TWO FUNDAMENTAL PROBABILISTIC MODELS

Contents

The following are two fundamental probabilistic models that can serve as building blocks for more complex models:

(a) The **uniform distribution** on $[0, 1]$, which assigns probability $b-a$ to every interval $[a, b] \subset [0, 1]$.

(b) A model of an infinite sequence of fair coin tosses that assigns equal probability, $1/2^n$, to every possible sequence of length n.

These two models are often encountered in elementary probability and used without further discussion. Strictly speaking, however, we need to make sure that these two models are well-posed, that is, consistent with the axioms of probability. To this effect, we need to define appropriate σ-fields and probability measures on the corresponding sample spaces. In what follows, we describe the required construction, while omitting the proofs of the more technical steps.

1 CARATHÉODORY'S EXTENSION THEOREM

The general outline of the construction we will use is as follows. We are interested in defining a probability measure with certain properties on a given measurable space $(\Omega, \mathcal{F})$. We consider a smaller collection, $\mathcal{F}_0 \subset \mathcal{F}$, of subsets of Ω, which is a field, and on which the desired probabilities are easy to define. (Recall that a field is a collection of subsets of the sample space that includes the empty set, and which is closed under taking complements and under *finite* unions.) Furthermore, we make sure that $\mathcal{F}_0$ is rich enough, so that the σ-field it generates is the same as the desired σ-field $\mathcal{F}$. We then extend the definition of the probability measure from $\mathcal{F}_0$ to the entire σ-field $\mathcal{F}$. This is possible, under few conditions, by virtue of the following fundamental result from measure theory.

Theorem 1. **(Carathéodory's extension theorem)** *Let $\mathcal{F}_0$ be a field of subsets of a sample space Ω, and let $\mathcal{F} = \sigma(\mathcal{F}_0)$ be the σ-field 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 $A \in \mathcal{F}_0$.

Remarks:

(a) The proof of the extension theorem is fairly long and technical; see, e.g., Appendix A of [Williams].

(b) The main hurdle in applying the extension theorem is the verification of the countable additivity property of $\mathbb{P}_0$ on $\mathcal{F}_0$; that is, one needs to show that if $\{A_i\}$ is a sequence disjoint sets in $\mathcal{F}_0$, and if $\cup_{i=1}^{\infty} A_i \in \mathcal{F}_0$, then $\mathbb{P}_0(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}_0(A_i)$. Alternatively, in the spirit of Theorem 1 from Lecture 1, it suffices to verify that if $\{A_i\}$ is a decreasing sequence of sets in $\mathcal{F}_0$ and if $\cap_{i=1}^{\infty} A_i$ is empty, then $\lim_{n\to\infty} \mathbb{P}_0(A_i) = 0$. Indeed, while Theorem 1 of Lecture 1 was stated for the case where $\mathcal{F}$ is a σ-field, an inspection of its proof indicates that it remains valid even if $\mathcal{F}$ is replaced by a field $\mathcal{F}_0$.

In the next two sections, we consider the two models of interest. We define appropriate fields, define probabilities for the events in those fields, and then use the extension theorem to obtain a probability measure.

2 LEBESGUE MEASURE ON $[0,1]$ AND ON $\mathbb{R}$

In this section, we construct the **uniform** probability measure on $[0,1]$, also known as **Lebesgue** measure. Under the Lebesgue measure, the measure assigned to any subset of $[0,1]$ is meant to be equal to its length. While the definition of length is immediate for simple sets (e.g., the set $[a,b]$ has length $b-a$), more general sets present more of a challenge.

We start by considering the sample space $\Omega = (0,1]$, which is slightly more convenient than the sample space $[0,1]$.

2.1 A σ-field and a field of subsets of $(0,1]$

Consider the collection $\mathcal{C}$ of all intervals $[a,b]$ contained in $(0,1]$, and let $\mathcal{F}$ be the σ-field generated by $\mathcal{C}$. As mentioned in the Lecture 1 notes, this is called the **Borel** σ-field, and is denoted by $\mathcal{B}$. Sets in this σ-field are called **Borel sets** or **Borel measurable sets**.

