University of Sheffield

School of Mathematics & and Statistics

Measure and Probability

MAS350/451/6352

Spring 2018

Chapter 1

Measure Spaces and Measure

1.1 What is Measure?

Measure theory is the abstract mathematical theory that underlies all models of measurement in the real world. This includes measurement of length, area and volume, mass but also chance/probability. Measure theory is on the one hand a branch of pure mathematics, but it also plays a key role in many applied areas such as physics and economics. In particular it provides a foundation for both the modern theory of integration and also the theory of probability. It is one of the milestones of modern analysis and is an invaluable tool for functional analysis.

To motivate the key definitions, suppose that we want to measure the lengths of several line segments. We represent these as closed intervals of the real number line $\mathbb{R}$ so a typical line segment is $[a, b]$ where $b > a$. We all agree that its length is $b - a$. We write this as

$$m([a, b]) = b - a$$

and interpret this as telling us that the measure m of length of the line segment $[a, b]$ is the number $b - a$. We might also agree that if $[a_1, b_1]$ and $[a_2, b_2]$ are two non-overlapping line segments and we want to measure their combined length then we want to apply m to the set-theoretic union $[a_1, b_1] \cup [a_2, b_2]$ and

$$m([a_1, b_1] \cup [a_2, b_2]) = (b_2 - a_2) + (b_1 - a_1) = m([a_1, b_1]) + m([a_2, b_2]). \tag{1.1.1}$$

An isolated point c has zero length and so

$$m(\{c\}) = 0.$$

and if we consider the whole real line in its entirety then it has infinite length, i.e.

$$m(\mathbb{R}) = \infty.$$

We have learned so far that if we try to abstract the notion of a measure of length, then we should regard it as a mapping m defined on subsets of the real line and taking values in the extended non-negative real numbers $[0, \infty]$.

Question Does it make sense to consider m on **all** subsets of $\mathbb{R}$?

Example 1.1. *The Cantor Set.* Start with the interval $[0, 1]$ and remove the middle third to create the set $C_1 = [0, 1/3) \cup (2/3, 1]$. Now remove the middle third of each remaining piece to get $C_2 = [0, 1/9) \cup (2/9, 1/3) \cup (2/3, 7/9) \cup (8/9, 1]$. Iterate this process so for $n > 2, C_n$ is obtained from C_{n-1} by removing the middle third of each set within that union. The *Cantor set* is $C = \bigcap_{n=1}^{\infty} C_n$. It turns out that C is uncountable. Does $m(C)$ make sense?

We'll see later that $m(C)$ does make sense and is a finite number (can you guess what it is?). But it turns out that there are even wilder sets in $\mathbb{R}$ than C which have no length. These are quite difficult to construct (they require the axiom of choice) so we won't try to describe them here.

Conclusion. The set of all subsets of $\mathbb{R}$ is its power set $\mathcal{P}(\mathbb{R})$. We've just learned that the power set is too large to support a good theory of measure of length. So we need to find a smaller class of subsets that we can work with.

1.2 Sigma Algebras

So far we have only discussed length but now we want to be more ambitious. Let S be an arbitrary set. We want to define mappings from subsets of S to $[0, \infty]$ which we will continue to denote by m. These will be called measures and they will share some of the properties that we've just been looking at for measures of length. Now on what type of subset of S can m be defined? The power set of S is $\mathcal{P}(S)$ and we have just argued that this could be too large for our purposes as it may contain sets that can't be measured.

Suppose that A and B are subsets of S that we can measure. Then we should surely be able to measure the complement A^c, the union $A \cup B$ and the whole set S. Note that we can then also measure $A \cap B = (A^c \cup B^c)^c$. This leads to a definition

Definition. Let S be a set. A *Boolean algebra* $\mathbf{B}$ is a subset of $\mathcal{P}(S)$ that has the following properties

B(i) $S \in \mathbf{B}$,

B(ii) If $A, B \in \mathbf{B}$ then $A \cup B \in \mathbf{B}$,

B(iii) If $A \in \mathbf{B}$ then $A^c \in \mathbf{B}$.

Boolean algebras are named after the British mathematician George Boole (1815-1864) who introduced them in his book *The Laws of Thought* published in 1854. They are wonderful mathematical objects that are extremely useful in logic and digital electronics. It turns out that they are inadequate for our purposes.

