2 Measure Theory

2.1 Measures

A lot of this exposition is motivated by Folland's wonderful text, "Real Analysis: Modern Techniques and Their Applications." Perhaps the most ubiquitous measure in our lives is the so-called *counting measure*, which, roughly speaking, gives every distinct object a value of 1. If I have ten distinct objects, then the measure of the union of these objects is precisely 10.

Now, this wonderful, but it is a very coarse measure. Suppose I want to take size or volume into account. For simplicity's sake, restrict yourself to subsets of $\mathbb{R}$. Probably the most natural definition of size in $\mathbb{R}$ is length. So, we define the measure of a subset $E \subset \mathbb{R}$ to be the length of the subset.

One very obvious problem with this is made clear in the following example: Let $A = [-100, -99] \cup [99, 100]$, and let $B = [-10, 10]$. Which set should have a greater measure? Intuition will probably lead you to say that B should have a greater measure. There really isn't much of a difficulty here, right? Well, which of the following sets should have a greater measure, $[0, 1]$ or $(0, 1]$? Things can get pretty complicated ...

Given a space, such as $\mathbb{R}^d$, we would love for there to be a type of function, $\mu : \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$, which would give us the "sizes" of sets in some meaningful way. That is, it would be grand if $\mu([0, 1]^d)$ was 1, μ was unaffected by rigid motions, and that disjoint union and addition played nicely together. Unfortunately, no such μ can exist.

Exercise 2.1. Find an example of a non-measurable subset of $\mathbb{R}$. (Hint: Consider the equivalence relation $\sim$, where $x \sim y$ iff $x - y \in \mathbb{Q}$. Use the Axiom of Choice to show the existence of the set N, a set consisting of one member from each equivalence class in $[0, 1)$. Now, consider all of the shifts, $N + r \pmod 1$, where r ranges through the rationals between zero and 1.)

Definition 2.1. A *measure* on is a non-negative function, on a subset $\mathcal{M} \subset \mathcal{P}(X)$, $\mu : \mathcal{M} \to [0, \infty]$, such that

- i) $\mu(\varnothing) = 0$,
- ii) For a sequence of sets $\{E_j\}$, $\mu(\bigcup_j U_j) \leq \sum_j \mu(E_j)$.

Since such nasty sets are out there, we probably don't want to let $\mathcal{M}$ be all of $\mathcal{P}(X)$ just yet.

Definition 2.2. Suppose X is a set, and $\mathcal{M} \subset \mathcal{P}(X)$ which contains $\varnothing$ and X, and is closed under complements and countable unions. Such an $\mathcal{M}$ is called a *σ-algebra*.

Now, suppose we restrict the domain of our measure to some suitable σ-algebra, $\mathcal{M} \subset \mathcal{P}(X)$. For this measure we will get the desired good behavior from rigid motions. In fact, if the E_j from ii) in Definition 2.1 are disjoint, we will get equality. There is more good news– we can, for most of what we want to do, extend a measure as in the definition above to include otherwise troublesome sets by defining the measure of such a set, $N \notin \mathcal{M}$, to be *completed*

$$\mu(N) = \inf \{\mu(M) : N \subset M \in \mathcal{M}\}. \tag{2.1}$$

Definition 2.3. One particularly useful σ-algebra called the *Borel sets* of $\mathbb{R}$. This is the smallest σ-algebra containing the open intervals. The Borel sets of $\mathbb{R}^d$ can be defined similarly.

Definition 2.4. The *Lebesgue measure* is the measure on the Borel sets of $\mathbb{R}$, which has been completed as in (2.1), which assigns a Borel set E the sum of the lengths of the intervals whose union is E. Here we use the convention that a singleton is an interval of length zero, and hence Lebesgue measure zero.

