measure can be characterized by  $N$  numbers  $m_n$  that sum to  $M$ , or equivalently a function that maps each element  $x_n$  to its allocated measure  $m_n$ .

If the total measure is infinite,  $M = \infty$ , then at least one of the elements in  $X$  *has* to receive an infinite allocation in order to empty the reservoir. We can also consistently allocate infinite measure to multiple elements, or even all of the elements, at the same time. Infinite measures can be very generous in their allocations.

The allocation  $\{x_n, m_n\}$  defined by a measure  $\mu$  can also be organized into a function that maps each element to its associated allocation,

$$\begin{aligned} \mu : X &\rightarrow [0, \infty] \\ x_n &\mapsto m_n = \mu(x_n). \end{aligned}$$

A nice conceptual benefit of this functional perspective is that instead of thinking about the global allocation  $m_1, \dots, m_N$  all at once we can isolate local allocations by evaluating  $\mu$  at only a single input  $m_n = \mu(x_n)$ .

In general there are an infinite number of ways to allocate a total measure to the elements of a finite set, and hence an infinite number of measures that we can define over that set. I will denote the set of all possible measures over  $X$  as  $\mathcal{M}(X)$ .

Within this set there are a few notable examples. For example a **singular measure** allocates the total measure  $M$  to a single element, leaving the rest with nothing (Figure 9a). On the other hand a **uniform measure** allocates the same measure  $M/N$  to each element (Figure 9b). On finite sets there are  $N$  distinct singular measures, one for each distinct element, and a single unique uniform measure.

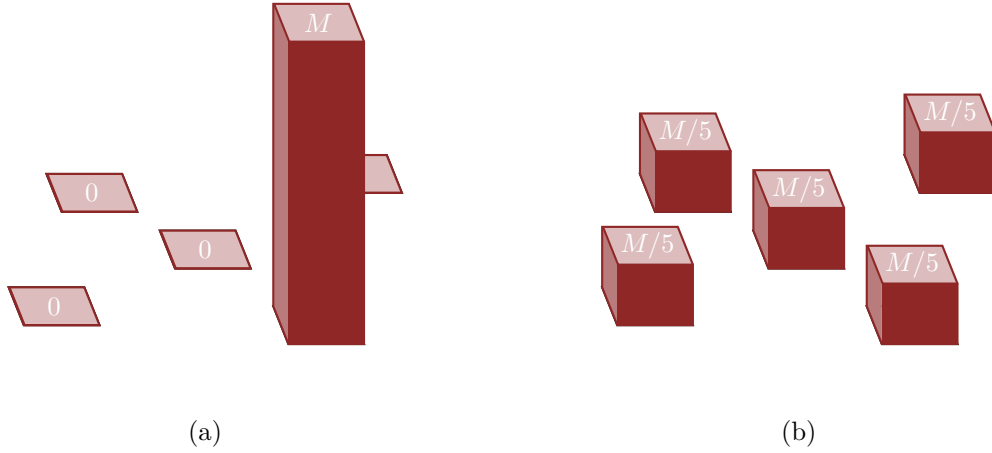


Figure 9: A singular measure (a) allocates the total measure to a single element while the uniform measure (b) spreads the total measure to each element evenly.

Perhaps the most important class of measures, however, are measures where the total measure is finite,  $M < \infty$ . Appropriately enough we refer to these as **finite measures**.

What makes finite measures so special is that we can always reframe the allocation they define into a *relative* one. Instead of considering the absolute measure allocated to each element  $m_n$  we can consider the *proportion* of the total measure allocated to each element, (Figure 10)

$$p_n = m_n/M.$$

By construction proportions are confined to the unit interval  $[0, 1]$ . As with any quantity taking values in  $[0, 1]$  we can represent proportions equally well with decimals, for example  $p_n = 0.2$ , and percentages,  $p_n = 20\%$ .

