Interpretations of probability come into play when its mechanics (measure theory) are used to model real situations. The goal of this document is to present probability as formally as I can.

- Probability theory for engineers
- Probability and Measure (Billingsey).

> Once the philosophy has been stripped away, probability theory is simply the study of an object, a probability distribution that assigns values to sets, and the transformations of that object.

Probability is a positive conserved quantity that we can distribute across a space. This is that more formal notion of spreading 'cream cheese'.

A measure space is a mathematical object that is defined by a triple:

$$(X, A, \mu)$$

where X is a set, A is a σ-algebra on the set and μ is a measure. A *measure* is a particular kind of function that maps from the A space to the real number line.

Measure theory aims to abstract the notion of 'size' [1].

The measure function assigns a size to each element in A. It is the function

$$\mu : A \rightarrow \Re$$

Intuitively each element of the σ-algebra is a part of a larger whole. So that if we add up each individual element we get some 'whole' (countable additivity). It's a particular set of subsets of X.

We define a *measurable space* as (X, A) the initial set and σ-algebra, the subsets that can be measured. The σ-algebra set is often the power set of X.

Measurable function (probability)

A function P is a probability measure for the probability space (Ω, F) if it satisfies:

- $0 \leq P(A) \leq 1$ for $A \in F$.
- $P(\emptyset) = 0, P(\Omega) = 1$
- $A_1, ...A_n$ is a disjoint sequence of F sets, then $P(\cap_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} P(A_k)$.

[1] https://mbernste.github.io/posts/measure_theory_1/

Basics

Here we start measuring the ‘size’ of uncertainty.

Define Ω as the unit interval (0, 1]. The length of the unit interval I is

$$|I| = b - a$$

,

$$A = \cup_{i=1}^{n} I_i$$

The set Ω can represent all future possible worlds.

Provided A is disjoint and finite and lies in Ω then we assign a measure of probability

$$P(A) = \sum_{i=1}^{n} (b_i - a_i)$$

A is a set of subsets that is a σ-algebra.

Functions/Transforms

A probability distribution is a mapping of the form

$$\pi : X \to [0, 1]$$

for each atomic element in X.

Expectation

A distribution as defined allows us to summarise the distribution function with numbers.

Reduction of functions of the form $f : X \to R$ to a single number.

If C is the space with all functions of the form $f : X \to R$ then expectation is a map

$$f$$

Definitions

A space is usually denoted by its set and structure (X, s). An example here might be the ordering of the set with some ordering rule (like in decision theory).

The indicator, or indicator function, of a set A is the function on Ω that assumes the value 1 on A and 0 on A^C it is denoted I_A.