Any set that can be formed by starting with intervals $[a,b]$ and using a countable number of set-theoretic operations (taking complements, or forming countable unions and intersections of previously formed sets) is a Borel set. For example, it can be verified that single-element sets, $\{a\}$, are Borel sets. Furthermore, intervals $(a,b]$ are also Borel sets since they are of the form $[a,b] \setminus \{a\}$. Every countable set is also a Borel set, since it is the union of countably many single-element sets. In particular, the set of rational numbers in $(0,1]$, as well as its complement, the set of irrational numbers in $(0,1]$, is a Borel set. While Borel sets can be fairly complicated, not every set is a Borel set; see Sections 5-6.

Directly defining a probability measure for all Borel sets directly is difficult, so we start by considering a smaller collection, $\mathcal{F}_0$, of subsets of $(0,1]$. We let $\mathcal{F}_0$ consist of the empty set and all sets that are finite unions of intervals of the form $(a,b]$. In more detail, if a set $A \in \mathcal{F}_0$ is nonempty, it is of the form

$$A = (a_1, b_1] \cup \cdots \cup (a_n, b_n],$$

where $0 \le a_1 < b_1 \le a_2 < b_2 \le \cdots \le a_n < b_n \le 1$, and $n \in \mathbb{N}$.

Lemma 1. *We have $\sigma(\mathcal{F}_0) = \sigma(\mathcal{C}) = \mathcal{B}$.*

Proof. We have already argued that every interval of the form $(a,b]$ is a Borel set. Hence, a typical element of $\mathcal{F}_0$ (a finite union of such intervals) is also a Borel set. Therefore, $\mathcal{F}_0 \subset \mathcal{B}$, which implies that $\sigma(\mathcal{F}_0) \subset \sigma(\mathcal{B}) = \mathcal{B}$. (The last equality holds because $\mathcal{B}$ is already a σ-field and is therefore equal to the smallest σ-field that contains $\mathcal{B}$.)

Note that for $a > 0$, we have $[a,b] = \cap_{n=1}^{\infty}(a-1/n, b]$. Since $(a-1/n, b] \in \mathcal{F}_0 \subset \sigma(\mathcal{F}_0)$, it follows that $[a,b] \in \sigma(\mathcal{F}_0)$. Thus, $\mathcal{C} \subset \sigma(\mathcal{F}_0)$, which implies that

$$\mathcal{B} = \sigma(\mathcal{C}) \subset \sigma\big(\sigma(\mathcal{F}_0)\big) = \sigma(\mathcal{F}_0) \subset \mathcal{B}.$$

(The second equality holds because the smallest σ-field containing $\sigma(\mathcal{F}_0)$ is $\sigma(\mathcal{F}_0)$ itself.) The first equality in the statement of the proposition follows. Finally, the equality $\sigma(\mathcal{C}) = \mathcal{B}$ is just the definition of $\mathcal{B}$. $\square$

Lemma 2.

(a) The collection $\mathcal{F}_0$ is a field.

(b) The collection $\mathcal{F}_0$ is not a σ-field.

Proof.

(a) By definition, Ø $\in \mathcal{F}_0$. Note that Ø$^c = (0,1] \in \mathcal{F}_0$. More generally, if A is of the form $A = (a_1, b_1] \cup \cdots \cup (a_n, b_n]$, its complement is $(0, a_1] \cup (b_1, a_2] \cup \cdots \cup (b_n, 1]$, which is also in $\mathcal{F}_0$. Furthermore, the union of two sets that are unions of finitely many intervals of the form $(a, b]$ is also a union of finitely many such intervals. For example, if $A = (1/8, 2/8] \cup (4/8, 7/8]$ and $B = (3/8, 5/8]$, then $A \cup B = (1/8, 2/8] \cup (3/8, 7/8]$.