In other words a proportional measure defines the function (Figure 11)

$$\begin{aligned} \pi : X &\rightarrow [0, 1] \\ x_n &\mapsto p_n = \pi(x_n) \end{aligned}$$

with

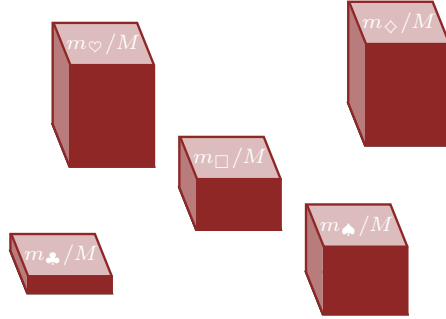
$$0 \leq p_n \leq 1$$

and

$$\sum_{n=1}^N p_n = 1.$$

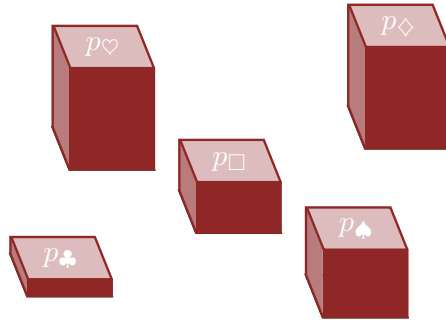
A collection of variables  $\{p_1, \dots, p_N\}$  satisfying these properties is referred to as a **simplex**.

More importantly a proportional measure  $\pi$  is also known as a **probability distribution** with the proportional allocations  $p_n$  denoted **probabilities**. While the term “probability” is



$$\frac{m_{\square}}{M} + \frac{m_{\clubsuit}}{M} + \frac{m_{\diamond}}{M} + \frac{m_{\heartsuit}}{M} + \frac{m_{\spadesuit}}{M} = \frac{M}{M} = 1$$

Figure 10: Every finite measure can be characterized by a proportional allocation.



$$p_{\square} + p_{\clubsuit} + p_{\diamond} + p_{\heartsuit} + p_{\spadesuit} = 1$$

Figure 11: A proportional allocation is also known as a probability distribution.

often encumbered with all kinds of interpretational and philosophical baggage its mathematical structure is really quite straightforward: on a finite set a probability is just the proportion of some finite quantity that is allocated to an individual element!

Philosophical tension arises only when we try to assign some meaning to these proportional allocations. This isn't a question of mathematics but rather *applying* mathematics to model a system of interest. We'll come back to the many different systems that can be consistently modeled with probability theory, and hence the many different interpretations of probability itself, in Chapter 5.

While we're on the topic of terminological issues the word "distribution" is not without its own problems. In mathematics the term "distribution" is heavily overloaded and can be used to refer to a variety of related but formally distinct concepts. One way to avoid confusion is to always refer to proportional measures as "probability distributions" and never just "distributions" alone.

### 3 Measure and Probability Over Subsets

On finite sets any allocation, absolute or proportional, over the individual elements  $x \in X$  also defines an allocation over entire subsets  $x \in 2^X$ . The measure allocated to a subset is just the sum of the measures allocated to the elements in that subset. For example the measure allocated to  $x = \{\square, \clubsuit, \heartsuit\}$  is  $m_{\square} + m_{\clubsuit} + m_{\heartsuit}$  (Figure 12).

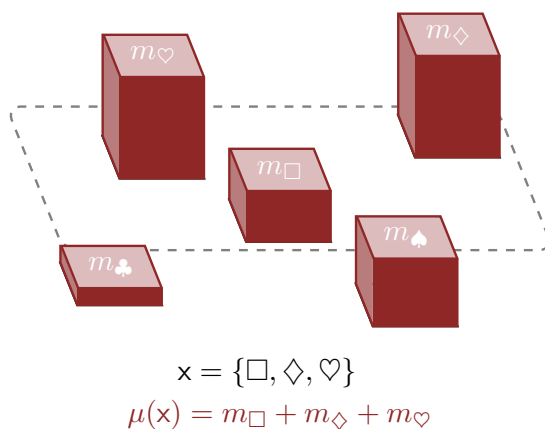


Figure 12: On a finite set an allocation over individual elements also defines an allocation over any subset.

