Course: Theory of Probability I
Term: Fall 2013
Instructor: Gordan Zitkovic

Lecture 2
Measures

Measure spaces

Definition 2.1 (Measure)**.** Let $(S, \mathcal{S})$ be a measurable space. A mapping $\mu : \mathcal{S} \to [0, \infty]$ is called a **(positive) measure** if

1. $\mu(\emptyset) = 0$, and
2. $\mu(\cup_n A_n) = \sum_{n\in\mathbb{N}} \mu(A_n)$, for all *pairwise disjoint* $\{A_n\}_{n\in\mathbb{N}}$ in $\mathcal{S}$.

A triple $(S, \mathcal{S}, \mu)$ consisting of a non-empty set, a σ-algebra $\mathcal{S}$ on it and a measure μ on $\mathcal{S}$ is called a **measure space**.

Remark 2.2.

1. A mapping whose domain is some nonempty set $\mathcal{A}$ of subsets of some set S is sometimes called a **set function**.
2. If the requirement 2. in the definition of the measure is weakened so that it is only required that $\mu(A_1 \cup \dots \cup A_n) = \mu(A_1) + \dots + \mu(A_n)$, for $n \in \mathbb{N}$, and pairwise disjoint $A_1, \dots, A_n$, we say that the mapping μ is a **finitely-additive measure**. If we want to stress that a mapping μ satisfies the original requirement 2. for *sequences* of sets, we say that μ is **σ-additive** (**countably additive**).

Definition 2.3 (Terminology)**.** A measure μ on the measurable space $(S, \mathcal{S})$ is called

1. a **probability measure**, if $\mu(S) = 1$,
2. a **finite measure**, if $\mu(S) < \infty$,
3. a **σ-finite measure**, if there exists a sequence $\{A_n\}_{n\in\mathbb{N}}$ in $\mathcal{S}$ such that $\cup_n A_n = S$ and $\mu(A_n) < \infty$,
4. **diffuse** or **atom-free**, if $\mu(\{x\}) = 0$, whenever $x \in S$ and $\{x\} \in \mathcal{S}$.

A set $N \in \mathcal{S}$ is said to be **null** if $\mu(N) = 0$.

Example 2.4 (Examples of measures)**.** Let S be a non-empty set, and let $\mathcal{S}$ be a σ-algebra on S.

1. **Measures on countable sets.** Suppose that S is a finite or countable set. Then each measure μ on $\mathcal{S} = 2^S$ is of the form

$$\mu(A) = \sum_{x \in A} p(x),$$

for some function $p : S \to [0, \infty]$ (why?). In particular, for a finite set S with N elements, if $p(x) = 1/N$ then μ is a probability measure called the **uniform measure**[1] on S.

2. **Dirac measure.** For $x \in S$, we define the set function δ_x on $\mathcal{S}$ by

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

It is easy to check that δ_x is indeed a measure on $\mathcal{S}$. Alternatively, δ_x is called the **point mass at** x (or an **atom on** x, or the **Dirac function**, even though it is not really a function). Moreover, δ_x is a probability measure and, therefore, a finite and a σ-finite measure. It is atom free only if $\{x\} \notin \mathcal{S}$.

3. **Counting Measure.** Define a set function $\mu : \mathcal{S} \to [0, \infty]$ by

$$\mu(A) = \begin{cases} \#A, & A \text{ is finite}, \\ \infty, & A \text{ is infinite}, \end{cases}$$

where, as above, $\#A$ denotes the number of elements in the set A. Again, it is not hard to check that μ is a measure - it is called the **counting measure**. Clearly, μ is a finite measure if and only is S is a finite set. μ could be σ-finite, though, even without S being finite. Simply take $S = \mathbb{N}$, $\mathcal{S} = 2^{\mathbb{N}}$. In that case $\mu(S) = \infty$, but for $A_n = \{n\}$, $n \in \mathbb{N}$, we have $\mu(A_n) = 1$, and $S = \cup_n A_n$. Finally, μ is never atom-free and it is a probability measure only if $\#S = 1$.

[1] In the finite case, it has the well-known property that $\mu(A) = \frac{\#A}{\#S}$, where # denotes the cardinality (number of elements).

Example 2.5 (A finitely-additive set function which is not a measure)**.** Let $S = \mathbb{N}$, and $\mathcal{S} = 2^S$. For $A \in S$ define $\mu(A) = 0$ if A is finite and $\mu(A) = \infty$, otherwise. For $A_1, \dots, A_n \subseteq S$, either

1. all A_i is finite, for $i = 1, \dots, n$. Then $\cup_{i=1}^n A_i$ is also finite and so

$$0 = \mu(\cup_{i=1}^n A_i) = \sum_{i=1}^n \mu(A_i), \text{ or}$$

2. at least one A_i is infinite. Then $\cup_{i=1}^n A_i$ is also infinite and so

$$\infty = \mu(\cup_{i=1}^n A_i) = \sum_{i=1}^n \mu(A_i).$$

On the other hand, take $A_i = \{i\}$, for $i \in \mathbb{N}$. Then $\mu(A_i) = 0$, for each $i \in \mathbb{N}$, and, so, $\sum_{i\in\mathbb{N}} \mu(A_i) = 0$, but $\mu(\cup_i A_i) = \mu(\mathbb{N}) = \infty$.

