Contents

1 Lecture 1

The notion of measure is going to generalize the real line's notion of **distance**. Recall that $\mathbb{Q}$ can be constructed from the integers as follows, defining $\mathbb{N}_+ = \{1, 2, 3, \dots\}$ and $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ and we can write $\mathbb{Q} = \{\frac{m}{n} : m \in \mathbb{Z}, n \in \mathbb{N}_+\}$. Recall also that $\mathbb{Q}$ is a countable set.

Recall that $\mathbb{Q}$ is a dense subset of the real line, which we will revisit. First, we define the notion of a **distance** (or **metric**) between two rational numbers, a function $d : \mathbb{Q} \times \mathbb{Q} \to [0, \infty)$:

$$d\left(\frac{m_1}{n_1}, \frac{m_2}{n_2}\right) = \frac{|m_1 n_2 - n_1 m_2|}{n_1 n_2}$$

the distance is also a rational number. However, not all Cauchy sequences in the rationals **converge** to a rational number, the metric space is not complete.

Definition 1.1
A **Cauchy sequence** is a sequence $\{x_n\}$ such that for all $\epsilon > 0$, there exists a threshold $n_0 \in \mathbb{N}$ such that if we have $n, m \geq n_0$, $d(x_n, x_m) < \epsilon$.

We can construct the reals by filling in absences in the rationals. To see these holes, we will represent real numbers with their decimal representation. Now, everyone has a unique representation, except, $0.999\cdots = 1.000\dots$. There are only countably many "awkward" points here (terminating decimals are a subset of the rationals), so it's not a big issue. We will just ban the 9s version, i.e. $1/2 = 0.5000$. If I select values for the decimal places at random, then with probability 0 I get a repeating decimal (rational number). This means the rationals are very slim among the reals.

Let's take π and write it as a Cauchy sequence of rationals $(3, 3.1, 3.14, 3.141, \dots)$. Since π is not rational, we have that the Cauchy sequence doesn't converge to a rational number. This seems to be a way to construct real numbers; why don't we identify π with this Cauchy sequence? But this isn't the only Cauchy sequence that converges to π. Also from infinite series we know that $\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots$; multiplying by 4 and taking partial sums forms another Cauchy sequences. However, defining the relation between two convergent Cauchy sequences (in the real numbers) x and y that $x \sim y$ if x and y have the same limit. One can check this is an equivalence relation. This implies:

Definition 1.2
The set of real numbers, $\mathbb{R}$ is the collection of equivalence classes of Cauchy sequences of rationals where for two sequences x, y, $x \sim y$ if $d(x_n, y_n) \to 0$.

Thus, $\mathbb{R}$ is the **completion** of the rational numbers $\mathbb{Q}$. Let's look at its properties

1. The distance function is as follows. Take $(x_n) \in X$, $(y_n) \in Y$

$$d : \mathbb{R} \times \mathbb{R} \to [0, \infty)$$
$$d(X, Y) = \lim_n d(x_n, y_n)$$

   This is also the distance function for any completion of a metric space (where on the right, the distance function is inherited from the original space). It turns out it doesn't matter what representation we use (we are modding out by everything in the equivalence classes).

2 Lecture 3

2.1 Weak Law of Large Numbers

Last time we discussed the weak law of large numbers, which states that for the experiment of flipping ∞ coins with the ith coin flip given as $B_i = \mathbf{1}\{i\text{th flip is H}\}$, that for all $\epsilon > 0$,

$$\lim_n \mathrm{P}\left(|\frac{1}{n}\sum_{i=1}^n B_i - \frac{1}{2}| > \epsilon \right) = 0$$

This means that looking at the sum of the first n terms, if we take big enough n, the sum of intervals that are bad (bijecting real binary sequences with a sequence of heads and tails) takes up an arbitrarily small portion of the real line.

2.2 Strong Law of Large Numbers

Now let us formulate the SLLN. Consider the sequence of functions $b_1 = \mathbf{1}_{[1/2,1)}$, $b_2 = \mathbf{1}_{[1/4,1/2)} + \mathbf{1}_{[3/4,1)}, \ldots$ such that for $x \in [0,1]$ we have $x = 0.b_1(x)b_2(x)\ldots$.