In other words a measure over individual elements  $\mu : X \rightarrow [0, \infty]$  implicitly defines a measure over subsets  $\mu : 2^X \rightarrow [0, \infty]$ . Similarly a probability distribution over individual elements  $\pi : X \rightarrow [0, 1]$  implicitly defines a probability distribution over subsets  $\pi : 2^X \rightarrow [0, 1]$ .

Note that we have made the cardinal sin of *overloading* our notation so that  $\mu$  refers to both types of measures, and the only way to differentiate them is through their different inputs;  $\mu(x)$  denotes the element-wise measure allocated to  $x \in X$  and  $\mu(\mathbf{x})$  denotes the subset-wise measure allocated to  $\mathbf{x} \in 2^X$ . Maintaining different typographical conventions for elements and subsets is critical to avoid any confusion when overloading notation like this.

By construction any subset measure and probability distribution satisfy a wealth of useful properties. For example for any measure

$$\mu(\emptyset) = 0$$

and

$$\mu(X) = \sum_{n=1}^N \mu(x_n) = M,$$

while for any probability distribution we have  $\pi(\emptyset) = 0$  and  $\pi(X) = 1$ .

Even better the subset allocations play well with the subset operations. Consider for example the two disjoint subsets  $\mathbf{x}_1 = \{\square, \diamond\}$  and  $\mathbf{x}_2 = \{\clubsuit, \spadesuit\}$ . Because the two subsets are disjoint their union simply combines all of their elements,

$$\mathbf{x}_1 \cup \mathbf{x}_2 = \{\square, \diamond\} \cup \{\clubsuit, \spadesuit\} = \{\square, \clubsuit, \diamond, \spadesuit\},$$

and the measure of that union is just the sum of the measures of the input subsets,

$$\begin{aligned} \mu(\mathbf{x}_1 \cup \mathbf{x}_2) &= \mu(\{\square, \clubsuit, \diamond, \spadesuit\}) \\ &= m_{\square} + m_{\clubsuit} + m_{\diamond} + m_{\spadesuit} \\ &= (m_{\square} + m_{\diamond}) + (m_{\clubsuit} + m_{\spadesuit}) \\ &= \mu(\mathbf{x}_1) + \mu(\mathbf{x}_2). \end{aligned}$$

More generally for any collection of subsets

$$\mathbf{x}_1, \dots, \mathbf{x}_K$$

that are mutually disjoint,

$$\mathbf{x}_k \cap \mathbf{x}_{k'} = \emptyset$$

for  $k \neq k'$ , we have

$$\mu(\cup_{k=1}^K \mathbf{x}_k) = \sum_{k=1}^K \mu(\mathbf{x}_k).$$

In words if we can decompose a subset into a disjoint collection of smaller subsets then we can also decompose the measure allocated to that initial subset into measures allocated to the component subsets. This consistency property is known as **additivity**.

A subset  $x$  and its complement  $x^c$  always disjoint,  $x \cap x^c = \emptyset$ . At the same time their union covers the entire set,  $x \cup x^c = X$ . Consequently additivity implies that

$$\begin{aligned} M &= \mu(X) \\ &= \mu(x \cup x^c) \\ &= \mu(x) + \mu(x^c), \end{aligned}$$

or

$$\mu(x^c) = M - \mu(x).$$

In words the measure allocated to the complement of a subset is the total measure less the measure allocated to that subset. For probability distributions this becomes even cleaner,

$$\pi(x^c) = 1 - \pi(x).$$

When two subsets overlap we have to take into consideration that the sum of their measures  $\mu(x_1) + \mu(x_2)$  double counts the allocations to any elements shared between them. For example if  $x_1 = \{\square, \heartsuit\}$  and  $x_2 = \{\square, \spadesuit\}$  then the union includes the overlapping element  $\square$  only once,

$$x_1 \cup x_2 = \{\square, \heartsuit\} \cup \{\square, \spadesuit\} = \{\square, \heartsuit, \spadesuit\}.$$

