Lecture 7: Borel Sets and Lebesgue Measure

Lecturer: Dr. Krishna Jagannathan *Scribes: Ravi Kolla, Aseem Sharma, Vishakh Hegde*

In this lecture, we discuss the case where the sample space is uncountable. This case is more involved than the case of a countable sample space, mainly because it is often not possible to assign probabilities to all subsets of Ω. Instead, we are forced to work with a smaller σ-algebra. We consider assigning a "uniform probability measure" on the unit interval.

7.1 Uncountable sample spaces

Consider the experiment of picking a real number at random from $\Omega = [0,1]$, such that every number is "equally likely" to be picked. It is quite apparent that a simple strategy of assigning probabilities to singleton subsets of the sample space gets into difficulties quite quickly. Indeed,

(i) If we assign some positive probability to each elementary outcome, then the probability of an event with infinitely many elements, such as $A = \{1, \frac{1}{2}, \frac{1}{3}, \cdots\}$, would become unbounded.

(ii) If we assign zero probability to each elementary outcome, this alone would not be sufficient to determine the probability of a uncountable subset of Ω, such as $\left[\frac{1}{2}, \frac{2}{3}\right]$. This is because probability measures are not additive over uncountable disjoint unions (of singletons in this case).

Thus, we need a different approach to assign probabilities when the sample space is uncountable, such as $\Omega = [0,1]$. In particular, we need to assign probabilities directly to specific subsets of Ω. Intuitively, we would like our 'uniform measure' μ on $[0,1]$ to possess the following two properties.

(i) $\mu((a,b)) = \mu((a,b]) = \mu([a,b)) = \mu([a,b])$

(ii) Translational Invariance. That is, if $A \in [0,1]$, then for any $x \in \Omega$, $\mu(A \oplus x) = \mu(A)$ where, the set $A \oplus x$ is defined as

$$A \oplus x = \{a + x | a \in A, a + x \leq 1\} \cup \{a + x - 1 | a \in A, a + x > 1\}$$

However, the following impossibility result asserts that there is no way to consistently define a uniform measure on all subsets of $[0,1]$.

Theorem 7.1 (Impossibility Result) *There does not exist a definition of a measure $\mu(A)$ for all subsets of $[0,1]$ satisfying (i) and (ii).*

Proof: Refer proposition 1.2.6 in [1].

Therefore, we must compromise, and consider a smaller σ-algebra that contains certain "nice" subsets of the sample space $[0,1]$. These "nice" subsets are the intervals, and the resulting σ-algebra is called the Borel σ-algebra. Before defining Borel sets, we introduce the concept of generating σ-algebras from a given collection of subsets.

7.2 Generated σ-algebra and Borel sets

The σ-algebra generated by a collection of subsets of the sample space is the smallest σ-algebra that contains the collection. More formally, we have the following theorem.

Theorem 7.2 *Let $\mathcal{C}$ be an arbitrary collection of subsets of Ω, then there exists a smallest σ-algebra, denoted by $\sigma(\mathcal{C})$, that contains all elements of $\mathcal{C}$. That is, if $\mathcal{H}$ is any σ-algebra such that $\mathcal{C} \subseteq \mathcal{H}$, then $\sigma(\mathcal{C}) \subseteq \mathcal{H}$. $\sigma(\mathcal{C})$ is called the σ-algebra generated by $\mathcal{C}$.*

Proof: Let $\{\mathcal{F}_i,\ i \in \mathcal{I}\}$ denote the collection of all σ-algebras that contain $\mathcal{C}$. Clearly, the collection $\{\mathcal{F}_i,\ i \in \mathcal{I}\}$ is non-empty, since it contains at least the power set, 2^Ω. Consider the intersection $\bigcap_{i\in\mathcal{I}} \mathcal{F}_i$. Since the intersection of σ-algebras results in a σ-algebra (homework problem!) and the intersection contains $\mathcal{C}$, it follows that $\bigcap_{i\in\mathcal{I}} \mathcal{F}_i$ is a σ-algebra that contains $\mathcal{C}$. Finally, if $\mathcal{C} \subseteq \mathcal{H}$, then $\mathcal{H}$ is one of $\mathcal{F}_i$'s for some $i \in \mathcal{I}$. Hence $\bigcap_{i\in\mathcal{I}} \mathcal{F}_i$ is the smallest σ-algebra generated by $\mathcal{C}$. ■