Definition 2.1
Call the set N of normal numbers is

$$N = \{x \in [0,1] : \lim_n \frac{1}{n}\sum_{i=1}^n b_i(x) = \frac{1}{2}\}$$

The informal strong law is thus if U is picked uniformly at random from the interval $[0,1]$, $\mathbb{P}[U \in N] = 1$, i.e. $m(N) = 1$ for a Lebesgue measure m (not defined yet). For sets that have Lebesgue measure 0 or 1, we can go for a more direct formulation.

Definition 2.2
$A \subseteq [0,1]$ is negligible if for all $\epsilon > 0$, we can find open intervals $O_i \subseteq \mathbb{R}$ for $i \in \mathbb{N}$ such that they form an open cover of A ($A \subseteq \bigcup_{i=1}^\infty O_i$) with $\sum_{i=1}^\infty \ell(O_i) < \epsilon$.

The negligible sets will be those with $m(A) = 0$. We claim $\mathbb{Q} \cap [0,1)$. Then taking an open interval around each rational $\frac{p}{q}$ of ϵe^{-q} suffices. Since the rationals are countable, this makes a countable open cover. In fact, any countable set is negligible; you can just order them as $a_1, a_2, \ldots$ and we can pick $\ell(O_i) = \epsilon/2^i$ to surround a_i.

Theorem 2.3
N^c is negligible.

We can also write, calling $\beta_k(x) = \frac{1}{k}\sum_{i=1}^k b_i(k)$:

$$N = \bigcap_{n=1}^\infty \bigcup_{j=1}^\infty \bigcap_{j}^\infty \{x \in [0,1] : |\beta_k(x) - \frac{1}{2}| < \frac{1}{n}\}$$

quantifiers can instead be replaced by union and interaction. $\bigcap$ is roughly for all, and $\bigcup$ is there exists.

Theorem 2.4
A countable union of negligible sets is negligible.

Proof 2.5
Consider negligible sets N_i. We wish to cover $N = \bigcup_i N_i$. This means we can choose covers for any ϵ we want, in particular,

$$\left\{ O_{n,i} : i \in C_n, N_n \subseteq \ bigcup_{i=1}^{\infty} O_{n,i}, \sum_{i=1}^{\infty} \ell(O_{n,i}) < \frac{\epsilon}{2^n} \right\}$$

But then $N \subseteq \bigcup_n \bigcup_i O_{n,i}$. But also $\sum_n \sum_i \ell(O_{n,i}) < \epsilon$.

But is N negligible? But this would imply $N \cup N^c = [0, 1]$ is negligible. This seems not possble. How do we prove it? Consider a number of intervals $O_i = (a_i, b_i)$ of length $d_i = b_i - a_i$. If $d = \sum d_i < \epsilon$, then clearly it should not be possible to cover the entire interval, just by stacking them on each other. But $\mathbb{Q}$ is dense in $[0, 1)$, yet we could still cover $\mathbb{Q} \subseteq [0, 1]$. The difference is $[0, 1]$ is compact.

3 Lecture 4

3.1 Showing $[0,1]$ is not negligible

Recall that last time we made an incorrect stacking argument (e.g. to cover an interval I with a bunch of open intervals, even if we stack them together we shouldn't be allowed to go further than the length). However, this stacking argument is indeed correct for a finite set of intervals:

Theorem 3.1
Let $\{O_a : a \in A\}$ be a finite collection of open intervals such that for some closed, bounded interval of the real line I, we have $I \subseteq \cup_{a\in A} O_a$. Then $\sum_{a\in A} \ell(O_a) > \ell(I)$.

Proof 3.2
By translating and stretching, we can without loss of generality we can just try to cover the unit interval $[0, 1]$. Let $n = |A|$. We will induct on n.

- If there is only 1 interval (λ, μ), we just need $\lambda < 0$ and $\mu > 1$. Clearly $\mu - \lambda > 1 - 0$, so the base case works.
- Suppose the theorem is true for $|A| = n$. Suppose now you have a collection of $n + 1$ intervals that cover $[0, 1]$. At least one of the intervals contain 1, call it $O_a = (\lambda, \mu)$. Clearly $\mu > 1$, but if $\lambda \le 0$, then $\ell((\lambda, \mu)) > 1$ and so the sum of the lengths of the intervals already exceeds 1. Now, if $0 < \lambda < 1$, then consider the interval $[0,1] \setminus (\lambda, \mu) = [0, \lambda]$. Clearly the rest of the intervals cover this new interval, and there are n of them. By the inductive hypothesis, this means $\sum_{b\in A\setminus\{a\}} |O_b| > \lambda$. Adding in our interval,
$$\sum_{a\in A} |O_a| > \lambda + (\mu - \lambda) = \mu > 1$$
showing our claim.