First note that if we use induction on B(ii) we can show that if $A_1, A_2, \ldots A_n \in \mathbf{B}$ then $A_1 \cup A_2 \cdots \cup \cdots A_n \in \mathbf{B}$. But we need to be able to do analysis and this requires us to be able to handle infinite unions. The next definition gives us what we need:

Definition. Let S be a set. A σ-algebra Σ is a subset of $\mathcal{P}(S)$ that has the following properties

S(i) $S \in \Sigma$,

S(ii) If (A_n) is a sequence of sets with $A_n \in \Sigma$ for all $n \in \mathbb{N}$ then $\bigcup_{n=1}^{\infty} A_n \in \Sigma$,

S(iii) If $A \in \Sigma$ then $A^c \in \Sigma$.

Facts about σ-algebras:

- By S(i) and S(iii), $\emptyset = S^c \in \Sigma$.
- We have seen in S(ii) that infinite unions of sets in Σ are themselves in Σ. The same is true of finite unions. To see this let $A_1, \ldots, A_m \in \Sigma$ and define the sequence (A'_n) by $A'_n = \begin{cases} A_n \text{ if } 1 \leq n \leq m \\ \emptyset \text{ if } n > m \end{cases}$ Now apply S(ii) to get the result. We can deduce from this that every σ-algebra is a Boolean algebra.
- Σ is also closed under infinite (or finite) intersections. To see this use de Morgan's law to write
$$\bigcap_{i=1}^{\infty} A_i = \left(\bigcup_{i=1}^{\infty} A_i^c\right)^c.$$
- Σ is closed under set theoretic differences $A - B$, since (by definition) $A - B = A \cap B^c$.

Examples of σ-algebras:

1. $\mathcal{P}(S)$ is a σ-algebra. If S is finite with n elements then $\mathcal{P}(S)$ has 2^n elements.

2. For any set S, $\{\emptyset, S\}$ is a σ-algebra which is called the *trivial* σ-algebra. It is the basic tool for modelling logic circuitry where $\emptyset$ corresponds to "OFF" and S to "ON".

3. If S is any set and $A \subset S$ then $\{\emptyset, A, A^c, S\}$ is a σ-algebra .

4. The most important σ-algebra for studying the measure of length is the *Borel σ-algebra* of $\mathbb{R}$ which is denoted $\mathcal{B}(\mathbb{R})$. It is named after the French mathematican Emile Borel (1871-1956) who was one of the founders of measure theory. It is defined rather indirectly and we postpone this definition until after the next section.

We close this section with a definition. A pair (S, Σ) where S is a set and Σ is a σ-algebra of subsets of S is called a *measurable space*. There are many possible choices of Σ to attach to S. For example we can always take Σ to be trivial or the power set. The choice of Σ is determined by what we want to measure. We sometimes say that a set $A \subset S$ is *measurable* if $A \in \Sigma$.

1.3 Measure

Let (S, Σ) be a measurable space. A measure on (S, Σ) is a mapping $m : \Sigma \to [0, \infty]$ which satisfies

M(i) $m(\emptyset) = 0$,

M(ii) (σ-*additivity*) If $(A_n, n \in \mathbb{N})$ is a sequence of sets where each $A_n \in \Sigma$ and if these sets are mutually disjoint, i.e. $A_n \cap A_m = \emptyset$ if $m \neq n$, then

$$m\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} m(A_n).$$

M(ii) may appear to be rather strong. Our earlier discussion about length led us to $m(A \cup B) = m(A) + m(B)$ and straightforward induction then extends this to *finite additivity*: $m(A_1 \cup A_2 \cup \cdots \cup A_n) = m(A_1) + m(A_2) + \cdots + m(A_n)$ but if we were to replace M(ii) by this weaker finite additivity condition, we would not have an adequate tool for applications to analysis or to probability, as we will see.