(b) To see that $\mathcal{F}_0$ is not a σ-field, note that $(0, n/(n+1)] \in \mathcal{F}_0$, for every $n \in \mathbb{N}$, but the union of these sets, which is $(0, 1)$, does not belong to $\mathcal{F}_0$. $\square$

2.2 The uniform measure on $(0, 1]$

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

$$A = (a_1, b_1] \cup \cdots \cup (a_n, b_n],$$

where $0 \leq a_1 < b_1 \leq a_2 < b_2 \leq \cdots \leq a_n < b_n \leq 1$, and $n \in \mathbb{N}$, we define

$$\mathbb{P}_0(A) = (b_1 - a_1) + \cdots + (b_n - a_n),$$

which is its total length. Note that $\mathbb{P}_0(\Omega) = \mathbb{P}\big((0, 1]\big) = 1$. Also $\mathbb{P}_0$ is *finitely additive*. Indeed if $A_1, \ldots, A_n$ are disjoint finite unions of intervals of the form $(a, b]$, then $A = \cup_{1 \leq i \leq n} A_i$ is also a finite union of such intervals and its total length is the sum of the lengths of the sets A_i. It turns out that $\mathbb{P}_0$ is also *countably additive* on $\mathcal{F}_0$. This essentially boils down to checking the following. If $(a, b] = \cup_{i=1}^{\infty} (a_i, b_i]$, where the intervals $(a_i, b_i]$ are disjoint, then $b - a = \sum_{i=1}^{\infty} (b_i - a_i)$. This may appear intuitively obvious, but a formal proof is nontrivial; see, e.g., [Williams, Chapter A1].

We can now apply the Extension Theorem and conclude that there exists a probability measure $\mathbb{P}$, called the **Lebesgue** or **uniform** measure, defined on the entire Borel σ-field $\mathcal{B}$, that agrees with $\mathbb{P}_0$ on $\mathcal{F}_0$. In particular, $\mathbb{P}\big((a,b]\big) = b-a$, for every interval $(a,b] \subset (0,1]$.

By augmenting the sample space Ω to include 0, and assigning zero probability to it, we obtain a new probability model with sample space $\Omega = [0,1]$. (Exercise: define formally the sigma-field on $[0,1]$, starting from the σ-field on $(0,1]$.)

Exercise 1. *Let A be the set of irrational numbers in $[0,1]$. Show that $\mathbb{P}(A) = 1$.*

Example. Let A be the set of points in $[0,1]$ whose decimal representation contains only odd digits. (We disallow decimal representations that end with an infinite string of nines. Under this condition, every number has a unique decimal representation.) What is the Lebesgue measure of this set?

Observe that $A = \cap_{n=1}^{\infty} A_n$, where A_n is the set of points whose first n digits are odd. It can be checked that A_n is a union of 5^n intervals, each with length $1/10^n$. Thus, $\mathbb{P}(A_n) = 5^n/10^n = 1/2^n$. Since $A \subset A_n$, we obtain $\mathbb{P}(A) \leq \mathbb{P}(A_n) = 1/2^n$. Since this is true for every n, we conclude that $\mathbb{P}(A) = 0$.

Exercise 2. *Let A be the set of points in $[0,1]$ whose decimal representation contains at least one digit equal to 9. Find the Lebesgue measure of that set.*

Note that there is nothing special about the interval $(0,1]$. For example, if we let $\Omega = (c,d]$, where $c < d$, and if $(a,b] \subset (c,d]$, we can define $\mathbb{P}_0\big((a,b]\big) = (b-a)/(d-c)$ and proceed as above to obtain a uniform probability measure on the set $(c,d]$, as well as on the set $[c,d]$.

On the other hand, a "uniform" probability measure on the entire real line, $\mathbb{R}$, that assigns equal probability to intervals of equal length, is incompatible with the requirement $\mathbb{P}(\Omega) = 1$. What we obtain instead, in the next section, is a notion of length which becomes infinite for certain sets.

2.3 The Lebesgue measure on $\mathbb{R}$

Let $\Omega = \mathbb{R}$. We first define a σ-field of subsets of $\mathbb{R}$. This can be done in several ways. It can be verified that the three alternatives below are equivalent.

(a) Let $\mathcal{C}$ be the collection of all intervals of the form $[a,b]$, and let $\mathcal{B} = \sigma(\mathcal{C})$ be the σ-field that it generates.