Intuitively, we can think of $\mathcal{C}$ as being the collection of subsets of Ω which are of interest to us. Then, $\sigma(\mathcal{C})$ is the smallest σ-algebra containing all the 'interesting' subsets.

We are now ready to define Borel sets.

Definition 7.3

(a) Consider $\Omega = (0, 1]$. Let $\mathcal{C}_0$ be the collection of all open intervals in $(0, 1]$. Then $\sigma(\mathcal{C}_0)$, the σ - algebra generated by $\mathcal{C}_0$, is called the Borel σ - algebra. It is denoted by $\mathcal{B}((0, 1])$.

(b) An element of $\mathcal{B}((0, 1])$ is called a Borel-measurable set, or simply a Borel set.

Thus, every open interval in $(0, 1]$ is a Borel set. We next prove that every singleton set in $(0, 1]$ is a Borel set.

Lemma 7.4 *Every singleton set $\{b\},\ 0 < b \leq 1$, is a Borel set, i.e., $\{b\} \in \mathcal{B}((0, 1])$.*

Proof: Consider the collection of sets set $\left\{\left(b - \frac{1}{n}, b + \frac{1}{n}\right), n \geq 1\right\}$. By the definition of Borel sets,

$$\left(b - \frac{1}{n}, b + \frac{1}{n}\right) \in \mathcal{B}((0, 1]).$$

Using the properties of σ-algebra,

$$\begin{aligned}
&\left(b - \frac{1}{n}, b + \frac{1}{n}\right)^c \in \mathcal{B}((0, 1]) \\
\implies &\bigcup_{n=1}^{\infty} \left(b - \frac{1}{n}, b + \frac{1}{n}\right)^c \in \mathcal{B}((0, 1]) \\
\implies &\left(\bigcap_{n=1}^{\infty} \left(b - \frac{1}{n}, b + \frac{1}{n}\right)\right)^c \in \mathcal{B}((0, 1]) \\
\implies &\bigcap_{n=1}^{\infty} \left(b - \frac{1}{n}, b + \frac{1}{n}\right) \in \mathcal{B}((0, 1]). \qquad (7.1)
\end{aligned}$$

Next, we claim that

$$\{b\} = \bigcap_{n=1}^{\infty} \left(b - \frac{1}{n}, b + \frac{1}{n}\right). \tag{7.2}$$

i.e., b is the only element in $\bigcap_{n=1}^{\infty} \left(b - \frac{1}{n}, b + \frac{1}{n}\right)$. We prove this by contradiction. Let h be an element in $\bigcap_{n=1}^{\infty} \left(b - \frac{1}{n}, b + \frac{1}{n}\right)$ other than b. For every such h, there exists a large enough n_0 such that $h \notin \left(b - \frac{1}{n_0}, b + \frac{1}{n_0}\right)$. This implies $h \notin \bigcap_{n=1}^{\infty} \left(b - \frac{1}{n}, b + \frac{1}{n}\right)$. Using (7.1) and (7.2), thus, proves that $\{b\} \in \mathcal{B}((0,1])$. ■

As an immediate consequence to this lemma, we see that every half open interval, $(a,b]$, is a Borel set. This follows from the fact that

$$(a,b] = (a,b) \cup \{b\},$$

and the fact that a countable union of Borel sets is a Borel set. For the same reason, every closed interval, $[a,b]$, is a Borel set.

Note: Arbitrary union of open sets is always an open set, but infinite intersections of open sets need not be open.