1.3.1 Basic Properties of Measures

1. (Finite additivity) If $A_1, A_2, \ldots, A_r \in \Sigma$ and are mutually disjoint then

$$m(A_1 \cup A_2 \cup \cdots \cup A_r) = m(A_1) + m(A_2) + \cdots + m(A_r).$$

   To see this define the sequence (A'_n) by $A'_n = \begin{cases} A_n \text{ if } 1 \leq n \leq r \\ \emptyset \text{ if } n > r \end{cases}$

   Then

$$m\left(\bigcup_{i=1}^{r} A_i\right) = m\left(\bigcup_{i=1}^{\infty} A'_i\right) = \sum_{i=1}^{\infty} m(A'_i) = \sum_{i=1}^{r} m(A_i),$$

   where we used M(ii) and then M(i) to get the last two expressions.

2. If $A, B \in \Sigma$ with $B \subset A$ and either $m(A) < \infty$, or $m(A) = \infty$ but $m(B) < \infty$, then

$$m(A - B) = m(A) - m(B). \tag{1.3.2}$$

   To prove this write the disjoint union $A = (A - B) \cup B$ and then use the result of (1) (with $r = 2$).

3. (Monotonicity) If $A, B \in \Sigma$ with $B \subseteq A$ then $m(B) \leq m(A)$.

   If $m(A) < \infty$ this follows from (1.3.2) using the fact that $m(A-B) \geq 0$. If $m(A) = \infty$, the result is immediate.

4. If $A, B \in \Sigma$ are arbitrary (i.e. not necessarily disjoint) then

$$m(A \cup B) + m(A \cap B) = m(A) + m(B). \tag{1.3.3}$$

   The proof of this is Problem 4(a). Note that if $m(A \cap B) < \infty$ we have

$$m(A \cup B) = m(A) + m(B) - m(A \cap B).$$

Now some concepts and definitions. A triple (S, Σ, m) where (S, Σ) is a measurable space and m is a measure on (S, Σ) is called a *measure space*. The extended real number $m(S)$ is called the *total mass* of m. The measure m is said to be *finite* if $m(S) < \infty$. A finite measure is called a *probability measure* if $m(S) = 1$. When we have a probability measure, we use a slightly different notation.

We write Ω instead of S and call it a *sample space.*

We write $\mathcal{F}$ instead of Σ. Sets in $\mathcal{F}$ are called *events.*

We use P instead of m.

The triple $(\Omega, \mathcal{F}, P)$ is called a *probability space.*

1.3.2 Examples of Measures

Example 1. **Counting Measure**

Let S be a finite set and take $\Sigma = \mathcal{P}(S)$. For each $A \subseteq S$ define

$$m(A) = \#(A) \text{ i.e. the number of elements in } A.$$

Example 2. **Dirac Measure**

This measure is named after the famous British physicist Paul Dirac (1902-84). Let (S, Σ) be an arbitrary measurable space and fix $x \in S$. The Dirac measure δ_x at x is defined by

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

Note that we can write counting measure in terms of Dirac measure, so if S is finite and $A \subseteq S$,

$$\#(A) = \sum_{x \in S} \delta_x(A).$$

Example 3. **Discrete Probability Measures**

Let Ω be a countable set and take $\mathcal{F} = \mathcal{P}(\Omega)$. Let $\{p_\omega, \omega \in \Omega\}$ be a set of real numbers which satisfies the conditions

$$p_\omega \geq 0 \text{ for all } \omega \in \Omega \text{ and } \sum_{\omega \in \Omega} p_\omega = 1.$$

Now define the discrete probability measure P by

$$P(A) = \sum_{\omega \in A} p_\omega = \sum_{\omega \in \Omega} p_\omega \delta_\omega(A),$$

for each $A \in \mathcal{F}$.

For example if $\#(\Omega) = n + 1$ and $0 < p < 1$ we can obtain the *binomial distribution* as a probability measure by taking $p_r = \binom{n}{r} p^r (1-p)^{n-r}$ for $r = 0, 1, \ldots, n$.

Example 4. **Measures via Integration**

Let (S, Σ, m) be an arbitrary measure space and $f : S \to [0, \infty)$ be a function that takes non-negative values. In Chapter 3, we will meet a powerful integration theory that allows us to cook up a new measure I_f from m and f (provided that f is suitably well-behaved - see Chapter 2) by the prescription:

$$I_f(A) = \int_A f(x) m(dx),$$

for all $A \in \Sigma$.