Now we will prove that [0,1] is not negligible, i.e. there exists no arbitrarily small open covers of $[0, 1]$. In fact, for any cover O, $\sum_{\alpha\in\mathcal{A}} |O_\alpha| > 1$. The reason for this is because by compactness, from O we can extract a finite open subcover $\mathcal{B}$. Thus, $\sum_{\alpha\in\mathcal{A}} |O_\alpha| \ge \sum_{\alpha\in\mathcal{B}} |O_\alpha| > 1$.

Corollary 3.3
[0,1] is uncountable.

This follows pretty fast from the above fact and the fact that any countable set is negligible (by a tight covering argument, similar to the rationals).

3.2 Constructing the Lebesgue Measure

To construct a general object, we often attempt to pick something which fits a few small examples we have. For example, let's take a sequence $(a_n : n \in \mathbb{N})$. To develop a general notion of convergence, we want to be able to assign an extended real number ($\mathbb{R} \cup \{\pm\infty\}$) to each sequence. One way to do that is with
$$\limsup_n a_n = \lim_{n\to\infty} \sup_{m\ge n} a_m = \inf_n \sup_{m\ge n} a_m$$
Since the sequence on the right is decreasing and bounded (or goes to $\pm\infty$ if unbounded), the limit converges. We could also pick
$$\liminf_n a_n = \lim_{n\to\infty} \inf_{m\ge n} a_m = \sup_n \inf_{m\ge n} a_m$$

If we want to define the notion of a limit, we could just say if these two agree, then $\lim_n a_n = \limsup_n a_n = \liminf_n a_n$ and that the limit does not exist otherwise. This extends the theory for monotone sequences to general sequences. What if we want a limit of intervals? Suppose we have a series of sets $A_i \subseteq [0, 1]$, $i \in \mathbb{N}$, we can define and inclusive limit

$$\limsup_n A_n = \{x : x \in A_i \text{ infinitely often}\} = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m$$

or a more restrictive limit

$$\liminf_n A_n = \{x : x \notin A_i \text{ for finitely many } i\} = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m$$

Similarly, we can define $\lim_n A_n = \liminf_n A_n = \limsup_n A_n$ if they share the same value, otherwise it doesn't exist. However, if the difference between the two is negligible we still want to say the limit is well-defined. We probably want to do some more business with equivalence classes, e.g. define $A \sim B \iff A \Delta B$ is negligible. The latter definition fits a bit better with the theory presented in this class.

4 Lecture 5

In Billingsley, he proves for disjoint sets, if $[0,1) = \bigcup_{i=1}^{\infty}[a_i, b_i)$, then $\sum_{i=1}^{\infty} d_i = 1$ where $d_i = b_i - a_i$. This result will be useful for us in the future.

4.1 Measure

For an open set (in universe $[0,1]$) $O = \bigcup O_i$ where the O_i are disjoint, define the measure $\ell(O) = \sum_{i=1}^{\infty} \ell(O_i)$. If the universal set is $[0,1]$, e.g. $(a,1]$ is an open subset (we can only make small movemenets to the left). For a closed set C, $\ell(C) = 1 - \ell([0,1] \setminus C)$. We want to now define the measure of a general set $A \subseteq [0,1]$. For $A \subseteq B$, we want $\lambda(A) \leq \lambda(O)$. We define the outer measure

$$\lambda^*(A) = \inf_{O \text{ open}: A \subseteq O} \lambda(O)$$

similarly, inner measure is:

$$\lambda_*(A) = 1 - \lambda^*(A^c)$$

if we switched the open to a closed set, then $\mathbb{Q} \cap [0,1]$ has outer measure 1 when it really ought to be 0. We can equivalently write:

$$\lambda_*(A) = \sup_{C \text{ closed}: C \subseteq A} \lambda(C)$$

Clearly $\lambda_*(A) \leq \lambda^*(A)$. For a “good” set A, these should be equal.

Definition 4.1
Let $A \subseteq [0,1]$. We say that A is **Lebesgue-measurable** and write $A \in \mathcal{L}$ if $\lambda_*(A) = \lambda^*(A)$.