Further reading for the enthusiastic: (try Wikipedia for a start)

- Non-Borel sets
- Non-measurable sets (Vitali set)
- Banach-Tarski paradox (a bizzare phenomenon about cutting up the surface of a sphere. See `https://www.youtube.com/watch?v=Tk4ubu7BlSk`
- The cardinality of the Borel σ-algebra (on the unit interval) is the same as the cardinality of the reals. Thus, the Borel σ-algebra is a much 'smaller' collection than the power set $2^{[0,1]}$. See `https://math.dartmouth.edu/archive/m103f08/public_html/borel-sets-soln.pdf`

7.3 Caratheodory's Extension Theorem

In this section, we discuss a formal procedure to define a probability measure on a general measurable space $(\Omega, \mathcal{F})$. Specifying the probability measure for all the elements of $\mathcal{F}$ directly is difficult, so we start with a smaller collection $\mathcal{F}_0$ of 'interesting' subsets of Ω, which need not be a σ-algebra. We should take $\mathcal{F}_0$ to be rich enough, so that the σ-algebra it generates is same as $\mathcal{F}$. Then we define a function $\mathbb{P}_0 : \mathcal{F}_0 \to [0,1]$, such that it corresponds to the probabilities we would like to assign to the interesting subsets in $\mathcal{F}_0$. Under certain conditions, this function $\mathbb{P}_0$ can be extended to a legitimate probability measure on $(\Omega, \mathcal{F})$ by using the following fundamental theorem from measure theory.

Theorem 7.5 (Caratheodory's extension theorem) *Let $\mathcal{F}_0$ be an algebra of subsets of Ω, and let $\mathcal{F} = \sigma(\mathcal{F}_0)$ be the σ-algebra that it generates. Suppose that $\mathbb{P}_0$ is a mapping from $\mathcal{F}_0$ to $[0,1]$ that satisfies $\mathbb{P}_0(\Omega) = 1$, as well as countable additivity on $\mathcal{F}_0$.*
Then, $\mathbb{P}_0$ can be extended uniquely to a probability measure on $(\Omega, \mathcal{F})$. That is, there exists a unique probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ such that $\mathbb{P}(A) = \mathbb{P}_0(A)$ for all $A \in \mathcal{F}_0$.

Proof: Refer Appendix A of [2]. ■

We use this theorem to define a uniform measure on $(0,1]$, which is also called the Lebesgue measure.

7.4 The Lebesgue measure

Consider $\Omega = (0,1]$. Let $\mathcal{F}_0$ consist of the empty set and all sets that are finite unions of the intervals of the form $(a,b]$. A typical element of this set is of the form

$$F = (a_1, b_1] \cup (a_2, b_2] \cup \ldots \cup (a_n, b_n]$$

where, $0 \le a_1 < b_1 \le a_2 < b_2 \le \ldots \le a_n < b_n$ and $n \in \mathbb{N}$.

Lemma 7.6

a) $\mathcal{F}_0$ is an algebra

b) $\mathcal{F}_0$ is not a σ-algebra

c) $\sigma(\mathcal{F}_0) = \mathcal{B}$

Proof:

a) By definition, $\Phi \in \mathcal{F}_0$. Also, $\Phi^C = (0,1] \in \mathcal{F}_0$. The complement of $(a_1,b_1] \cup (a_2,b_2]$ is $(0,a_1] \cup (b_1,a_2] \cup (b_2,1]$, which also belongs to $\mathcal{F}_0$. Furthermore, the union of finitely many sets each of which are finite unions of the intervals of the form $(a,b]$, is also a set which is the union of finite number of intervals, and thus belongs to $\mathcal{F}_0$.

b) To see this, note that $\left(0, \frac{n}{n+1}\right] \in \mathcal{F}_0$ for every n, but $\bigcup\limits_{n=1}^{\infty} \left(0, \frac{n}{n+1}\right] = (0,1) \notin \mathcal{F}_0$.

c) First, the null set is clearly a Borel set. Next, we have already seen that every interval of the form $(a,b]$ is a Borel set. Hence, every element of $\mathcal{F}_0$ (other than the null set), which is a finite union of such intervals, is also a Borel set. Therefore, $\mathcal{F}_0 \subseteq \mathcal{B}$. This implies $\sigma(\mathcal{F}_0) \subseteq \mathcal{B}$.