1.4 The Borel σ-algebra and Lebesgue Measure

In this section we take S to be the real number line $\mathbb{R}$. We want to describe a measure λ that captures the notion of length as we discussed at the beginning of this chapter. So we should have $\lambda((a,b)) = b - a$. The first question is - which σ-algebra should we use? We have already argued that the power set $\mathcal{P}(\mathbb{R})$ is too big. Our σ-algebra should contain open intervals, and also unions, intersections and complements of these.

Definition. The *Borel σ-algebra* of $\mathbb{R}$ to be denoted $\mathcal{B}(\mathbb{R})$ is the smallest σ-algebra that contains all open intervals (a,b) where $-\infty \leq a < b \leq \infty$. Sets in $\mathcal{B}(\mathbb{R})$ are called *Borel sets.*

Note that $\mathcal{B}(\mathbb{R})$ also contains isolated points $\{a\}$ where $a \in \mathbb{R}$. To see this first observe that $(a,\infty) \in \mathcal{B}(\mathbb{R})$ and also $(-\infty,a) \in \mathcal{B}(\mathbb{R})$. Now by S(iii),$(-\infty,a] = (a,\infty)^c \in \mathcal{B}(\mathbb{R})$ and $[a,\infty) = (-\infty,a)^c \in \mathcal{B}(\mathbb{R})$. Finally as σ-algebras are closed under intersections, $\{a\} = [a,\infty) \cap (-\infty,a] \in \mathcal{B}(\mathbb{R})$. You can show that $\mathcal{B}(\mathbb{R})$ also contains all closed intervals (see Problem 8).

Notes.

1. $\mathcal{B}(\mathbb{R})$ is defined quite indirectly and there is no "formula" that can be used to give the most general element in it. However it is very hard to construct a subset of $\mathbb{R}$ that isn't in $\mathcal{B}(\mathbb{R})$.

2. $\mathcal{B}(S)$ makes sense on any set S for which there are subsets that can be called "open" in a sensible way. In particular this works for metric spaces. The most general type of S for which you can form $\mathcal{B}(S)$ is a *topological space.*

The measure that precisely captures the notion of length is called *Lebesgue measure* in honour of the French mathematician Henri Lebesgue (1875-1941), who founded the modern theory of integration. We will denote it by λ. First we need a definition.

Let $A \in \mathcal{B}(\mathbb{R})$ be arbitrary. A *covering* of A is a finite or countable collection of open intervals $\{(a_n,b_n), n \in \mathbb{N}\}$ so that

$$A \subseteq \bigcup_{n=1}^{\infty} (a_n, b_n).$$

Let $\mathcal{C}_A$ be the set of all coverings of the set A. Then Lebesgue measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is defined by the prescription:

$$\lambda(A) = \inf_{\mathcal{C}_A} \sum_{n=1}^{\infty} (b_n - a_n), \tag{1.4.4}$$

where the inf is taken over all possible coverings of A.

It would take a long time to prove that λ really is a measure so we'll omit that from the course. For the proof, see the standard text books e.g. by Cohn, Schilling or Tao.

Let's check that the definition (1.4.4) agrees with our intuitive ideas about length.

1. If $A = (a, b)$ then $\lambda((a, b)) = b - a$ as expected, since (a, b) is a covering of itself and any other cover will have greater length.

2. If $A = \{a\}$ then choose any $\epsilon > 0$. Then $(a - \epsilon/2, a + \epsilon/2)$ is a cover of a and so $\lambda(\{a\}) \leq (a + \epsilon/2) - (a - \epsilon/2) = \epsilon$. But ϵ is arbitrary and so we conclude that $\lambda(\{a\}) = 0$.

   From (1) and (2), and using M(ii), we deduce that for $a < b$,

   $$\lambda([a, b)) = \lambda(\{a\} \cup (a, b)) = \lambda(\{a\}) + \lambda((a, b)) = b - a.$$

3. If $A = [0, \infty)$ write $A = \bigcup_{n=1}^{\infty} [n - 1, n)$. Then by M(ii), $\lambda([0, \infty)) = \infty$. By a similar argument, $\lambda((-\infty, 0)) = \infty$ and so $\lambda(\mathbb{R}) = \lambda((-\infty, 0)) + \lambda([0, \infty)) = \infty$.