Exercise 2.2. Let m denote the Lebesgue measure, and let μ_c denote the counting measure. Compute the following:

- a) $m((a, b)), a < b$,
- b) $m\left(\bigcup_{j=0}^{\infty} E_j\right), E_j = \left[j, j + \frac{1}{j!}\right)$,
- c) $\mu_c\{5, 6, \pi$, a box of twelve apples, $-1\}$.

Definition 2.5. We will call a subset $A \subset X$ *μ-measurable* for a measure μ if, for every $E \subset X$,

$$\mu(E) = \mu(E \cap A) + \mu(E \setminus A).$$

Definition 2.6. We will often say that some condition holds *almost everywhere*. This means that within whatever space we are working, this condition holds everywhere except for possibly a set of measure zero. This is often abbreviated as *a.e.* in text.

2.2 Integration

While we have just scratched the surface of measure theory, hopefully this will suffice for what follows. One of the main tools for which measure theory is used is that of integration. Often times, we don't need anything fancier than Riemann integration. However, sometimes, we need a little bit more power. For example, suppose that we wanted to integrate the constantly 1 function over all of the rational numbers between 0 and 1. What should the integral be? Well, by the hitherto tried-and-true Riemann sums, we will get a lower sum of 0, and an upper sum of 1. This is a problem, but Lebesgue integration will give us an answer. Before we can get to that, here are some useful concepts for what follows.

Definition 2.7.

$$\liminf x_n = \lim_{n\to\infty}\left(\inf_{m\geq n} x_m\right), \text{ and } \limsup x_n = \lim_{n\to\infty}\left(\sup_{m\geq n} x_m\right).$$

Definition 2.8. Given a set X, and a subset $E \subset X$, the *characteristic* or *indicator* function of the set E is defined to be $\chi_E : X \to \{0, 1\}$,

$$\chi_e(x) = \begin{cases} 0, & x \notin E \\ 1, & x \in E \end{cases}$$

Definition 2.9. Given a set X, a *simple function* is a function $f : X \to \mathbb{R}$ of the form

$$f(x) = \sum_{j=1}^{N} a_j \chi_{E_j}(x),$$

where the E_j are subsets of X and the a_j are real numbers.

Now, we can compute the Lebesgue integral of a simple function over a set E as follows:

$$\int_E f = \sum_{j=1}^{N} a_j m(E \cap E_j).$$

This is great, but what about functions which are NOT simple? Well, it just so happens that we can still do something like this if we have a nice enough function.

Definition 2.10. A function $f : X \to \mathbb{R}$ is *measurable* if the pre-image of every element in its image is a measurable set. In other words, for every $y \in f(X)$, we have that $f^{-1}((-\infty, y))$ is a measurable set. Note, this is not the standard definition, but it will get the idea across.

Theorem 2.11. *If $f : X \to \mathbb{C}$ is measurable, then there is a sequence, $\{\phi_n\}$, of simple functions such that $0 \leq |\phi_1| \leq |\phi_2| \leq \cdots \leq |f|, \phi_n \to f$ pointwise, and $\phi_n \to f$ uniformly on any set on which f is bounded.*

Proof. First, write $f = g + ih$, where $i = \sqrt{-1}$. Now, without loss of generality, we can assume that f is real and non-negative. Now, let $n = 0, 1, 2, \dots$ and $0 \le k \le 2^{2n} - 1$. Define two families of sets,

$$E_n^k = f^{-1}\left(\left(k2^{-n}, (k+1)2^{-n}\right]\right), \text{ and } F_n = f^{-1}\left((2^n, \infty]\right).$$

Now, set

$$\phi_n = \sum_{k=0}^{n^{2n}-1} k2^{-n}\chi_{E_n^k} + 2^n\chi_{F_n}.$$

The rest follows pretty quickly. □

Theorem 2.12. *[Monotone Convergence Theorem] If $\{f_n\}$ is a sequence of non-negative measurable functions such that $f_j \le f_{j+1}$ for all j, and $f_n \to f$, then*

$$\int f = \lim_{n\to\infty} \int f_n.$$

Proof. This is merely a sketch. First, note that $\{\int f_n\}$ is an increasing sequence of numbers, so its limit exists, even though it may be infinite. Now, we know that $\int f_n \le \int f$ for every n, so this must hold in the limit as well. Thus, we have that LHS $\le$ RHS.

