

# 1 Outer measure and Lebesgue measure cont.

## 1.1 Properties of Lebesgue measure

This collection is a section of facts about Lebesgue measure and the set of Lebesgue measurable sets. We start with looking at  $\mathcal{M}$  the  $\sigma$ -algebra of Lebesgue measurable sets.

**Lemma 1.1** (Null sets are all Lebesgue measurable). *If  $A$  in  $\mathcal{P}(\mathbb{R})$  and  $\lambda^*(A) = 0$  then  $A \in \mathcal{M}$ .*

*Proof.* This is on the assignment. □

We can actually characterise all Lebesgue measurable sets in terms of Null sets and Borel sets both of which we have shown are measurable. We have the following proposition which we won't prove in the course. We will show a similar result in the optional exercises.

**Proposition 1.2.** *A set  $S \subseteq \mathbb{R}$  is Lebesgue measurable if and only there exists a Borel set  $B$  and a null set  $N$  such that  $S = B \Delta N$ .*

The most important thing we want to prove about  $\mathcal{M}$  is that there exists a non-Lebesgue measurable set. Before we do this we need to explore a few more properties of Lebesgue measure itself.

**Proposition 1.3.** *Lebesgue measure is regular that is to say*

- $\lambda(A) = \inf\{\lambda(U) : U \text{ is open, } A \subseteq U\},$
- $\lambda(A) = \sup\{\lambda(K) : K \text{ is compact, } K \subseteq A\}.$

*Proof.* By monotonicity we can see that  $\lambda(A) \leq \inf\{\lambda(U) : U \text{ is open, } A \subseteq U\}$ . Furthermore we can find a sequence of half open rectangles  $R_k$  such that  $A \subseteq \bigcup_n R_n$  and  $\sum_n \lambda(R_n) \leq \lambda(A) + \epsilon$ . By slightly enlarging each of the half open rectangles we can produce another sequence of fully open rectangles  $\tilde{R}_n$  such that  $A \subseteq \bigcup_n \tilde{R}_n$  and  $\lambda(A) \geq \sum_n \lambda(\tilde{R}_n) - 2\epsilon$ . The set,  $\bigcup_n \tilde{R}_n$  is open and  $\epsilon$  can be made arbitrarily small so this shows  $\lambda(A) \geq \inf\{\lambda(U) : U \text{ is open, } A \subseteq U\}$ .

Monotonicity shows that  $\lambda(A) \geq \sup\{\lambda(K) : K \text{ is compact, } K \subseteq A\}$ . First let us assume that  $A$  is contained in some ball,  $B$  around 0. Now use the first part to find some open set  $U$  such that  $B \setminus A \subseteq U$  and  $\lambda(U) \leq \lambda(B \setminus A) + \epsilon$ . Now let  $K = B \setminus U$  then we have  $K \subseteq A \subseteq B$  and  $\lambda(K) \geq \lambda(B) - \lambda(U) \geq \lambda(B) - \lambda(B \setminus A) - \epsilon = \lambda(A) - \epsilon$  (here we use the fact that  $B, A, U, K$  will all have finite measure as they are inside  $B$ ). As  $\epsilon$  is arbitrary this concludes the proof when  $A$  is contained in a ball.

Now suppose that  $A$  is unbounded. Then let  $A_n = A \cap B_n$  where  $B_n$  is the closed ball of radius  $n$ . We have that  $\lambda(A_n) \rightarrow \lambda(A)$ . If  $\lambda(A) = \infty$  then we can find  $K_n \subseteq A_n$  with  $\lambda(K_n)$  arbitrarily close to  $\lambda(A_n)$  therefore we can find such a sequence with  $\lambda(K_n) \rightarrow \infty$ . If  $\lambda(A) \neq \infty$  then, given  $\epsilon$ , there exists  $N$  such that  $\lambda(A_n) \geq \lambda(A) - \epsilon$  for  $n \geq N$ . Then we can find  $K_N \subseteq A_N$  such that  $\lambda(K_N) \geq \lambda(A_N) - \epsilon$  therefore  $\lambda(K_N) \geq \lambda(A) - 2\epsilon$ . This shows we can a compact set which is contained in  $A$ , with measure arbitrarily close to that of  $A$ . □

We now want to show that Lebesgue measure is the only which assigns each interval the correct measure.