(b) Let $\mathcal{D}$ be the collection of all intervals of the form $(-\infty,b]$, and let $\mathcal{B} = \sigma(\mathcal{D})$ be the σ-field that it generates.

(c) For any n, we define the Borel σ-field of $(n,n+1]$ as the σ-field generated by sets of the form $[a,b] \subset (n,n+1]$. We then say that A is a Borel subset of $\mathbb{R}$ if $A \cap (n,n+1]$ is a Borel subset of $(n,n+1]$, for every n.

Exercise 3. *Show that the above three definitions of $\mathcal{B}$ are equivalent.*

Let $\mathbb{P}_n$ be the uniform measure on $(n, n+1]$ (defined on the Borel sets in $(n, n+1]$). Given a set $A \subset \mathbb{R}$, we decompose it into countably many pieces, each piece contained in some interval $(n, n+1]$, and define its "length" $\mu(A)$ using countable additivity:

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

It turns out that μ is a measure on $(\mathbb{R}, \mathcal{B})$, called again **Lebesgue measure**. However, it is not a probability measure because $\mu(\mathbb{R}) = \infty$.

Exercise 4. *Show that μ is a measure on $(\mathbb{R}, \mathcal{B})$. Hint: Use the countable additivity of the measures $\mathbb{P}_n$ to establish the countable additivity of μ. You can also the fact that if the numbers a_{ij} are nonnegative, then $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} a_{ij}$.*

3 COIN TOSSES: A "UNIFORM" MEASURE ON $\{0,1\}^\infty$

Consider an infinite sequence of fair coin tosses. We wish to construct a probabilistic model of this experiment under which every possible sequence of results of the first n tosses has the same probability, $1/2^n$.

The sample space for this experiment is the set $\{0,1\}^\infty$ of all infinite sequences $\omega = (\omega_1, \omega_2, \ldots)$ of zeroes and ones (we use zeroes and ones instead of heads and tails).

For technical reasons, it is not possible to assign a probability to every subset of the sample space. Instead, we proceed as in Section 2, i.e., first define a field of subsets, assign probabilities to sets that belong to this field, and then extend them to a probability measure on the σ-field generated by that field.

3.1 A field and a σ-field of subsets of $\{0,1\}^n$

Let $\mathcal{F}_n$ be the collection of events whose occurrence can be decided by looking at the results of the first n tosses. For example, the event $\{\omega \mid \omega_1 = 1 \text{ and } \omega_2 \neq \omega_4\}$ belongs to $\mathcal{F}_4$ (as well as to $\mathcal{F}_k$ for every $k \geq 4$).

Let B be an arbitrary subset of $\{0,1\}^n$. Consider the set

$$A = \{\omega \in \{0,1\}^\infty \mid (\omega_1, \omega_2, \ldots, \omega_n) \in B\}.$$

We can express $A \subset \{0,1\}^\infty$ in the form $A = B \times \{0,1\}^\infty$. (This is simply saying that any sequence in A can be viewed as a pair consisting of a n-long sequence that belongs to B, followed by an arbitrary infinite sequence. The event A belongs to $\mathcal{F}_n$, and all elements of $\mathcal{F}_n$ are of this form, for some A. It is easily verified that $\mathcal{F}_n$ is a σ-field.

Exercise 5. *Provide a formal proof that $\mathcal{F}_n$ is a σ-field.*

The σ-field $\mathcal{F}_n$, for any fixed n, is too small; it can only serve to model the first n coin tosses. We are interested instead in sets that belong to $\mathcal{F}_n$, for arbitrary n, and this leads us to define $\mathcal{F}_0 = \cup_{n=1}^{\infty}\mathcal{F}_n$, the collection of sets that belong to $\mathcal{F}_n$ for some n. Intuitively, $A \in \mathcal{F}_0$ if the occurrence or nonoccurrence of A can be decided after a fixed number of coin tosses.[1]

Example. Let $A_n = \{\omega \mid \omega_n = 1\}$, the event that the nth toss results in a "1". Note that $A_n \in \mathcal{F}_n$. Let $A = \cup_{i=1}^{\infty} A_n$, which is the event that there is at least one "1" in the infinite toss sequence. The event A does not belong to $\mathcal{F}_n$, for any n. (Intuitively, having observed a sequence of n zeroes does not allow us to decide whether there will be a subsequent "1" or not.) Consider also the complement of A, which is the event that the outcome of the experiment is an infinite string of zeroes. Once more, we see that A^c does not belong to $\mathcal{F}_0$.