Consequently

$$\begin{aligned} \mu(x_1 \cup x_2) &= \mu(\{\square, \heartsuit, \spadesuit\}) \\ &= m_{\square} + m_{\heartsuit} + m_{\spadesuit}. \end{aligned}$$

Adding the measures allocated to the two subsets individually, however, gives

$$\begin{aligned} \mu(x_1) + \mu(x_2) &= (m_{\square} + m_{\heartsuit}) + (m_{\square} + m_{\spadesuit}) \\ &= m_{\square} + m_{\square} + m_{\heartsuit} + m_{\spadesuit} \\ &= m_{\square} + \mu(x_1 \cup x_2). \end{aligned}$$

The element that is double counted here, however, is exactly the lone element in the intersection of the two subsets (Figure 13),

$$m_{\square} = \mu(\{\square\}) = \mu(x_1 \cap x_2).$$

In other words we can write

$$\mu(x_1) + \mu(x_2) = \mu(x_1 \cap x_2) + \mu(x_1 \cup x_2).$$

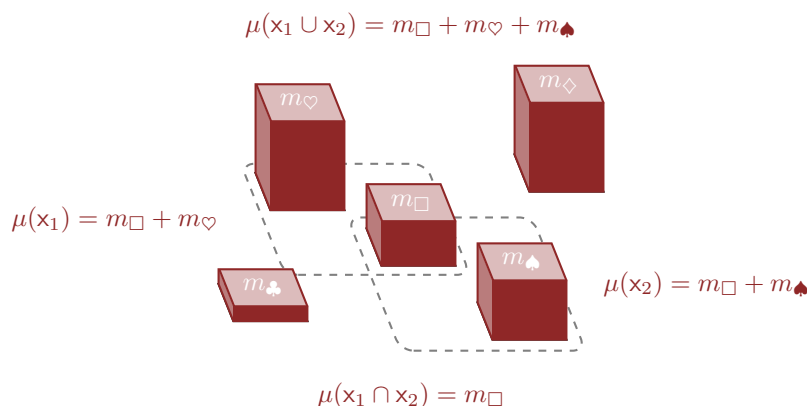


Figure 13: When two subsets overlap the measure allocated to each double counts the measure allocated to any overlapping elements, here  $\square$ , but the measure allocated to their union does not. This results in an important relationship between the measures allocated to the two subsets, the measure allocated to their union, and the measure allocated to their intersection.

This relationship is quite general and holds for *any* two subsets regardless of their overlap.

These subset properties allow us to build up a given measure in many different ways, each of which can be useful in different circumstances. This provides convenient flexibility when trying to apply measure theory and probability theory in practice.

For example we can always specify a measure *globally* by specifying the individual allocations at the same time (Figure 14). Alternatively we can specify the allocation *locally* by allocating measure to each element one at a time (Figure 15).

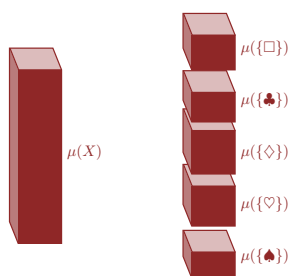


Figure 14: Measures can be constructed by specifying the individual element allocations all at once.



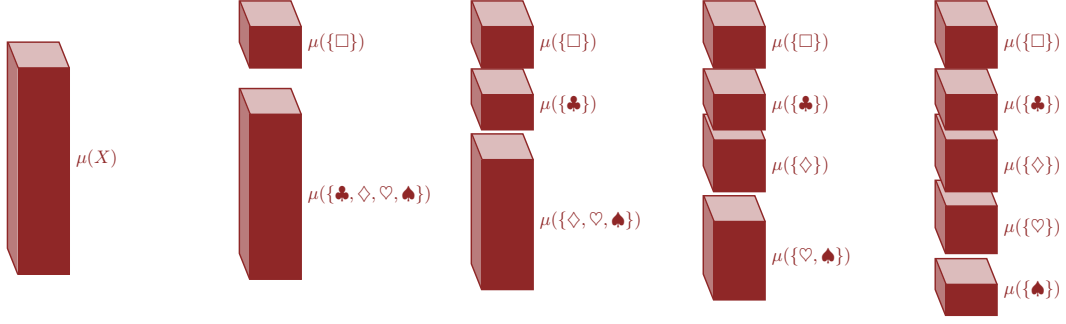


Figure 15: At the same time measures can be constructed by specifying the individual element allocations one by one.

