Measure Theory and the Radon–Nikodym Theorem

Sai Nallani
sai.s.nallani@gmail.com

Euler Circle

July 9, 2024

Table of Contents

The Essence of Measure Theory

Lengths, Areas, and Volumes

Let us say we have a set X. How can we compare sizes of subsets of X? One way is cardinalities, but this approach is not useful when we are dealing two uncountable subsets.

Lengths, Areas, and Volumes

Let us say we have a set X. How can we compare sizes of subsets of X?
One way is cardinalities, but this approach is not useful when we are dealing two uncountable subsets.
For example, in the real line the subsets $[0, 1]$ and $[0, 2]$ have the same cardinality though $[0, 2]$ is, in a way, bigger.
For intervals in $\mathbb{R}$, we might focus on their lengths. How about $\mathbb{R}^2$ and $\mathbb{R}^3$?

Lengths, Areas, and Volumes

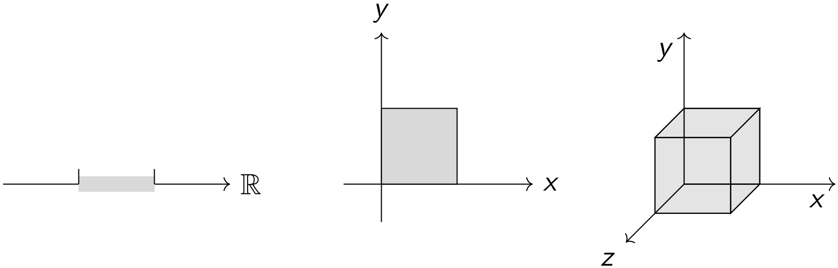
Let us say we have a set X. How can we compare sizes of subsets of X? One way is cardinalities, but this approach is not useful when we are dealing two uncountable subsets.

For example, in the real line the subsets $[0, 1]$ and $[0, 2]$ have the same cardinality though $[0, 2]$ is, in a way, bigger.

For intervals in $\mathbb{R}$, we might focus on their lengths. How about $\mathbb{R}^2$ and $\mathbb{R}^3$? We use areas and volumes, respectively.

Lengths, Areas, and Volumes

Let us say we have a set X. How can we compare sizes of subsets of X? One way is cardinalities, but this approach is not useful when we are dealing two uncountable subsets.
For example, in the real line the subsets $[0, 1]$ and $[0, 2]$ have the same cardinality though $[0, 2]$ is, in a way, bigger.
For intervals in $\mathbb{R}$, we might focus on their lengths. How about $\mathbb{R}^2$ and $\mathbb{R}^3$? We use areas and volumes, respectively.



Measuring Subsets of $\mathbb{R}$

How can we find the "length" of the rationals $\mathbb{Q} \subset \mathbb{R}$? How about the Cantor set $C \subseteq \mathbb{R}$? We can't just measure its length as easily as for the intervals.

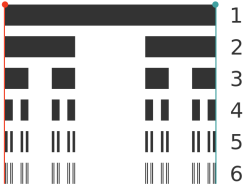


Figure 1: The Cantor set is created by repeatedly taking the middle third interval.

What We Want

We want a function μ taking subsets of $\mathbb{R}$ to $[0, +\infty]$ such that the function satisfies the following properties:

1. The domain of μ is $\mathcal{P}(\mathbb{R})$. In other words, μ can take any subset of $\mathbb{R}$ as an input.
2. If $I \subset \mathbb{R}$ is an interval, then $\mu(I)$ is the length of the interval.
3. If $S_1, S_2, \ldots$ is a countable collection of pairwise disjoint sets, then
$$\mu\left(\bigcup_{n=1}^{\infty} S_n\right) = \sum_{n=1}^{\infty} \mu(S_n).$$
4. (Translation Invariant) Let $A \subseteq \mathbb{R}$ and define $A + x = \{a + x : a \in A\}$ for all $x \in \mathbb{R}$. Then we want $\mu(A) = \mu(A + x)$. Moving the set on the number line shouldn't change its measure.

Bad News

See Folland's *Real Analysis* for a proof on why such a function does not exist on $\mathbb{R}$ (or google it up)!