The preceding example shows that $\mathcal{F}_0$ is not a σ-field. On the other hand, it can be verified that $\mathcal{F}_0$ is a field.

Exercise 6. *Prove that $\mathcal{F}_0$ is a field.*

We would like to have a probability model that assigns probabilities to all of the events in $\mathcal{F}_n$, for every n. This means that we need a σ-field that includes $\mathcal{F}_0$. On the other hand, we would like our σ-field to be as small as possible, i.e., contain as few subsets of $\{0,1\}^n$ as possible, to minimize the possibility that it includes pathological sets to which probabilities cannot be assigned. This leads us to define $\mathcal{F}$ as the sigma-field $\sigma(\mathcal{F}_0)$ generated by $\mathcal{F}_0$.

3.2 A probability measure on $\big(\{0,1\}^{\infty}, \mathcal{F}\big)$

We start by defining a finitely additive function $\mathbb{P}_0$ on the field $\mathcal{F}_0$ that also satisfies $\mathbb{P}_0(\{0,1\}^{\infty}) = 1$. This is accomplished as follows. Every set A in $\mathcal{F}_0$ is of the form $B \times \{0,1\}^{\infty}$, for some n and some $B \subset \{0,1\}^n$. We then let $\mathbb{P}_0(A) = |B|/2^n$.[2] Note that the event $\{\omega_1, \omega_2, \ldots, \omega_n\} \times \{0,1\}^n$, which is the event that the first n tosses resulted in a particular sequence $\{\omega_1, \omega_2, \ldots, \omega_n\}$,

[1] The union $\cup_{i=1}^{\infty}\mathcal{F}_i = \mathcal{F}_0$ is not the same as the collection of sets of the form $\cup_{i=1}^{\infty} A_i$, for $A_i \in \mathcal{F}_i$. For an illustration, if $\mathcal{F}_1 = \{\{a\},\{b,c\}\}$ and $\mathcal{F}_2 = \{\{d\}\}$, then $\mathcal{F}_1 \cup \mathcal{F}_2 = \{\{a\},\{b,c\},\{d\}\}$. Note that $\{b,c\} \cup \{d\} = \{b,c,d\}$ is not in $\mathcal{F}_1 \cup \mathcal{F}_2$.

[2] For any set A, $|A|$ denotes its cardinality, the number of elements it contains.

is assigned probability $1/2^n$. In particular, all possible sequences of length are assigned equal probability, as desired.

Before proceeding further, we need to verify that the above definition is *consistent*. Note that same set A can belong to $\mathcal{F}_n$ for several values of n. We therefore need to check that when we apply the definition of $\mathbb{P}_0(A)$ for different choices of n, we obtain the same value. Indeed, suppose that $A \in \mathcal{F}_m$, which implies that $A \in \mathcal{F}_n$, for $n > m$. In this case, $A = B \times \{0,1\}^\infty = C \times \{0,1\}^\infty$, where $B \subset \{0,1\}^n$ and $C \subset \{0,1\}^m$. Thus, $B = C \times \{0,1\}^{n-m}$, and $|B| = |C| \cdot 2^{n-m}$. One application of the definition yields $\mathbb{P}_0(A) = |B|/2^n$, and another yields $\mathbb{P}_0(A) = |C|/2^m$. Since $|B| = |C| \cdot 2^{n-m}$, they both yield the same value.