Critically we do not always need to start with individual allocations. Instead we can always start by allocating the total measure to disjoint subsets and then iteratively *refining* that allocation to smaller and smaller subsets until we reach the individual elements (Figure 16).

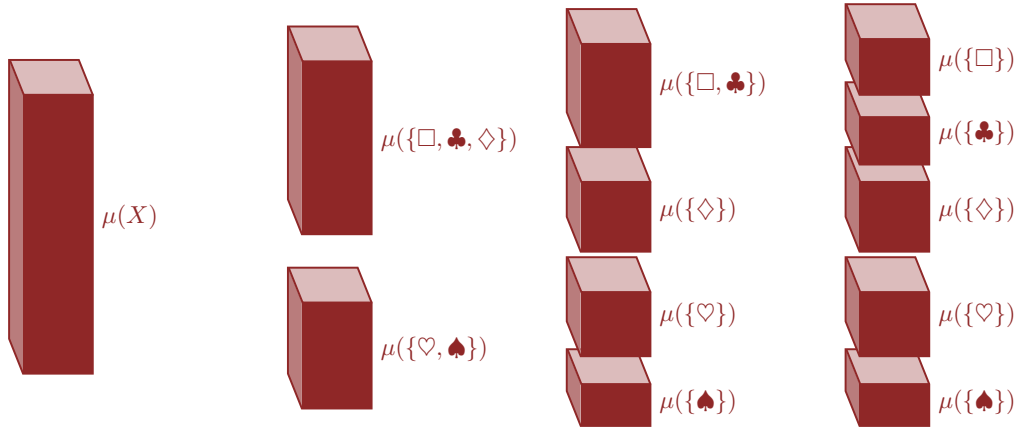


Figure 16: Measures can also be constructed by allocating the total measure to disjoint subsets and then iteratively refining that allocation to smaller and smaller subsets.

In addition to providing flexible constructions the subset definition of a measure  $\mu : 2^X \rightarrow [0, \infty]$  is also critical for generalizing measure theory beyond finite sets. Specifically it becomes *necessary* when trying to consistently define measures on more mathematically complicated sets like the real line. This will be the topic of Chapter 4.

## 4 Acknowledgements

I thank Simon Duane and Edgar Merkle for helpful comments.