Bad News

See Folland's *Real Analysis* for a proof on why such a function does not exist on $\mathbb{R}$ (or google it up)!

This means that we need to focus on certain subsets of $\mathcal{P}(\mathbb{R})$...

Measures and σ-algebras

σ-algebras

Definition 2.1

A σ-**algebra** $\mathcal{A} \subseteq \mathcal{P}(X)$ on X is a set such that

1. $\varnothing \in \mathcal{A}$
2. If $S \in \mathcal{A}$, then $S^c \in \mathcal{A}$, where $S^c = X \setminus S$ is the complement of S.
3. If $S_1, S_2, \ldots$ is a countable collection of elements in $\mathcal{A}$, then

$$\bigcup_{n=1}^{\infty} S_n \in \mathcal{A} \text{ and } \bigcap_{n=1}^{\infty} S_n \in \mathcal{A}.$$

The pair $(X, \mathcal{A})$ is called a **measurable space**.

Essentially, $\mathcal{A}$ is a set of subsets that contain the empty set, and is closed under complements and countable unions.

Measures

Definition 2.2

Given a set X and a σ-algebra $\mathcal{A}$, a **measure** on $(X, \mathcal{A})$ is a function $\mu : \mathcal{A} \to [0, +\infty]$ such that

1. $\mu(\varnothing) = 0$
2. If $S_1, S_2, \ldots$ is a countable collection of pairwise disjoint elements in $\mathcal{A}$, then
$$\mu\left(\bigcup_{n=1}^{\infty} S_n\right) = \sum_{n=1}^{\infty} \mu(S_n).$$

The triplet $(X, \mathcal{A}, \mu)$ is called a **measure space**.

Examples of Measure Spaces

Example 1

Let X be non-empty, and take $x \in X$. Let us again take $\mathcal{A} = \mathcal{P}(X)$ Then the **point-mass measure** $\mu_x : \mathcal{A} \to [0, \infty]$ is defined as

$$\mu_x(A) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}.$$

The measure space is $(X, \mathcal{P}(X), \mu_x)$.

Examples of Measure Spaces

Example 1

Let X be non-empty, and take $x \in X$. Let us again take $\mathcal{A} = \mathcal{P}(X)$ Then the **point-mass measure** $\mu_x : \mathcal{A} \to [0, \infty]$ is defined as

$$\mu_x(A) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}.$$

The measure space is $(X, \mathcal{P}(X), \mu_x)$.

Example 2

Given a set X and σ-algebra $\mathcal{A} = \mathcal{P}(X)$ we can define a measure $\mu : \mathcal{A} \to [0, \infty]$ by $\mu(A) = |A|$ for finite sets A. If A is infinite, we can let $\mu(A) = \infty$. μ is known as the **counting measure**. The measure space is $(X, \mathcal{P}(X), \mu)$.

Borel σ-Algebra

Definition 2.3

The **Borel σ-algebra** $\mathcal{B}(\mathbb{R})$ on $\mathbb{R}$ is the smallest σ-algebra generated by all open sets of $\mathbb{R}$. In general, the Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$ on the n-dimensional Euclidean space $(\mathbb{R}^n, d)$, where d is the n-dimensional Euclidean norm is the smallest σ-algebra generated by all open sets of $\mathbb{R}^n$. A set is a **Borel** set if it is in the Borel σ-algebra.

Borel σ-Algebra

Definition 2.3

The **Borel** σ**-algebra** $\mathcal{B}(\mathbb{R})$ on $\mathbb{R}$ is the smallest σ-algebra generated by all open sets of $\mathbb{R}$. In general, the Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$ on the n-dimensional Euclidean space $(\mathbb{R}^n, d)$, where d is the n-dimensional Euclidean norm is the smallest σ-algebra generated by all open sets of $\mathbb{R}^n$. A set is a **Borel** set if it is in the Borel σ-algebra.

$\mathcal{B}(\mathbb{R})$ is the smallest σ-algebra containing all intervals.