In simple practical examples on Lebesgue measure, it is generally best not to try to use (1.4.4), but to just apply the properties (1) to (3) above:

e.g. to find $\lambda((-3, 10) - (-1, 4))$, use (1.3.2) to obtain

$$\begin{aligned} \lambda((-3, 10) - (-1, 4)) &= \lambda((-3, 10)) - \lambda((-1, 4))) \\ &= (10 - (-3)) - (4 - (-1)) = 13 - 5 = 8. \end{aligned}$$

If I is a closed interval (or in fact any Borel set) in $\mathbb{R}$ we can similarly define $\mathcal{B}(I)$, the Borel σ-algebra of I, to be the smallest σ-algebra containing all open intervals in I. Then Lebesgue measure on $(I, \mathcal{B}(I))$ is obtained by restricting the sets A in (1.4.4) to be in $\mathcal{B}(I)$.

Sets of measure zero play an important role in measure theory. Here are some interesting examples of quite "large" sets that have Lebesgue measure zero

1. **Countable Subsets of $\mathbb{R}$ have Lebesgue Measure Zero**

   Let $A \subset \mathbb{R}$ be countable. Write $A = \{a_1, a_2, \ldots\} = \bigcup_{n=1}^{\infty}\{a_n\}$. Since A is an infinite union of point sets, it is in $\mathcal{B}(\mathbb{R})$. Then

   $$\lambda(A) = \lambda\left(\bigcup_{n=1}^{\infty}\{a_n\}\right) = \sum_{n=1}^{\infty}\lambda(\{a_n\}) = 0.$$

   It follows that

   $$\lambda(\mathbb{N}) = \lambda(\mathbb{Z}) = \lambda(\mathbb{Q}) = 0.$$

   The last of these is particularly intriguing as it tells us that the only contribution to length of sets of real numbers comes from the irrationals.

2. **The Cantor Set has Lebesgue Measure Zero**

   Recall the construction of the Cantor set $C = \bigcap_{n=1}^{\infty} C_n$ given earlier in this chapter. Since C_n is a union of intervals, $C_n \in \mathcal{B}(\mathbb{R})$ for all $n \in \mathbb{N}$. Hence $C \in \mathcal{B}(\mathbb{R})$. We easily see that $\lambda(C_1) = 1 - \frac{1}{3}$ and $\lambda(C_2) = 1 - \frac{1}{3} - \frac{2}{9}$. Iterating we deduce that $\lambda(C_n) = 1 - \sum_{r=1}^{n} \frac{2^{r-1}}{3^r}$ and so

   $$\lambda(C) = 1 - \sum_{r=1}^{\infty} \frac{2^{r-1}}{3^r} = 1 - \frac{1}{3}.\frac{1}{1 - 2/3} = 0.$$

   Some details of the last part of this calculation need more careful justification and this is left for you to do in Problem 9(c). You will need to use results from section 1.5 below.

1.5 Two Useful Theorems About Measure

In this section we return to the consideration of arbitrary measure spaces (S, Σ, m). Let (A_n) be a sequence of sets in Σ. We say that it is *increasing* if $A_n \subseteq A_{n+1}$ for all $n \in \mathbb{N}$ and we write $A = \bigcup_{n=1}^{\infty} A_n$.

A useful technique is the *disjoint union trick* whereby we can write $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n$ where the B_ns are all mutually disjoint by defining $B_1 = A_1$ and for $n > 1, B_n = A_n - A_{n-1}$. e.g. $\mathbb{R} = \bigcup_{n=1}^{\infty}[-n, n]$ and here $B_1 = [-1, 1], B_2 = [-2, -1) \cup (1, 2]$ etc.