Definition 4.2
Lebesgue measure is the function $\lambda : \mathcal{L} \to [0,1]$ where $\lambda(A) = \lambda_*(A)$.

Definition 4.3
Let X be a set. An algebra $\mathcal{A}$ is a collection of subsets of X.

1. $X \in \mathcal{A}$
2. if $A \in \mathcal{A}$, then $A \setminus X \in \mathcal{A}$
3. if $A_1, \ldots, A_n \in \mathcal{A}$ for finite n, then $\bigcup_{i=1}^{n} A_i \in \mathcal{A}$.

Note that $\emptyset \in \mathcal{A}$.

Definition 4.4
A σ-algebra $\mathcal{A}$ is an algebra which is closed under countable union, so if we have for $i \in \mathbb{N}$ that $A_i \in \mathcal{A}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

By DeMorgan’s laws, we could replace the unions with intersections. In fact, $\mathcal{L}$ is a sigma algebra of $[0,1]$ (we will show this).

Definition 4.5
A set X equipped with a σ-algebra $\mathcal{A}$ is called a **measurable space**. The sets in $\mathcal{A}$ are called **measurable**. A **measure** $\mu : \mathcal{A} \to [0,\infty]$ following the following two properties:

1. $\mu(\emptyset) = 0$
2. If for $i \in \mathbb{N}$ we have $A_i \in \mathcal{A}$ and the A_i are pairwise disjoint, then $\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$.

Definition 4.6
A set X with a σ-algebra $\mathcal{A}$ and a measure μ is called a **measure space** (or probability space).

To introduce some more vocabulary, if $\mu(X) < \infty$, then it's called a finite measure space. If there exists a countable collection of sets A_i for $i \in \mathbb{N}$ with $\mu(A_i) < \infty$ and $X = \bigcup_{i=1}^{\infty} A_i$, then μ is called a σ-finite measure.

Theorem 4.7
$([0,1], \mathcal{L}, \lambda)$ is a measure space which contains all open intervals in $[0, 1]$. Furthermore, λ is the unique measure which extends ℓ for open intervals.

Here's a proof sketch of this theorem. Call $O = (a, b)$. Then $\lambda_*((a, b)) = \lambda^*((a, b)) = b - a$ by a mix of definition and simple sequence arguments. Closure under complement just follows by definition. Finally, to prove the countable union property. Given a sequence of sets, we can "disjointify" them by working greedily, (the nth set is the leftover from the original nth set that hasn't been claimed yet). To cover A_i, we can cover it with an open set that can be arbitrarily small, e.g. $\lambda(A_i) \leq \lambda^*(O_i) + \frac{\epsilon}{2^i}$. Using these gives us the correct upper bound on λ.

Take finite union of closed sets, we can apprxoimate the union up to arbitrary approximation.

5 Lecture 6

Recall the definition of a measure space. The $\mu(\emptyset) = 0$, then the only thing this excludes is μ being ∞ for every set $A \in \mathcal{A}$. Let's talk about a few examples of σ-algebras.

Example 5.1
Let X be a set which we will detail below. Then the following $\mathcal{A}$ are σ-algebras under X:

1. $\mathcal{A} = \mathcal{P}(X)$.
2. $\mathcal{A} = \{A \subset X :$ either A is countable, or A^c is countable.$\}$
3. $X = (0, 1]$. Define a dyadic interval of rank n as $(\frac{i}{2^n}, \frac{i+1}{2^n}]$, where $0 \leq i \leq 2^n - 1$ and let D_n be the set of such intervals. $\mathcal{A}$ = the arbitrary union of elements of D_n (this is an ordinary algebra).

Lemma 5.2
Let $\{A_\alpha : \alpha \in I\}$ be a collection of σ-algebras. Then $\bigcap_{\alpha \in I} A_\alpha$ is a σ-algebra.

Proof 5.3
If $A \in \bigcap_{\alpha \in I} A_\alpha$, meanis $A \in A_\alpha$ for all $\alpha \in I$, which means that $A^c \in A_\alpha$ for all α, meaning that $A^c \in \bigcap_{\alpha \in I} A_\alpha$. The other axiom is verified equally trivially.