Lebesgue Measure

Outer Measures

Definition 3.1 (Outer measure)

Let X be a set. A function $\mu^* : \mathcal{P}(X) \to [0, \infty]$ is an **outer measure** if it has the following properties:

1. $\mu^*(\varnothing) = 0$
2. μ^* is **monotonic**, meaning if $A \subseteq B \subseteq X$, then $\mu^*(A) \leq \mu^*(B)$.
3. If $A_1, A_2, A_3, \ldots$ is a countable collection of sets, then we have

$$\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu^*(A_n).$$

This property is called countable subadditivity.

Lebesgue Outer Measure

The **Lebesgue outer measure** on $\mathbb{R}$, denoted by λ^*, is defined as follows. For each subset A of $\mathbb{R}$, let $\mathcal{C}_A$ be the set of all infinite sequences $\{(a_i, b_i)\}$ of bounded open intervals such that $A \subseteq \cup_i (a_i, b_i)$. In other words, $\mathcal{C}_A$ contains sequences of all open intervals such that their union contains A. Then, $\lambda^* : \mathcal{P}(\mathbb{R}) \to [0, \infty]$ is defined by

$$\lambda^*(A) = \inf \left\{ \sum_i (b_i - a_i) : \{(a_i, b_i)\} \in \mathcal{C}_A \right\}.$$

We take the sum of lengths of sequences of open intervals such that their union just barely cover the set we are trying to measure.

See proof in the paper for how Lebesgue outer measure is actually an outer measure, and it satisfies the properties of a measure we want on $\mathbb{R}$.

Carathéodory's Condition

Definition 3.2

(Carathéodory's Condition or μ^*-measurability) Let X be a set, and let $\mu^* : \mathcal{P}(X) \to [0, \infty]$ be an outer measure on X. A set $S \subseteq X$ is μ^*-**measurable** (or just measurable if the measure we are using in the context is clear), if for each arbitrary set $A \subseteq X$, we have

$$\mu^*(A) = \mu^*(A \cap S) + \mu^*(A \cap S^c).$$

A set is measurable if we can divide the set in such a way that the sizes of the pieces add properly.

Lebesgue Measure

Proposition 3.3

Let $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be a measurable space. All sets $S \in \mathcal{B}(\mathbb{R})$ are λ^*-measurable.

Theorem 3.4

Let X be a set and μ^* be an outer measure on X. Let us denote $\mathscr{M}_{\mu^*}$ as the set of μ^*-measurable sets. Then $\mathscr{M}_{\mu^*}$ is a σ-algebra. Furthermore, μ^* restricted to $\mathscr{M}_{\mu^*}$ is a measure.

We call $\lambda : \mathscr{M}_{\lambda^*} \to [0, +\infty]$ defined by $\lambda(A) = \lambda^*(A)$ the **Lebesgue measure**. We call sets $S \in \mathscr{M}_{\lambda^*}$ **Lebesgue sets**.

Integration

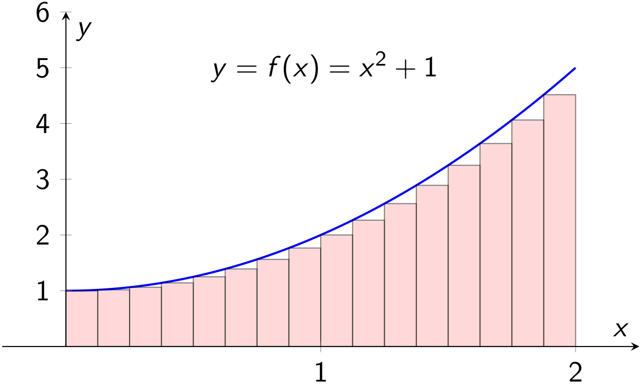
Riemann Integral

In a Calculus course, we may have been introduced to the Riemann integral:

$$\int_a^b f(x)\,dx.$$

It is typically defined as a summation of rectangles of partitions of $[a, b]$ as the partition gets finer and finer.

Riemann Integral Visualization



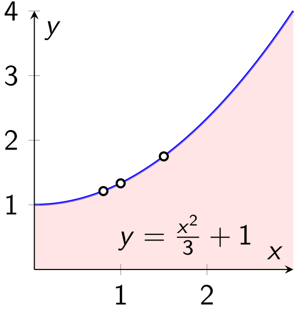
As the width of the rectangles gets smaller, the sum of the areas approaches the area under the curve.