**Proposition 1.4.** *Lebesgue measure is the only measure on  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  which assigns each half open interval its length. This is equally true with half open hyper-rectangles in  $\mathbb{R}^d$ .*

*Proof.* The collection of half open intervals is a  $\pi$ -system which generates the Borel  $\sigma$ -algebra. Therefore, if  $\lambda(\mathbb{R})$  had been finite we could use Dynkin's uniqueness of extension Lemma to get that any other measure which agrees with Lebesgue measure on the half open intervals must agree with Lebesgue

measure on the whole of the Borel  $\sigma$ -algebra. Instead let  $E_n = [-n, n]^d$  then, by Dynkin's uniqueness of extension lemma we have that  $\lambda$  is the only measure on  $E_n$ , assigning each rectangle inside  $E_n$  its measure. Since every rectangle is bounded, eventually it is inside some  $E_n$  so if  $\mu$  is a measure such that  $\mu(R) = \lambda(R)$  for every rectangle then the restriction of  $\mu$  to  $E_n$  must agree with the restriction of  $\lambda$  to  $E_n$ . We also have that, for any  $A$ ,  $\mu(A) = \lim_n \mu(A \cap E_n)$  by continuity of measure. So  $\mu(A) = \lim_n \mu(A \cap E_n) = \lim_n \lambda(A \cap E_n) = \lambda(A)$ .  $\square$

**Corollary 1.5.** *Lebesgue measure is translation invariant. That is to say if we define the set  $x + A = \{x + y, y \in A\}$  then  $\lambda(x + A) = \lambda(A)$*

*Proof.* Define a new measure  $\lambda_x$  by  $\lambda_x(A) = \lambda(x + A)$  then  $\lambda_x((a, b]) = \lambda((a + x, b + x]) = b + x - (a + x) = b - a$ . Therefore  $\lambda_x$  agrees with  $\lambda$  on the half open intervals and therefore agrees with  $\lambda$  on the whole of  $\mathcal{B}(\mathbb{R})$ . Again it is straightforward to extend this to  $\mathbb{R}^d$ .  $\square$

Lastly, in the construction of Lebesgue measure we show that  $\mathcal{M}$  is not the whole of  $\mathcal{P}(\mathbb{R})$  and that there exist non-Lebesgue measurable sets.

**Proposition 1.6.** *There exists sets that are in  $\mathcal{P}(\mathbb{R})$  which are not in  $\mathcal{M}$ .*

*Proof.* This proof involves the use of the axiom of choice. In fact it is known that it is necessary to use some form of the axiom of choice to prove the existence of a non-Lebesgue measurable set in  $\mathbb{R}$ .

We use an argument by contradiction, we begin by assuming every subset of  $\mathbb{R}$  is Lebesgue measurable. We define an equivalence relation on  $[0, 1)$  by saying  $x \sim y$  exactly when  $x - y \in \mathbb{Q}$ . Using the axiom of choice we find a subset  $S$  of  $[0, 1)$  which contains exactly one representative of each equivalence class. Next we define the set  $S + q = \{s + q \pmod{1} : s \in S\}$  for each  $q \in \mathbb{Q} \cap [0, 1)$ . Then by our choice of  $S$  we have that

$$[0, 1) = \bigcup_{q \in \mathbb{Q} \cap [0, 1)} (S + q),$$

where this union is disjoint. We can also see by translation invariance of  $\lambda$  that if  $S$  were Lebesgue measurable then we would have

$$\lambda(S) = \lambda(S + q)$$

for every  $q$ . Therefore, by countable additivity we would have

$$\lambda([0, 1)) = \sum_{q \in \mathbb{Q} \cap [0, 1)} \lambda(S + q) = \sum_{q \in \mathbb{Q} \cap [0, 1)} \lambda(S) = \infty.$$

$\square$

## 2 Measurable Functions

A big part of measure theory is the study of functions which are compatible with the measure spaces. We begin with a basic definition which will be satisfied by all the functions we are interested in.

**Definition 2.1** (Measurable functions). If  $(E, \mathcal{E})$  and  $(F, \mathcal{F})$  are two measurable spaces and  $f$  is a function  $E \rightarrow F$ , then we say  $f$  is *measurable* if for every  $A \in \mathcal{F}$  we have  $f^{-1}(A) \in \mathcal{E}$ .