It is easily verified that $\mathbb{P}_0(\Omega) = 1$, and that $\mathbb{P}_0$ is finitely additive: if $A, B \subset \mathcal{F}_n$ are disjoint, then $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$. It also turns out that $\mathbb{P}_0$ is also countably additive on $\mathcal{F}_0$. (The proof of this fact is more elaborate, and is omitted.) We can now invoke the Extension Theorem and conclude that there exists a unique probability measure on $\mathcal{F}$, the σ-field generated by $\mathcal{F}_0$, that agrees with $\mathbb{P}_0$ on $\mathcal{F}_0$. This probability measure assigns equal probability, $1/2^n$, to every possible sequence of length n, as desired. This confirms that the intuitive process of an infinite sequence of coin flips can be captured rigorously within the framework of probability theory.

Exercise 7. *Consider the probability space $(\{0,1\}^\infty, \mathcal{F}, \mathbb{P})$. Let A be the set of all infinite sequences ω for which $\omega_n = 0$ for every odd n.*

(a) Establish that $A \notin \mathcal{F}_0$, but $A \in F$.

(b) Compute $\mathbb{P}(A)$.

Similar to the case of Borel sets in $[0,1]$, there exist subsets of $\{0,1\}^\infty$ that do not belong to $\mathcal{F}$. In fact the similarities between the models of Sections 2 and 3 are much deeper; the two models are essentially equivalent, although we will not elaborate on the meaning of this. Let us only say that the equivalence relies on the one-to-one correspondence of the sets $[0,1]$ and $\{0,1\}^\infty$ obtained through the binary representation of real numbers. Intuitively, generating a real number at random, according to the uniform distribution (Lebesgue measure) on $[0,1]$, is probabilistically equivalent to generating each bit in its binary expansion at random.

4 COMPLETION OF A PROBABILITY SPACE

Starting with a field $\mathcal{F}_0$ and a countably additive function $\mathbb{P}_0$ on that field, the Extension Theorem leads to a measure on the smallest σ-field containing $\mathcal{F}_0$.

Can we extend the measure further, to a larger σ-field? If so, is the extension unique, or will there have to be some arbitrary choices? We describe here a generic extension that assigns probabilities to certain additional sets A for which there is little choice.

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Suppose that $B \in \mathcal{F}$, and $\mathbb{P}(B) = 0$. Any set B with this property is called a **null** set. (Note that in this context, "null" is not the same as "empty.") Suppose now that $A \subset B$. If the set A is not in $\mathcal{F}$, it is not assigned a probability; were it to be assigned one, the only choice that would not lead to a contradiction is a value of zero.

The first step is to augment the σ-field $\mathcal{F}$ so that it includes all subsets of null sets. This is accomplished as follows:

(a) Let $\mathcal{N}$ be the collection of all subsets of null sets;

(b) Define $\mathcal{F}^* = \sigma(\mathcal{F} \cup \mathcal{N})$, the smallest σ-field that contains $\mathcal{F}$ as well as all subsets of null sets.

(c) Extend $\mathbb{P}$ in a natural manner to obtain a new probability measure $\mathbb{P}^*$ on $(\Omega, \mathcal{F}^*)$. In particular, we let $\mathbb{P}^*(A) = 0$ for every subset $A \subset B$ of every null set $B \in \mathcal{F}$. It turns out that such an extension is always possible and unique.

The resulting probability space is said to be **complete**. It has the property that all subsets of null sets are included in the σ-field and are also null sets.

When $\Omega = [0, 1]$ (or $\Omega = \mathbb{R}$), $\mathcal{F}$ is the Borel σ-field, and $\mathbb{P}$ is Lebesgue measure, we obtain an augmented σ-field $\mathcal{F}^*$ and a measure $\mathbb{P}^*$. The sets in $\mathcal{F}^*$ are called **Lebesgue measurable** sets. The new measure $\mathbb{P}^*$ is referred to by the same name as the measure $\mathbb{P}$ ("Lebesgue measure").

5 FURTHER REMARKS

We record here a few interesting facts related to Borel σ-fields and the Lebesgue measure. Their proofs tend to be fairly involved.

(a) There exist sets that are Lebesgue measurable but not Borel measurable, i.e., $\mathcal{F}$ is a proper subset of $\mathcal{F}^*$.