Limitations of Riemann Integral

The Riemann integral is good at integrating continuous functions, and even some discontinuous functions as long as the set of points of discontinuity have measure zero.

Limitations of Riemann Integral

The Riemann integral is good at integrating continuous functions, and even some discontinuous functions as long as the set of points of discontinuity have measure zero.



Limitations of Riemann Integral Continued

Consider the function $f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$. Since it is discontinuous almost everywhere (only continuous at $x = 0$), it is not Riemann-integrable on the interval $[0, 1]$. But there are only a few points with $f(x) = 1$ on $[0, 1]$, as there are a lot more irrationals than rationals. It seems like the area under $f(x)$ on $[0, 1]$ should just be zero.

Integral of Indicator functions

Using measure theory, we can expand the definition of integration to allow functions like the one in the previous slide to also be integrated.

Integral of Indicator functions

Using measure theory, we can expand the definition of integration to allow functions like the one in the previous slide to also be integrated.

Definition 4.1

Let X be a set, and let $A \subseteq X$. An **indicator function** $\mathbb{1}_A$ is defined as follows:

$$\mathbb{1}_A = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}.$$

Integral of Indicator functions

Using measure theory, we can expand the definition of integration to allow functions like the one in the previous slide to also be integrated.

Definition 4.1

Let X be a set, and let $A \subseteq X$. An **indicator function** $\mathbb{1}_A$ is defined as follows:

$$\mathbb{1}_A = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}.$$

Definition 4.2

Let $(X, \mathcal{A}, \mu)$ be a measure space, and let $A \in \mathcal{A}$. The **(Lebesgue) integral** of $\mathbb{1}_A$ is defined as

$$\int_X \mathbb{1}_A \, d\mu = \mu(A).$$

Resolving Integral of $\mathbb{1}_{\mathbb{Q}}$

We see that $\int_X \mathbb{1}_{\mathbb{Q}}\, d\mu = \mu(\mathbb{Q}) = 0$, and this lines up with our original intuition, as $\mathbb{1}_{\mathbb{Q}}$ is basically the zero function for most points.

Now What Do We Do?

How can we extend this from indicator functions? Since we want this new theory of integration to be better than Riemann's, we still want to preserve the Riemann integral's nice properties of linearity.

Namely, we want

$$\int_X a\mathbb{1}_A d\mu = a\int_X \mathbb{1}_A d\mu = a\mu(A)$$

for real $a \in \mathbb{R}$ and we also want

$$\int_X (\mathbb{1}_A + \mathbb{1}_B) d\mu = \int_X \mathbb{1}_A d\mu + \int_X \mathbb{1}_B d\mu = \mu(A) + \mu(B).$$

Therefore, we can extend our (and Lebesgue's) integral by defining **simple functions**, which are finite linear combinations of indicator functions!

Simple Functions

Definition 4.3 (Simple functions)

Let $(X, \mathcal{A}, \mu)$ be a measure space. A **simple function** is a function $f : X \to \mathbb{R}$ such that there exist finitely many measurable subsets $A_1, A_2, \ldots, A_n \in \mathcal{A}$ and real numbers $c_1, \ldots, c_n$ such that

$$f(x) = \sum_{i=1}^{n} c_i \mathbb{1}_{A_i}(x).$$

Simple Functions Continued

The integral of a simple function f agrees with out intuition laid out previously.

Definition 4.4 (Lebesgue Integral of a simple function)

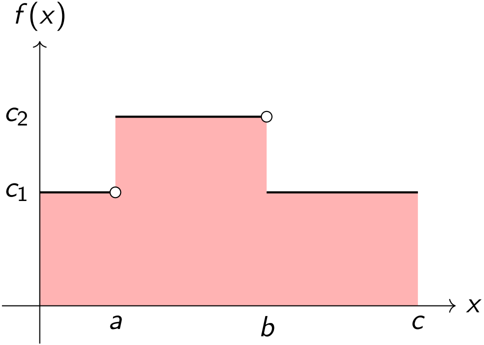
Let $(X, \mathcal{A}, \mu)$ be a measure space, and let $f(x) = \sum_{i=1}^{n} c_i \mathbb{1}_{A_i}(x)$ be a simple function. Then we define the integral of f to be

$$\int_X f d\mu = \sum_{i=1}^{n} c_i \mu(A_i).$$

Simple Functions Continued

Simple functions look like piecewise horizontal functions. For example, this is the graph of

$$f = c_1 \mathbb{1}_{[0,a)} + c_2 \mathbb{1}_{[a,b)} + c_1 \mathbb{1}_{[b,c]}.$$