**Lemma 2.2.** *Suppose that  $\mathcal{A} \subset \mathcal{F}$  is such that  $\sigma(\mathcal{A}) = \mathcal{F}$ . If  $f$  is a function such that for every  $A \in \mathcal{A}$  we have  $f^{-1}(A) \in \mathcal{E}$  then  $f$  is measurable.*

*Proof.* First we note that

$$f^{-1}\left(\bigcup_i A_i\right) = \bigcup_i f^{-1}(A_i),$$

and

$$f^{-1}(B \setminus A) = f^{-1}(B) \setminus f^{-1}(A).$$

Now if we consider  $\{A \in \mathcal{F} : f^{-1}(A) \in \mathcal{E}\}$  then this is a  $\sigma$ -algebra, as  $\mathcal{E}$  is a  $\sigma$ -algebra and  $f^{-1}$  preserves set operations. Therefore,  $\{A \in \mathcal{F} : f^{-1}(A) \in \mathcal{E}\}$  is a  $\sigma$ -algebra containing  $\mathcal{A}$  therefore  $\mathcal{F} \subseteq \{A \in \mathcal{F} : f^{-1}(A) \in \mathcal{F}\}$  so  $f$  is measurable.  $\square$

*Remark 2.3.* In particular note that the above lemma means that whenever we have  $f : E \rightarrow \mathbb{R}$  and  $\mathbb{R}$  is equipped with the Borel  $\sigma$  algebra, we know that  $f$  is measurable if  $f^{-1}((-\infty, b])$  is a measurable set for every  $b$ .

**Lemma 2.4.** *If  $E, F$  are topological spaces, equipped with their Borel  $\sigma$ -algebras, and we have  $f : E \rightarrow F$  is a continuous map then  $f$  is measurable.*

*Proof.* This is on the exercise sheet.  $\square$

**Lemma 2.5.** *If  $f_n$  is a sequence of measurable function taking values in  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  then the following functions are all measurable:*

- $-f_1$
- $\lambda f_1$  for  $\lambda > 0$  a fixed constant.
- $f_1 \wedge f_2$
- $f_1 \vee f_2$
- $f_1 + f_2$ ,
- $f_1 f_2$ ,
- $\sup_n f_n$ ,
- $\inf_n f_n$ ,
- $\limsup_n f_n$ ,
- $\liminf_n f_n$ .

*Proof.* We only show two result. The rest are on the assignment.

In order to show that any of these functions are measurable we want to look at  $f^{-1}((-\infty, b])$  or a similar set.  $(f_1 \vee f_2)^{-1}((-\infty, b]) = \{x : \max\{f_1(x), f_2(x)\} \leq b\} = \{x : f_1(x) \leq b \text{ and } f_2(x) \leq b\} = \{x : f_1(x) \leq b\} \cap \{x : f_2(x) \leq b\} = f_1^{-1}((-\infty, b]) \cap f_2^{-1}((-\infty, b])$ . Now since  $f_1$  and  $f_2$  are both measurable the sets  $f_1^{-1}((-\infty, b])$  and  $f_2^{-1}((-\infty, b])$  are both measurable. We also know that the intersection of two measurable sets is measurable.

$(f_1 + f_2)^{-1}((b, \infty)) = \{x : f_1(x) + f_2(x) > b\}$ . Now if  $f_1(x) > b - f_2(x)$  then there exists a  $q \in \mathbb{Q}$  such that  $f_1(x) > q > b - f_2(x)$ . Let us define the set  $A = \bigcup_{q \in \mathbb{Q}} \{x : f_1(x) > q\} \cap \{x : f_2(x) > b - q\}$ . Since  $f_1, f_2$  are both measurable  $A$  is a countable union of measurable sets so measurable. We can also see that if  $x \in A$  then  $f_1(x) + f_2(x) > b$  and our observation shows that in fact  $A = \{x : f_1(x) + f_2(x) > b\}$ . Therefore,  $f_1 + f_2$  is measurable.  $\square$

**Definition 2.6** (Image measure). We can use a measurable function  $f$  to define an image measure. Suppose  $\mu$  is a measure on  $(E, \mathcal{E})$  and  $f$  is a measurable function  $(E, \mathcal{E}) \rightarrow (F, \mathcal{F})$  then we can define a new measure  $\nu$  by saying that

$$\nu(A) = \mu(f^{-1}(A)),$$

for every  $A \in \mathcal{F}$ . We write  $\nu = \mu \circ f^{-1}$ .

We can use the notion of image measure to construct further measures from Lebesgue measure.

**Lemma 2.7.** Suppose  $g : \mathbb{R} \rightarrow \mathbb{R}$  and that  $g$  is non-constant, right-continuous and non-decreasing. Let us define  $g(-\infty) = \lim_{x \rightarrow -\infty} g(x)$  and  $g(\infty) = \lim_{x \rightarrow \infty} g(x)$  and let us call the interval  $I := (g(-\infty), g(\infty))$  (this might be the whole of  $\mathbb{R}$ ). Define a partial inverse to  $g$  by  $f : I \rightarrow \mathbb{R}$  by

$$f(x) = \inf_y \{x \leq g(y)\}.$$

Then  $f$  is left-continuous and non-decreasing and  $f(x) \leq y$  if and only if  $x \leq g(y)$ .