Next we show that $\mathcal{B} \subseteq \sigma(\mathcal{F}_0)$. For any interval of the form (a,b) in $\mathcal{C}_0$, we can write $(a,b) = \bigcup\limits_{n=1}^{\infty} \left(\left(a, b - \frac{1}{n}\right] \cap \Omega\right)$. Since every interval of the form $\left(a, b - \frac{1}{n}\right] \in \mathcal{F}_0$, a countable number of unions of such intervals belongs to $\sigma(\mathcal{F}_0)$. Therefore, $(a,b) \in \sigma(\mathcal{F}_0)$ and consequently, $\mathcal{C}_0 \subseteq \sigma(\mathcal{F}_0)$. This gives $\sigma(\mathcal{C}_0) \subseteq \sigma(\mathcal{F}_0)$. Using the fact that $\sigma(\mathcal{C}_0) = \mathcal{B}$ proves the required result.

■

For every $F \in \mathcal{F}_0$ of the form

$$F = (a_1, b_1] \cup (a_2, b_2] \cup \ldots \cup (a_n, b_n],$$

we define a function $\mathbb{P}_0 : \mathcal{F}_0 \to [0,1]$ such that

$$\mathbb{P}_0(\Phi) = 0 \text{ and } \mathbb{P}_0(F) = \sum_{i=1}^{n} (b_i - a_i).$$

Note that $\mathbb{P}_0(\Omega) = \mathbb{P}_0((0,1]) = 1$. Also, if $(a_1, b_1], (a_2, b_2], \ldots, (a_n, b_n]$ are disjoint sets, then

$$\mathbb{P}_0\left(\bigcup_{i=1}^{n} ((a_i, b_i])\right) = \sum_{i=1}^{n} \mathbb{P}_0((a_i, b_i]) = \sum_{i=1}^{n} (b_i - a_i)$$

implying finite additivity of $\mathbb{P}_0$. It turns out that $\mathbb{P}_0$ is countably additive on $\mathcal{F}_0$ as well i.e., if $(a_1, b_1], (a_2, b_2], \ldots$ are disjoint sets such that $\bigcup_{i=1}^{\infty} ((a_i, b_i]) \in \mathcal{F}_0$, then $\mathbb{P}_0\left(\bigcup_{i=1}^{\infty} ((a_i, b_i])\right) = \sum_{i=1}^{\infty} \mathbb{P}_0((a_i, b_i]) = \sum_{i=1}^{\infty} (b_i - a_i)$. The proof is non-trivial and beyond the scope of this course (see [Williams] for a proof). Thus, in view of Theorem 7.5, there exists a unique probability measure $\mathbb{P}$ on $((0,1], \mathcal{B})$ which is the same as $\mathbb{P}_0$ on $\mathcal{F}_0$. This unique probability measure on $(0,1]$ is called the **Lebesgue** or **uniform** measure.

The Lebesgue measure formalizes the notion of length. This suggests that the Lebesgue measure of a singleton should be zero. This can be shown as follows. Let $b \in (0,1]$. Using (7.2), we write

$$\mathbb{P}(\{b\}) = \mathbb{P}\left(\bigcap_{n=1}^{\infty} \left(b - \frac{1}{n}, b\right] \cap \Omega\right)$$

Let $A_n = \left(b - \frac{1}{n}, b\right]$. For each n, the lebesgue measure of A_n is

$$\mathbb{P}(A_n) = \frac{1}{n} \tag{7.3}$$

Since A_n is a decreasing sequence of nested sets,

$$\begin{aligned}\mathbb{P}(\{b\}) &= \mathbb{P}\left(\bigcap_{n=1}^{\infty} A_n\right) \\ &= \lim_{n\to\infty} \mathbb{P}(A_n) \\ &= \lim_{n\to\infty} \frac{1}{n} \\ &= 0\end{aligned}$$

where the second equality follows from the continuity of probability measures.