$\mathcal{A}$-measurable Functions

Definition 4.5 (Measurable extended real-valued functions)

Let $(X, \mathcal{A})$ be a measurable space, and let $f : X \to [-\infty, +\infty]$. We say f is $\mathcal{A}$-**measurable** (or just **measurable**) if for each real number $t \in \mathbb{R}$ the set $\{x \in A : f(x) \leq t\}$ belongs to $\mathcal{A}$.

Definition 4.6

Let $(X, \mathcal{A}, \mu)$ be a measure space. Denote $\mathscr{S}$ to be the set of measurable real-valued simple functions, and denote $\mathscr{S}_+ \subseteq \mathscr{S}$ to be the set of measurable non-negative simple functions.

Measurable Functions are Limits of Simple Functions

The following Lemma gives us an indication that simple functions are the correct framework for defining our (well, actually Lebesgue's) new integral.

Lemma 4.7 (Measurable functions are limits of simple functions)

Let $(X, \mathcal{A})$ be a measurable space, let A be a subset of X that belongs to $\mathcal{A}$, and let f be a $[0, +\infty]$-valued measurable function on A. Then there is a sequence $\{f_n\}$ of simple $[0, +\infty)$-valued measurable functions on A that satisfy

$$f_1(x) \leq f_2(x) \leq \dots$$

and

$$f(x) = \lim_n f_n(x)$$

at each x in A.

Proof of Lemma 4.7

Proof.

For each positive integer n and for $k = 1, 2, \ldots, n2^n$ let $A_{n,k} = \{x \in A : (k-1)/2^n \leq f(x) < k/2^n\}$. The measurability of f implies that each $A_{n,k}$ belongs to $\mathcal{A}$. Define a sequence $\{f_n\}$ of functions from A to $\mathbb{R}$ by requiring f_n to have value $(k-1)/2^n$ at each point in $A_{n,k}$ (for $k = 1, 2, \ldots, n2^n$) and to have value n at each point in $A - \bigcup_k A_{n,k}$. The functions so defined are simple and measurable, and it is easy to check that they satisfy (1) and (2) at each x in A. ■

Picture of Lemma 4.7

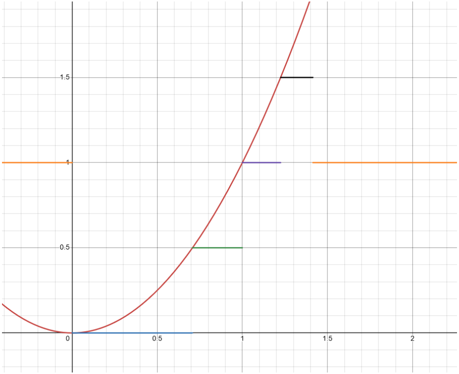


Figure 2: Approximating the function $f(x) = x^2$ with $n = 1$ by using the proof, we have these four "steps".

Lebesgue Integral of $\mathcal{A}$-Measurable functions

Definition 3 (Lebesgue integral of measurable functions $f : X \to [0, +\infty)$)

Let(X, A, μ) be a measure space, and let $f : X \to [0, +\infty]$ to be a measurable function. We define its **Lebesgue integral** to be

$$\int_X f \, d\mu = \sup \left\{ \int_X s \, d\mu : s \in \mathscr{S} \text{ and } s(x) \leq f(x) \text{ for all } x \right\},$$

if this suprenum is finite. When the suprenum is finite, we say that f is **integrable**.

We define the Lebesgue integral of f over A as

$$\int_A f \, d\mu = \int_X f \cdot \mathbb{1}_A \, d\mu.$$

Integrating $[-\infty, +\infty]$-valued functions

What if we want to integrate over functions that take up negative values?