*Proof.* Fix  $x \in I$  and consider the set  $J_x = \{y \in \mathbb{R} : x \leq g(y)\}$  by definition of  $I$  we know that  $J_x$  is non empty and is not the whole of  $\mathbb{R}$  (this shows that  $f$  is well defined). As  $g$  is non-decreasing, if  $y \in J_x$  and  $y' \geq y$ , then  $y' \in J_x$ . As  $g$  is right-continuous, if  $y_n \in J_x$  and  $y_n \downarrow y$  then  $y \in J_x$  (noting the  $\leq$  sign in  $J_x$ ). Now using this we have that if  $x \leq x'$  then  $J_x \supseteq J_{x'}$  so  $f(x) \leq f(x')$ . We also have that if  $x_n \uparrow x$  then  $J_x = \bigcap_n J_{x_n}$ , so  $f(x_n) \rightarrow f(x)$ .  $\square$

**Theorem 2.8.** Let  $g$  be a non-constant, right-continuous and non-decreasing function from  $\mathbb{R} \rightarrow \mathbb{R}$ . There exists a unique Radon measure on  $\mathbb{R}$  such that for all  $a, b \in \mathbb{R}$  with  $a < b$

$$dg((a, b]) = g(b) - g(a).$$

We call this measure the Lebesgue Steitjles measure associated with  $g$ . Furthermore, every Radon measure on  $\mathbb{R}$  can be represented this way.

*Proof.* Define  $I$  and  $f$  as in the Lemma above. Then we can construct  $dg$  as the image measure of Lebesgue measure on  $I$ . That is to say we can let  $dg = \lambda \circ f^{-1}$ . If this is the case then

$$dg((a, b]) = \lambda(\{x : f(x) > a, f(x) \leq b\}) = \lambda((g(a), g(b)]) = g(b) - g(a).$$

The standard argument for uniqueness of measures (as for that of Lebesgue measure) gives uniqueness of this measure.

Finally, if  $\nu$  is a Radon measure on  $\mathbb{R}$  then we can define a function  $g$ , by

$$g(y) = \nu((0, y]), \quad y \geq 0, \quad g(y) = -\nu((y, 0]), \quad y < 0.$$

Then  $\nu = dg$  by uniqueness.  $\square$

## 2.1 Random variables and the measure theoretic formulation of probability - in brief

The structure of measure theory allows us to put probability theory on a firm footing.

**Definition 2.9.** We call a measure space  $(\Omega, \mathcal{F}, \mathbb{P})$  a *probability space* (and use different letters for the different bits) if  $\mathbb{P}(\Omega) = 1$ . In this setting we still have  $\Omega$  is a set,  $\mathcal{F}$  is a  $\sigma$ -algebra and  $\mathbb{P}$  is a measure. We call  $\mathbb{P}$  a probability measure.

The set  $\Omega$  is not really described and we view it as the *space of all possible events*. We call  $A \subset \Omega$  and event, and  $\mathbb{P}(A)$  the probability of an event happening.

**Definition 2.10.** A *random variable*,  $X$  is a measurable function from a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  to another measurable space  $(E, \mathcal{A})$ .

Under this way of writing things we have  $\mathbb{P}(X \in B) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in B\}) = \mathbb{P}(X^{-1}(B))$ , where  $X^{-1}(B) \in \mathcal{F}$  as  $X$  is measurable. We call  $X^{-1}(B)$  the event that  $X \in B$ . When working with probability people usually suppress the argument  $\omega$ .

**Definition 2.11.** The *law of a random variable*,  $X$  is the image measure of  $\mathbb{P}$  under the measurable function  $X$ . I.e. if  $X : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (E, \mathcal{A})$  then we define a measure  $\mu_X$  on  $(E, \mathcal{A})$  by  $\mu_X(B) = \mathbb{P}(X \in B)$ .

*Remark 2.12.* The law of a random variable is an object which allows us to understand both probability densities and discrete probability distributions in the same way.

If  $X$  takes values in  $\mathbb{R}$  then the distribution function of  $X$ ,  $F_X(x) = \mu_X((-\infty, x])$ . If  $X$  has density  $f(x)$  then  $\mu_X$  is equal to the measure given by  $\mu_X(A) = \int_A f(x)dx$ , though we still don't know how to integrate over complicated sets.