Theorem 1.5.1 *If (A_n) is increasing then*

$$m(A) = \lim_{n\to\infty} m(A_n).$$

Proof. We use the disjoint union trick and M(ii) to find that

$$\begin{aligned} m(A) &= m\left(\bigcup_{n=1}^{\infty} B_n\right) \\ &= \sum_{n=1}^{\infty} m(B_n) \\ &= \lim_{N\to\infty} \sum_{n=1}^{N} m(B_n) \\ &= \lim_{N\to\infty} m\left(\bigcup_{n=1}^{N} B_n\right) \\ &= \lim_{N\to\infty} m(A_N). \end{aligned}$$

since $A_N = B_1 \cup B_2 \cup \cdots \cup B_N$. $\square$

Theorem 1.5.2 *If (A_n) is an arbitrary sequence of sets with $A_n \in \Sigma$ for all $n \in \mathbb{N}$ then*

$$m\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} m(A_n).$$

Proof. In Problem 4, you showed that $m(A_1 \cup A_2) + m(A_1 \cap A_2) = m(A_1) + m(A_2)$ from which we deduce that $m(A_1 \cup A_2) \leq m(A_1) + m(A_2)$. By induction we then obtain for all $N \geq 2$,

$$m\left(\bigcup_{n=1}^{N} A_n\right) \leq \sum_{n=1}^{N} m(A_n).$$

Now define $X_N = \bigcup_{n=1}^{N} A_n$. Then $X_N \subseteq X_{N+1}$ and so (X_N) is increasing to $\bigcup_{n=1}^{\infty} X_n = \bigcup_{n=1}^{\infty} A_n$. By Theorem 1.5.1 we have

$$\begin{aligned} m\left(\bigcup_{n=1}^{\infty} A_n\right) &= m\left(\bigcup_{n=1}^{\infty} X_n\right) \\ &= \lim_{N\to\infty} m(X_N) \\ &= \lim_{N\to\infty} m\left(\bigcup_{n=1}^{N} A_n\right) \\ &\leq \lim_{N\to\infty} \sum_{n=1}^{N} m(A_n) \\ &= \sum_{n=1}^{\infty} m(A_n). \end{aligned}$$

1.6 Product Measures

We calculate areas of rectangles by multiplying products of lengths of their sides. This suggests trying to formulate a theory of products of measures. Let (S_1, Σ_1, m_1) and (S_2, Σ_2, m_2) be two measure spaces. Form the Cartesian product $S_1 \times S_2$. We can similarly try to form a product of σ-algebras

$$\Sigma_1 \times \Sigma_2 = \{A \times B; A \in \Sigma_1, B \in \Sigma_2\},$$

but it turns out that $\Sigma_1 \times \Sigma_2$ is not a σ-algebra (or even a Boolean algebra) e.g. take $S_1 = S_2 = \mathbb{R}$ and consider $((0,1) \times (0,1))^c$. Instead we need $\Sigma_1 \otimes \Sigma_2$ which is defined to be the smallest σ-algebra which contains all the sets in $\Sigma_1 \times \Sigma_2$. We state but do not prove:

Theorem 1.6.1 *There exists a measure $m_1 \times m_2$ on $(S_1 \times S_2, \Sigma_1 \otimes \Sigma_2)$ so that for all $A \in \Sigma_1, B \in \Sigma_2$,*

$$(m_1 \times m_2)(A \times B) = m_1(A)m_2(B).$$

The measure $m_1 \times m_2$ is called the *product* of the two measures m_1 and m_2.

For example, consider $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$. We equip it with the Borel σ-algebra, $\mathcal{B}(\mathbb{R}^2) = \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$. Then the product Lebesgue measure $\lambda_2 = \lambda \times \lambda$. It has the property that

$$\lambda_2((a,b) \times (c,d)) = (b-a)(d-c).$$

Given n-measure spaces $(S_1, \Sigma_1, m_1), (S_2, \Sigma_2, m_2), \ldots, (S_n, \Sigma_n, m_n)$, we can iterate the above procedure to define the product σ-algebra $\Sigma_1 \otimes \Sigma_2 \otimes \cdots \otimes \Sigma_n$ and the product measure $m_1 \times m_2 \times \cdots \times m_n$ so that for $A_i \in \Sigma_i, 1 \leq i \leq n$,

$$(m_1 \times m_2 \times \cdots \times m_n)(A_1 \times A_2 \times \cdots \times A_n) = m_1(A_1)m_2(A_2) \cdots m_n(A_n).$$

In particular n-dimensional Lebesgue measure on $\mathbb{R}^n$ may be defined in this way ($n = 3$ yields volume measure).

Of course there are many measures that one can construct on $(S_1 \times S_2, \Sigma_1 \times \Sigma_2)$ and not all of these will be product measures. For probability spaces, product measures are closely related to the notion of independence as we will see later.