To see the other inequality, let ϕ be a simple function between 0 and f, inclusively, and consider an $\alpha \in (0,1)$. Now, define E_n to be the subset where $f_n(x) \ge \alpha\phi(x)$. We can see that $\int f_n \ge \int_{E_n} f_n \ge \alpha \int_{E_n} \phi$. If we check carefully (i.e. skipping some details here), we will see that $\int_{E_n} \phi \to \int \phi$, so $\lim_{n\to\infty} \int f_n \ge \alpha \int \phi$ for every $\alpha < 1$. This means we can get it for $\alpha = 1$ as well, and we have RHS $\le$ LHS.

□

Lemma 2.13 (Fatou's Lemma)**.** Suppose $\{f_n\}$ is a sequence of measurable, non-negative functions, then

$$\int (\liminf f_n) \le \liminf \int f_n.$$

Proof. Notice first that when $j \le k$,

$$\inf_{n\ge k} f_n \le f_j$$

$$\int \inf_{n\ge k} f_n \le \int f_j \le \inf_{j\ge k} \int f_j.$$

Now, by the monotone convergence theorem, I get

$$\int (\liminf f_n) = \lim_{k\to\infty} \int \inf_{n\ge k} f_n \le \int \liminf \int f_n.$$

□

Definition 2.14. We say that a function, f, is *integrable* iff $\int |f| \leq \infty$.

Theorem 2.15 (Dominated Convergence Theorem)**.** *Let $\{f_n\}$ be a sequence of integrable functions with $f_n \to f$ a.e., and suppose that there exists a non-negative, integrable function g such that $f_n \leq g$ a. e., then f is integrable and*

$$\int f = \lim_{n\to\infty} \int f_n.$$

Proof. We know that f will be measurable (read the details elsewhere...), and since $|f| \leq g$ a.e., we get that f will be integrable. Without loss of generality, suppose that f_n and f are real (why is this okay?) and get that $g + f_n \geq 0$ a.e. and $g - f_n \geq 0$ a.e. Now, we use Fatou ...

$$\int g + \int f \leq \liminf \int (g + f_n) = \int g + \liminf \int f_n, \text{ and}$$

$$\int g - \int f \leq \limsup \int (g - f_n) = \int g - \limsup \int f_n.$$

This gives us that $\liminf \int f_n \geq \int f \geq \limsup \int f_n$, so we are done. □

Definition 2.16. We say that $(X, \mathcal{M}, \mu)$ is *σ-finite* if it can be written as a countable union of measurable, finite-measure subsets.

Notice that this depends on the associated measure. The real numbers equipped with the Lebesgue measure is σ-finite, where the real numbers equipped with the counting measure is not.

Theorem 2.17 (Fubini)**.** *Suppose that $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are σ-finite measure spaces. If f is a non-negative, integrable function on $X \times Y$, then*

$$\int f \ d(\mu \times \nu) = \int \left(\int f(x,y) d\nu(y) \right) d\mu(x) = \int \left(\int f(x,y) d\mu(x) \right) d\nu(y).$$

Definition 2.18. Given a set X, two positive measures, μ and ν are said to be *absolutely singular* if for every set, E, on which μ is positive, ν is zero, and vice versa. We write $\mu \perp \nu$. We say that ν is *absolutely continuous with respect to* μ if for every set, E, on which μ is zero, ν is also zero. We write $\nu \ll \mu$.

Theorem 2.19 (Radon-Nikodym Derivative)**.** *If μ and ν are two positive, σ-finite measures, with $\nu \ll \mu$, then there exists a measurable function f, called the Radon-Nikodym Derivative, such that*

$$\nu(E) = \int_E f d\mu.$$