(b) There are as many Borel measurable sets as there are points on the real line (this is the "cardinality of the continuum"), but there are as many Lebesgue measurable sets as there are subsets of the real line (which is a higher cardinality) [Billingsley]

(c) There exist subsets of $[0, 1]$ that are not Lebesgue measurable; see Section 6 and [Williams, p. 192].

(d) It is not possible to construct a probability space in which the σ-field includes all subsets of $[0,1]$, with the property that $\mathbb{P}(\{x\}) = 0$ for every $x \in (0,1]$ [Billingsley, pp. 45-46].

6 APPENDIX: ON STRANGE SETS (optional reading)

In this appendix, we provide some evidence that not every subset of $(0,1]$ is Lebesgue measurable, and, furthermore, that Lebesgue measure cannot be extended to a measure defined for all subsets of $(0,1]$.

Let "$+$" stand for addition modulo 1 in $(0,1]$. For example, $0.5+0.7 = 0.2$, instead of 1.2. You may want to visualize $(0,1]$ as a circle that wraps around so that after 1, one starts again at 0. If $A \subset (0,1]$, and x is a number, then $A + x$ stands for the set of all numbers of the form $y + x$ where $y \in A$.

Define x and y to be *equivalent* if $x + r = y$ for some rational number r. Then, $(0,1]$ can be partitioned into equivalence classes. (That is, all elements in the same equivalence class are equivalent, elements belonging to different equivalent classes are not equivalent, and every $x \in (0,1]$ belongs to exactly one equivalence class.) Let us pick exactly one element from each equivalence class, and let H be the set of the elements picked this way. (This fact that a set H can be legitimately formed this way involves the Axiom of Choice, a generally accepted axiom of set theory.) We will now consider the sets of the form $H + r$, where r ranges over the rational numbers in $(0,1]$. Note that there are countably many such sets.

The sets $H + r$ are disjoint. (Indeed, if $r_1 \neq r_2$, and if the two sets $H + r_1$, $H + r_2$ share the point $h_1 + r = h_2 + r_2$, with $h_1, h_2 \in H$, then h_1 and h_2 differ by a rational number and are equivalent. If $h_1 \neq h_2$, this contradicts the construction of H, which contains exactly one element from each equivalence class. If $h_1 = h_2$, then $r_1 = r_2$, which is again a contradiction.) Therefore, $(0,1]$ is the union of the countably many disjoint sets $H + r$.

The sets $H + r$, for different r, are "translations" of each other (they are all formed by starting from the set H and adding a number, modulo 1). Let us say that a measure is *translation-invariant* if it has the following property: if A and $A + x$ are measurable sets, then $\mathbb{P}(A) = \mathbb{P}(A + x)$. Suppose that $\mathbb{P}$ is a translation invariant probability measure, defined on all subsets of $(0,1]$. Then,

$$1 = \mathbb{P}\big((0,1]\big) = \sum_r \mathbb{P}(H + r) = \sum_r \mathbb{P}(H),$$

where the sum is taken over all rational numbers in $(0,1]$. But this impossible. We conclude that a translation-invariant measure, defined on all subsets of $(0,1]$ does not exist.

On the other hand, it can be verified that the Lebesgue measure is translation-invariant on the Borel σ-field, as well as its extension, the Lebesgue σ-field. This implies that the Lebesgue σ-field does not include all subsets of $(0, 1]$.

An even stronger, and more counterintuitive example is the following. It indicates, that the ordinary notion of area or volume cannot be applied to arbitrary sets.

The Banach-Tarski Paradox. Let S be the two-dimensional surface of the unit sphere in three dimensions. There exists a subset F of S such that for any $k \geq 3$,

$$S = (\tau_1 F) \cup \cdots \cup (\tau_k F),$$

where each τ_i is a rigid rotation and the sets $\tau_i F$ are disjoint. For example, S can be made up by three rotated copies of F (suggesting probability equal to 1/3, but also by four rotated copies of F, suggesting probability equal to 1/4). Ordinary geometric intuition clearly fails when dealing with arbitrary sets.