Since any countable set is a countable union of singletons, the probability of a countable set is zero. For example, under the uniform measure on $(0,1]$, the probability of the set of rationals is zero, since the rational numbers in $(0,1]$ form a countable set.

For $\Omega = (0,1]$, the Lebesgue measure is also a probability measure. For other intervals (for example $\Omega = (0,2]$), it will only be a finite measure, which can be normalized as appropriate to obtain a uniform probability measure.

Definition 7.7 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. An event A is said to occur almost surely (a.s) if* $\mathbb{P}(A) = 1$.

Caution: $\mathbb{P}(A) = 1$ does not mean $A = \Omega$.

Lebesgue Measure of the Cantor set: Consider the cantor set K. It is created by repeatedly removing the open middle thirds of a set of line segments. Consider its complement. It contains countable number of disjoint intervals. Hence we have:

$$\mathbb{P}(K^c) = \frac{1}{3} + \frac{2}{9} + \frac{4}{27} + \cdots = \frac{\frac{1}{3}}{1 - \frac{2}{3}} = 1.$$

Therefore $\mathbb{P}(K) = 0$. It is very interesting to note that though the Cantor set is equicardinal with $(0, 1]$, its Lebesgue measure is equal to 0 while the Lebesgue measure of $(0, 1]$ is equal to 1.
We now extend the definition of Lebesgue measure on $[0, 1]$ to the real line, $\mathbb{R}$. We first look at the definition of a Borel set on $\mathbb{R}$. This can be done in several ways, as shown below.

Definition 7.8 *Borel sets on* $\mathbb{R}$*:*

- *Let* $\mathcal{C}$ *be a collection of open intervals in* $\mathbb{R}$*. Then* $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{C})$ *is the Borel set on* $\mathbb{R}$*.*
- *Let* $\mathcal{D}$ *be a collection of semi-infinite intervals* $\{(-\infty, x]; x \in \mathbb{R}\}$*, then* $\sigma(\mathcal{D}) = \mathcal{B}(\mathbb{R})$*.*
- $A \subseteq \mathbb{R}$ *is said to be a Borel set on* $\mathbb{R}$*, if* $A \cap (n, n+1]$ *is a Borel set on* $(n, n+1]$ $\forall n \in \mathbb{Z}$*.*

Exercise: Verify that the three statements are equivalent definitions of Borel sets on $\mathbb{R}$.

Definition 7.9 *Lebesgue measure of* $A \subseteq \mathbb{R}$*:*

$$\lambda(A) = \sum_{n=-\infty}^{\infty} \mathbb{P}_n(A \cap (n, n+1])$$

Theorem 7.10 $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ *is an infinite measure space.*

Proof: We need to prove following:

- $\lambda(\mathbb{R}) = \infty$
- $\lambda(\Phi) = 0$
- The countable additivity property

We see that

$$\mathbb{P}_n(\mathbb{R} \cap (n, n+1]) = 1, \forall n \in \mathbb{I}$$

Hence we have

$$\lambda(\mathbb{R}) = \sum_{n=-\infty}^{\infty} 1 = \infty$$

Now consider $\Phi \cap (n, n+1]$. This is a null set for all n. Hence we have,

$$\mathbb{P}_n(\Phi \cap (n, n+1]) = 0, \forall n \in \mathbb{I}$$

which implies,

$$\lambda(\Phi) = \sum_{n=-\infty}^{\infty} \mathbb{P}_n(\Phi \cap (n, n+1]) = 0$$