Take $\mathcal{C} \subseteq \mathcal{P}(X)$. The above lemma implies that there exists a smallest σ-algebra that contains $\mathcal{C}$. We shall define

$$\sigma(\mathcal{C}) = \bigcap_{\sigma\text{-algebra}\mathcal{A}:\mathcal{C}\subseteq\mathcal{A}} \mathcal{A}$$

Note that $\sigma(\sigma(\mathcal{C}))$. If $\mathcal{C}_1 \subseteq \mathcal{C}_2$, then $\sigma(\mathcal{C}_1) \subseteq \sigma(\mathcal{C}_2)$. The probability point of view is to call the elements of $\mathcal{A}$ as events. We can then think about points in X as outcomes; it participates in some events in $\mathcal{A}$ and makes them true. Likewise, we can restrict a space to one set in $\mathcal{A}$; then we can treat the new space as a conditional probability. We "revealed" some information.

Suppose X is a topological space (i.e. it has open subsets). Call $\mathcal{G} = \{\text{all open sets}\}$. Then we call $\mathcal{B}(X) := \sigma(\mathcal{G})$ the Borel σ-algebra on X. A Borel set is a

Theorem 5.4
Let $X = \mathbb{R}$. The Borel σ-algebrais generated by each of the following collections.

1. $e_1 = \{(a, b)\}$
2. $e_2 = \{[a, b]\}$
3. $e_3 = \{(a, b]\}$
4. $e_4 = \{(a, \infty)\}$

Proof 5.5
It's clearly true that $\sigma(e_1) \subseteq \mathcal{B}(\mathbb{R})$ because each of the sets in e_1 is open. Furthermore, for any open set $O \subseteq \mathbb{R}$ it can be decomposed into a countable number of open intervals (to see countability, note that it's clear that the lengths of each of these are positive). So $\sigma\{\text{open sets}\} \subseteq \sigma(e_1)$.
Now, let's show $\sigma(e_1) = \sigma(e_2)$. $[a, b]^c = (-\infty, a) \cup (b, \infty)$. We can decompose $\bigcup_{n\in\mathbb{N},n>b}(b, n) = (b, \infty)$. We can do the same thing with $(-\infty, a)$. By the σ-algebraproperties, this means $[a, b] \in \sigma(e_1)$, so $\sigma(e_1) \supseteq \sigma(e_2)$. We could've also written $[a, b] = \bigcap_{n\in\mathbb{N}}(a - \frac{1}{n}, b + \frac{1}{n})$. Finally $(a, b) = \bigcup_{n=1}^{\infty}[a + \frac{1}{n}, b - \frac{1}{n}]$.
$(a, b]$ and (a, ∞) aren't too conceptually different. The only trick is to make $A \setminus B = A \cap B^c$. In fact (a, ∞) where $a \in \mathbb{Q}$ is even good enough, since for $a \in \mathbb{R}$, $(a, \infty) = \bigcup_{n=1}^{\infty}(2^{-n} \lfloor 2^n a \rfloor, \infty)$.

Recall the definition of measure space. X with a σ-algebra$\mathcal{A}$ is a measurable space; adding a function which acts as a measure μ, we get a measure space. If a measure space has total measure ∞, it's an infinite measure space; if it's 1, then it's a probability space. For $\mathcal{A} = \mathcal{P}(X)$ and $\mu(A) = |A|$ then this is the counting measure. Consider the following measure, the Dirac measure:

$$\delta_x(A) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

Recall that under countable union of disjoint sets,

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

Note that we cannot hope for an uncountable measure, as $[0, 1] = \bigcup_{x\in[0,1]}\{x\}$, making $1 = 0+0+\cdots+0$. Uncountable sums don't make much sense! One can see that this is a measure. Another measure would be $\sum_{i\in I} a_i\delta_{x_i}$.

6 Lecture N (10/11)

We continue the proof of the following theorem:

Theorem 6.1
Suppose f_n is Cauchy in measure. $\exists f$ measurable such that $f_n \to f$ in measureand there exists a subsequence $n(j)$ where $f_{n(j)} \to f$ a.e. If $f_n \to g$ in measure then $f = g$ a.e.