A very special thanks to everyone supporting me on Patreon: Aapeli Nevala, Adam Bartonicek, Adam Fleischhacker, Adan, Adriano Yoshino, Alan Chang, Alessandro Varacca, Alexander Bartik, Alexander Noll, Alexander Rosteck, Anders Valind, Andrea Serafino, Andrew Mascioli, Andrew Rouillard, Andrew Vigotsky, Angie\_Hyunji Moon, Ara Winter, asif zubair, Austin Rochford, Austin Rochford, Avraham Adler, Ben Matthews, Ben Swallow, Brynjolfur Gauti Jónsson, Cameron Smith, Canaan Breiss, Cat Shark, Charles Naylor, Chase Dwelle, Chris Zawora, Christopher Mehrvarzi, Chuck Carlson, Colin Carroll, Colin McAuliffe, Cruz, Damien Mannion, Damon Bayer, dan mackinlay, Dan Muck, Dan W Joyce, Dan Waxman, Dan Weitzenfeld, Daniel, Daniel Edward Marthaler, Daniel Rowe, Darshan Pandit, Darthmaluus, David Burdelski, David Galley, David Humeau, David Wurtz, dilsher singh dhillon, Doug Rivers, Dr. Jobo, Dr. Omri Har Shemesh, Ed Cashin, Ed Henry, Edgar Merkle, edith darin, Eric LaMotte, Erik Banek, Ero Carrera, Eugene O’Friel, Evan Cater, Fabio Pereira, Fabio Zotte, Felipe González, Fergus Chadwick, Finn Lindgren, Florian Wellmann, Francesco Corona, Geoff Rollins, George Ho, Granville Matheson, Greg Sutcliffe, Guido Biele, Hamed Bastan-Hagh, Haonan Zhu, Hector Munoz, Henri Wallen, Hugo Botha, Huib Meulenbelt, Håkan Johansson, Ian Costley, Ian Koller, idontgetoutmuch, Ignacio Vera, Ilaria Prosdocimi, Isaac S, Isaac Vock, J, J Michael Burgess, Jair Andrade, James Hodgson, James McInerney, James Wade, Janek Berger, Jason Martin, Jason Pekos, Jeff Burnett, Jeff Dotson, Jeff Helzner, Jeffrey Burnett, Jeffrey Erlich, Jesse Wolfhagen, Jessica Graves, Joe Wagner, John Flournoy, Jon, Jonathan H. Morgan, Jonathan St-onge, Jonathon Vallejo, Joran Jongerling, Joseph Despres, Josh Weinstock, Joshua Duncan, Joshua Griffith, Josué Mendoza, JU, Justin Bois, Karim Naguib, Karim Osman, Keith O’Rourke, Kejia Shi, Kevin Foley, Kristian Gårdhus Wichmann, Kádár András, lizzie, LOU ODETTE, Luiz Pessoa, Marc Dotson, Marc Trunjer Kusk Nielsen, Marcel Lüthi, Marek Kwiatkowski, Mark Donoghoe, Mark Worrall, Markus P., Martin Modrák, Matt Moores, Matthew, Matthew Kay, Matthieu LEROY, Maurits van der Meer, Merlin Noel Heidemanns, Michael Colaresi, Michael DeWitt, Michael Dillon, Michael Lerner, Mick Cooney, Márton Vaitkus, N Sanders, Name, Nathaniel Burbank, Nic Fishman, Nicholas Clark, Nicholas Cowie, Nick S, Nicolas Frisby, Octavio Medina, Ole Rogeberg, Oliver Crook, Olivier Ma, Pablo León Villagrà, Palwasha Khan, Patrick Kelley, Patrick Boehnke, Pau Pereira Batlle, Peter Smits, Pieter van den Berg, ptr, Putra Manggala, Ramiro Barrantes Reynolds, Ravin Kumar, Raúl Peralta Lozada, Riccardo Fusaroli, Richard Nerland, Robert Frost, Robert Goldman, Robert kohn, Robert Mitchell V, Robin Taylor, Rong Lu, Ross McCullough, Ryan Grossman, Rémi, S Hong, Sam Levy, Scott Block, Scott Brown, Sean Pinkney, Sean Wilson, Serena, Seth Axen, shira, Simon Duane, Simon Lilburn, Srivatsa Srinath, sssz, Stan\_user, Stefan, Stephanie Fitzgerald, Stephen Lienhard, Steve Bertolani, Stone Chen, Sus, Susan Holmes, Svilup, Sören Berg, Tagir Akhmetshin, Tao Ye, Tate Tunstall, Tatsuo Okubo, Teresa Ortiz, Thiago de Paula Oliveira, Thomas Lees, Thomas Vladeck, Tiago Cabaço, Tim Radtke, Tom McEwen, Tomáš Frýda, Tony Wuersch, Virginia Fisher, Vitaly Druker, Vladimir

Markov, Wil Yegelwel, Will Farr, Will Kurt, Will Tudor-Evans, woejozney, yolhaj , yureq , Zach A, Zad Rafi, and Zhengchen Cai.

## License

A repository containing all of the files used to generate this chapter is available on [GitHub](#).

The text and figures in this chapter are copyrighted by Michael Betancourt and licensed under the CC BY-NC 4.0 license:

<https://creativecommons.org/licenses/by-nc/4.0/>