We now need to prove the countable additivity property. For this we consider $A_i \in \mathcal{B}(\mathbb{R})$ such that the sequence $A_1, A_2, \ldots, A_n, \ldots$ are arbitrary pairwise disjoint sets in $\mathcal{B}(\mathbb{R})$. Therefore we obtain,

$$\begin{aligned}
\lambda(\bigcup_{i=1}^{\infty} A_i) &= \sum_{n=-\infty}^{\infty} \mathbb{P}_n(\bigcup_{i=1}^{\infty} A_i \cap (n, n+1]) \\
&= \sum_{n=-\infty}^{\infty} \sum_{i=1}^{\infty} \mathbb{P}_n(A_i \cap (n, n+1]) \\
&= \sum_{i=1}^{\infty} \sum_{n=-\infty}^{\infty} \mathbb{P}_n(A_i \cap (n, n+1])
\end{aligned}$$

The second equality above comes from the fact that the probability measure has countable additivity property. The last equality above comes from the fact that the summations can be interchanged (from Fubini's theorem). We also have the following:

$$\lambda(A_i) = \sum_{n=-\infty}^{\infty} \mathbb{P}_n(A_i \cap (n, n+1])$$

We now immediately see that

$$\lambda(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \lambda(A_i)$$

Hence proved. ■

7.5 Exercises

1. Let $\mathcal{F}$ be a σ-algebra corresponding to a sample space Ω. Let H be a subset of Ω that does not belong to $\mathcal{F}$. Consider the collection $\mathcal{G}$ of all sets of the form $(H \cap A) \cup (H^c \cap B)$, where A and $B \in \mathcal{F}$.
   (a) Show that $H \cap A \in \mathcal{G}$.
   (b) Show that $\mathcal{G}$ is a σ-algebra.
2. Show that $\mathcal{C} = \sigma(\mathcal{C})$ iff $\mathcal{C}$ is a σ-algebra.
3. Let $\mathcal{C}$ and $\mathcal{D}$ be two collections of subsets of Ω such that $\mathcal{C} \subseteq \mathcal{D}$. Prove that $\sigma(\mathcal{C}) \subseteq \sigma(\mathcal{D})$.
4. Prove that the following subsets of $(0, 1]$ are Borel-measurable.
   (a) any countable set
   (b) the set of irrational numbers
   (c) the Cantor set (Hint: rather than defining it in terms of ternary expansions, it's easier to use the equivalent definition of the Cantor set that involves sequentially removing the "middle-third" open intervals; see Wikipedia for example).
   (d) The set of numbers in $(0, 1]$ whose decimal expansion does not contain 7.
5. Let $\mathcal{B}$ denote the Borel σ-algebra as defined in class. Let $\mathcal{C}_c$ denote the set of all closed intervals contained in $(0, 1]$. Show that $\sigma(\mathcal{C}_c) = \mathcal{B}$. In other words, we could have very well defined the Borel σ-algebra as being generated by closed intervals, rather than open intervals.

6. Let $\Omega = [0, 1]$, and let $\mathcal{F}_3$ consist of all countable subsets of Ω, and all subsets of Ω having a countable complement. It can be shown that $\mathcal{F}_3$ is a σ-algebra (Refer Lecture 4, Exercises, 6(d)). Let us define $\mathbb{P}(A) = 0$ if A countable, and $\mathbb{P}(A) = 1$ if A has a countable complement. Is $(\Omega, \mathcal{F}_3, \mathbb{P})$ a legitimate probability space?

7. We have seen in 4(c) that the Cantor set is Borel-measurable. Show that the Cantor set has zero Lebesgue measure. Thus, although the Cantor set can be put into a bijection with $[0, 1]$, it has zero Lebesgue measure!

References

[1] Rosenthal, J. S. (2006). A first look at rigorous probability theory (Vol. 2). Singapore: World Scientific.

[2] Williams, D. (1991). Probability with martingales. Cambridge university press.