Here is where we began the proof. We found a subsequence $n(j)$ with $g_j := f_{n(j)}$ such that the "bad sets" $E_j = \{x : |g_j(x) - g_{j+1}(x)| \geq 2^{-j}\}$ has $\mu(E_j) \leq 2^{-j}$ then $F_j = \bigcup_{k \geq j} E_k$ has $\mu(F_j) \leq 2^{1-j}$. Define $F = \bigcup_{j=1}^{\infty} F_j$. Clearly $\mu(F) = 0$, so we can define the limit as

$$f(x) = \begin{cases} \lim_j g_j(x) & \text{on } F^c \\ 0 & \text{o/w} \end{cases}$$

Then f is measurable and $g_j \to f$ almost everywhere. Then since $\mu(F_k) \to 0$ by definition, $g_j \to f$ in measure. To bring this back to f_n, note

$$\{x : |f_n(x) - f(x)| > \epsilon\} \subseteq \{x : |f_n(x) - g_j(x)| \geq \frac{\epsilon}{2}\} \cup \{x : |g_j(x) - f(x)| \geq \frac{\epsilon}{2}\}$$

Since f_n is Cauchy in measure, the measure of the first set goes to 0. The second set is 0 in measure because $g_j \to f$ in measure.

Finally, suppose $f_n \to g$ in measure. $\{x : |f(x) - g(x)| \geq \epsilon\} \subseteq \{x : |f_n(x) - f(x)| \geq \frac{\epsilon}{2}\} \cup \{x : |f_n(x) - g(x)| > \frac{\epsilon}{2}\}$ Since this is true for every n, we can show that $\mu(\{|f(x) - g(x)| \geq \epsilon\}) = 0$ for any $\epsilon > 0$, thus $\mu(\{|f(x) - g(x)| > 0\}) = 0$ and so $f = g$ almost everywhere.

Theorem 6.2
If $f_n \to f$ a.e. and $\mu(X) < \infty$, then $f_n \to f$ in measure.

Call $E_n(\epsilon) = \{x : |f_n(x) - f(x)| > \epsilon\}$ and $F_n(\epsilon) = \bigcup_{m=n}^{\infty} E_m(\epsilon)$. Since $f_n \to f$ almost everywhere, for $F = \bigcap_{n=1}^{\infty} F_n(\epsilon)$, $\mu(F) = 0$. Since $\mu(X) < \infty$ (to make sure measure limits work), this means $\mu(F_n(\epsilon)) \to 0$ as $n \to \infty$. Now, $E_n(\epsilon) \subset F_n(\epsilon)$ so $0 \leq \mu(E_n(\epsilon)) \leq \mu(F_n(\epsilon)) \to 0$, so we've shown $f_n \to f$ in measure.

Similarly, if $f_n \to f$ in L^1 then $f_n \to f$ in measure then f_n is Cauchy in measure, so $f_{n(j)} \to f$ a.e..

6.1 Egorov's Theorem

Theorem 6.3
Suppose $\mu(X) < \infty$ and let $f_n : X \to \mathbb{R}$ and $f_n \to f$ a.e. Then for all $\epsilon > 0$, there exist an "error set" R with $\mu(E) < \epsilon$ then $f_n \to f$ uniformly on E^c.

Without loss of generality, $f_n \to f$ pointwise (we have issues on only a measure zero set, so we can union it with E at the end). For $k, n \in \mathbb{N}$ consider $E_n(k) = \bigcup_{m=n}^{\infty} \{x \in X : |f_m(x) - f(x)| \geq \frac{1}{k}\}$. as $n \to \infty$, $E_n(k)$ decrease to the empty set. Then since $\mu(X) < \infty$, $\mu(E_n(k)) \to 0$ as $n \to \infty$. There exists n_k such that $\mu(E_{n_k}(k)) \leq \frac{\epsilon}{2^k}$. Call $E = \bigcup_{k=1}^{\infty} E_{n_k}(k)$, means $\mu(E) < \epsilon$. Now for $x \notin E$, then for all $k \geq 1$, for all $n \geq n_k$, then $|f_n(x) - f(x)| < \frac{1}{k}$. So $f_n \to f$ uniformly for all $x \notin E$, so we're finished.

6.2 Littlewood's three principles

Three principles from 1944 by Littlewood give us some intuition about real analysis.

1. Every measurable set is ALMOST a finite union of open intervals. (Inner and outer measures coincide)
2. Every integrable/measurable function is ALMOST continuous. (Lusin, L^1 Approximation)
3. Every convergent sequence of functions is ALMOST uniformly convergent. (Egorov)

All of these ALMOSTs mean throw away a measure-0 set.