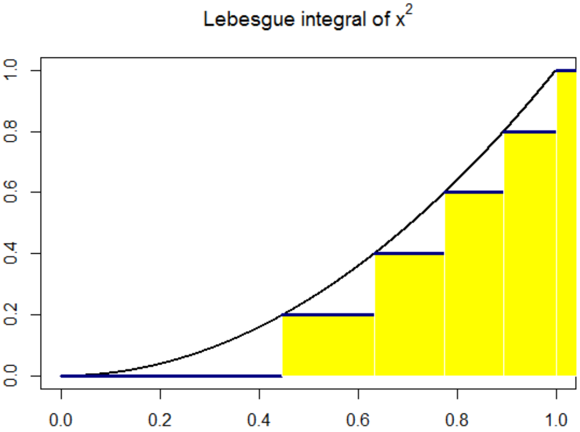
Suppose $f : X \to [-\infty, +\infty]$ is measurable. Define two functions, splitting f into a positive part and negative part, as follows:

$$f^+(x) = \max(f(x), 0), \qquad f^-(x) = \max(0, -f(x)).$$

Now we define the integral of f to be

$$\int_X f d\mu = \int_X f^+ d\mu - \int_X f^- d\mu.$$

Simple Functions Partition Measurable Functions



Limit Theorems

Almost Everywhere and Almost Nowhere

Let $(X, \mathcal{A}, \mu)$ be a measure space. A property holds true almost everywhere if the set of points where the property doesn't hold true has measure zero. A property holds true almost nowhere if the set of points where the property holds true has measure zero.

Monotone Convergence Theorem

Theorem 5.1

(Monotone Convergence Theorem) Let $(X, \mathcal{A}, \mu)$ *be a measure space, let* f *and* $f_1, f_2, \ldots$ *be* $[0, +\infty]$*-valued* $\mathcal{A}$*-measurable functions on* X*, Suppose that the relations*

$$f_1(x) \leq f_2(x) \leq \ldots \tag{1}$$

and

$$f(x) = \lim_n f_n(x) \tag{2}$$

hold at almost every x *in* X*. Then* $\int_X f \, d\mu = \lim_n \int f_n \, d\mu$*.*

Signed Measures

Signed Measures

Definition 6.1 (Signed Measure)

Let $(X, \mathcal{A})$ be a measurable space, and let $\mu : \mathcal{A} \to [-\infty, +\infty]$. If the function μ is **countably additive**, meaning the identity

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$$

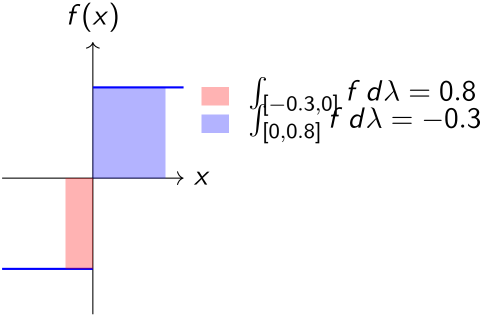
holds for each sequence of pairwise disjoint sets $A_1, A_2, \ldots$ and $\mu(\varnothing) = 0$, then it is a **signed measure**.

Example

Consider the positive measure space $(X, \mathcal{A}, \mu)$, and let $f : X \to [-\infty, +\infty]$ be integrable. Then $\nu(A) = \int_A f \, d\mu$ is a signed measure!

Another Example of Signed Measures

Let $(\mathbb{R}, \mathcal{A}, \lambda)$ be a measure space under the Lebesgue measure. Define $f(x) = 1$ if $x \geq 0$ and $f(x) = -1$ if $x < 0$. Then, we can see that $\nu(A) = \int_A f \, d\lambda$ is a measure. Furthermore, we can interpret ν as being a measure that gives negative weight to sets with negative numbers. For example, we can see that $\nu([-0.3, 0.8]) = 0.5$. We might ask ourselves if every signed measure can be split up into a positive and negative measure.



