Measure Theory

A very general overview

Measure Theory: Why do we care as data scientists?

- Measure theory is the foundation of probabilities in the continuous setting.
    - It's concerned with what sets can form the universe of events in a given setting, and how to assign probabilities to these sets in a way that makes sense.

Measure Theory: Why do we care as data scientists?

- Measure theory helps statisticians handle integrations and limits or convergences of these integrations, common when statisticians try to figure out asymptotic distribution, the distribution of a statistic when the sample size goes to infinity.

- Statisticians do a lot of integrations, probability, moments, distribution are all integrals, and these integrations get very nasty very quickly. Not only are you dealing with multiple integrals with respect to multiple probability distributions or measures and then later stochastic processes, you are also at the same time handling sums of infinite series, then throw in a good measure of countable and uncountable unions and intersections, you have roughly what theoretical statisticians see daily.

Measure Theory: Why do we care as data scientists?

- Once simple things like switching integrals, switching integrals and sums, switching limits with integrals, etc, now requires "heavy machineries" to move. Measure theory provides the "heavy machineries".

- For example, something as "trivial" as the **Central Limit Theorem** would be impossible to prove without measure theory.

- Let's hear Joseph Wang, Chief Scientist at Bitquant Research, give us an example...
  https://www.quora.com/Why-is-measure-theory-so-important-in-probability-theory-and-is-this-also-the-case-for-applications

Measure Theory: Basics

- Explain what the set of all subsets is.
- Explain what a measure is.
- Explain what integration with respect to a measure is and how to compute an integral numerically.

Let's talk about all the things that we can do with two sets.

1. $A \cup B$ is the **union** of A and B and consists of all elements that are either in A or in B (OR Logic)
2. $A \cap B$ is the **intersection** of A and B and contains all elements that are in A and B (AND Logic)
3. $A \setminus B$ is the **difference** and consists of all points in A that are not in B.
4. Two sets A and B are called disjoint if $A \cap B = \varnothing$ which means that any point in A is not in B and vice versa.
5. Let $A \cap B = \varnothing$ be two disjoint sets. We call $A \uplus B$ the **disjoint union** of A and B. (This notation means the same as 1. We are only emphasising that A and B are disjoint)

Let's take any set that we would like and call it $\mathbb{A}$. We can now look at a very special set namely the set of all subsets of $\mathbb{A}$ which is denoted by

$$\mathbb{P}(\mathbb{A}) = \{\text{All subsets of } \mathbb{A}\}$$