Positive and Negative Sets

Definition 6.2 (Positive and negative sets)

Let μ be a signed measure on the measurable space $(X, \mathcal{A})$. A subset A of X is a **positive set** for μ if $A \in \mathcal{A}$ and eqach $\mathcal{A}$-measurable subset E of A satisfies $\mu(E) \geq 0$. Likewise A is a **negative set** for μ if $A \in \mathcal{A}$ and for each $\mathcal{A}$-measurable subsets E of A we have $\mu(E) \leq 0$.

A natural question to now ask is if we can split up X into sets P and N such that P is positive and N is negative.

Splitting Signed Measure Spaces

Theorem 6.3 (Hahn Decomposition Theorem)

Let $(X, \mathcal{A})$ be a measurable space, and let μ be a signed measure on $(X, \mathcal{A})$. Then there are disjoint subsets P and N of X such that P is a positive set for μ, N is a negative set for μ, and $X = P \cup N$.

Corollary 6.4 (Jordan Decomposition Theorem)

Every signed measure is the difference of two positive measures, at least one of which is finite.

Radon–Nikodym Theorem

Absolute Continuity

Definition 7.1

Let $(X, \mathcal{A})$ be a measurable space, and let μ and ν be positive measures on $(X, \mathcal{A})$. Then ν is **absolutely continuous with respect to** μ if for each set A in $\mathcal{A}$ that satisfies $\mu(A) = 0$ also satisfies $\nu(A) = 0$.

Example

Define $\nu(A) = \int_A f \, d\mu$, then ν is absolutely continuous with respect to μ. This is because if $\mu(A) = 0$, the integral vanishes; $\nu(A) = \int_X f \mathbb{1}_A \, d\mu$ and $f \mathbb{1}_A = 0$ almost everywhere, so $\nu(A) = 0$ too.

Definition 7.2 (σ-finite measures)

Let $(X, \mathcal{A})$ be a measurable space, and let μ be a positive measure on $(X, \mathcal{A})$. The measure μ is σ-finite if X is a countable union of sets of finite measure. That is, $X = \bigcup_{i=1}^{\infty} A_i$ where $\mu(A_i) < +\infty$ for all i.

Radon–Nikodym Theorem

If we make a new measure ν by integrating over a function in the measure space $(X, \mathcal{A}, \mu)$, we have that ν is absolutely continuous with respect to μ.

What about the converse? If ν is absolutely continuous with respect to μ, is ν the integral of some function?

Theorem 7.3 (Radon–Nikodym theorem)

Let $(X, \mathcal{A})$ be a measurable space, and let μ and ν be σ-finite positive measures on $(X, \mathcal{A})$. If ν is absolutely continuous with respect to μ, then there is an $\mathcal{A}$-measurable function $g : X \to [0, +\infty)$ such that $\nu(A) = \int_A g d\mu$ holds for each A in $\mathcal{A}$. The function g is unique up to μ-almost everywhere equality.

We call g the Radon–Nikodym derivative.

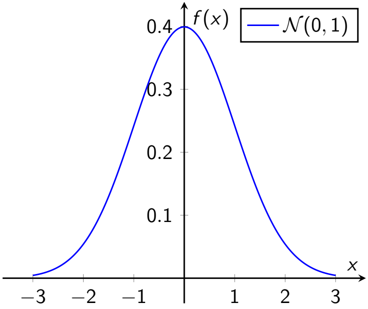
Significance of Radon–Nikodym Theorem

Let us just zoom in on the measurable space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Any measure μ we come up with on that space, if $\mu(A) = 0$ whenever $\lambda(A) = 0$, we can write μ as an integral of some function g.

The theorem also has various applications in probability theory. The **Conditional expectation** of a random variable is its expected value evaluated with respect to the conditional probability distribution. Essentially, it is what we expect a random variable X to be, given another random variable Y.

Probability Density Functions

What we see below is a probability density function (or pdf) for a random variable X. More specifically, it is a normal distribution. Let the $f(x)$ represent the curve below. Define a new measure on the space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ by defining $\nu(A) = \int_A f \, d\mu$. We have $\nu(A)$ to give us the probability of A. We can say that f is a Radon–Nikodym derivative.



Thank you for listening. Please reach out to me on Discord (or email) if there are any questions.