Note: It is possible to construct very simple-looking finite-additive measures which are not σ-additive. For example, there exist $\{0,1\}$-valued finitely-additive measures on all subsets of $\mathbb{N}$, which are not σ-additive. Such objects are called **ultrafilters** and their existence is equivalent to a certain version of the Axiom of Choice.

Proposition 2.6 (First properties of measures)**.** *Let $(S, \mathcal{S}, \mu)$ be a measure space.*

1. *For $A_1, \dots, A_n \in \mathcal{S}$ with $A_i \cap A_j = \emptyset$, for $i \neq j$, we have*

$$\sum_{i=1}^{n} \mu(A_i) = \mu(\cup_{i=1}^{n} A_i) \qquad \textit{(Finite additivity)}$$

2. *If $A, B \in \mathcal{S}$, $A \subseteq B$, then*

$$\mu(A) \leq \mu(B) \qquad \textit{(Monotonicity of measures)}$$

3. *If $\{A_n\}_{n\in\mathbb{N}}$ in $\mathcal{S}$ is increasing, then*

$$\mu(\cup_n A_n) = \lim_n \mu(A_n) = \sup_n \mu(A_n).$$

(Continuity with respect to increasing sequences)

4. *If $\{A_n\}_{n\in\mathbb{N}}$ in $\mathcal{S}$ is decreasing and $\mu(A_1) < \infty$, then*

$$\mu(\cap_n A_n) = \lim_n \mu(A_n) = \inf_n \mu(A_n).$$

(Continuity with respect to decreasing sequences)

5. *For a sequence $\{A_n\}_{n\in\mathbb{N}}$ in $\mathcal{S}$, we have*

$$\mu(\cup_n A_n) \leq \sum_{n\in\mathbb{N}} \mu(A_n). \qquad \textit{(Subadditivity)}$$

Proof.

1. Note that the sequence $A_1, A_2, \dots, A_n, \emptyset, \emptyset, \dots$ is pairwise disjoint, and so, by σ-additivity,

$$\begin{aligned}\mu(\cup_{i=1}^{n} A_i) &= \mu(\cup_{i\in\mathbb{N}} A_i) = \sum_{i\in\mathbb{N}} \mu(A_i) \\ &= \sum_{i=1}^{n} \mu(A_i) + \sum_{i=n+1}^{\infty} \mu(\emptyset) = \sum_{i=1}^{n} \mu(A_i).\end{aligned}$$

2. Write B as a disjoint union $A \cup (B \setminus A)$ of elements of $\mathcal{S}$. By (1) above, $\mu(B) = \mu(A) + \mu(B \setminus A) \geq \mu(A)$.

3. Define $B_1 = A_1$, $B_n = A_n \setminus A_{n-1}$ for $n > 1$. Then $\{B_n\}_{n\in\mathbb{N}}$ is a pairwise disjoint sequence in $\mathcal{S}$ with $\cup_{k=1}^{n} B_k = A_n$ for each $n \in \mathbb{N}$ (why?). By σ-additivity we have

$$\begin{aligned}\mu(\cup_n A_n) &= \mu(\cup_n B_n) = \sum_{n\in\mathbb{N}} \mu(B_n) = \lim_n \sum_{k=1}^{n} \mu(B_k) \\ &= \lim_n \mu(\cup_{k=1}^{n} B_k) = \lim_n \mu(A_n).\end{aligned}$$

4. Consider the increasing sequence $\{B_n\}_{n\in\mathbb{N}}$ in $\mathcal{S}$ given by $B_n = A_1 \setminus A_n$. By De Morgan laws, finiteness of $\mu(A_1)$ and (3) above, we have

$$\begin{aligned}\mu(A_1) - \mu(\cap_n A_n) &= \mu(A_1 \setminus (\cap_n A_n)) = \mu(\cup_n B_n) = \lim_n \mu(B_n) \\ &= \lim_n \mu(A_1 \setminus A_n) = \mu(A_1) - \lim_n \mu(A_n).\end{aligned}$$

   Subtracting both sides from $\mu(A_1) < \infty$ produces the statement.

5. We start from the observation that for $A_1, A_1 \in \mathcal{S}$ the set $A_1 \cup A_2$ can be written as a disjoint union

$$A_1 \cup A_2 = (A_1 \setminus A_2) \cup (A_2 \setminus A_1) \cup (A_1 \cap A_2),$$

   so that

$$\mu(A_1 \cup A_2) = \mu(A_1 \setminus A_2) + \mu(A_2 \setminus A_1) + \mu(A_1 \cap A_2).$$

   On the other hand,

$$\begin{aligned}\mu(A_1) + \mu(A_2) &= (\mu(A_1 \setminus A_2) + \mu(A_1 \cap A_2)) \\ &\quad + \Big(\mu(A_2 \setminus A_1) + \mu(A_1 \cap A_2)\Big) \\ &= \mu(A_1 \setminus A_2) + \mu(A_2 \setminus A_1) + 2\mu(A_1 \cap A_2),\end{aligned}$$

   and so

$$\mu(A_1) + \mu(A_2) - \mu(A_1 \cup A_2) = \mu(A_1 \cap A_2) \geq 0.$$

   Induction can be used to show that

$$\mu(A_1 \cup \dots \cup A_n) \leq \sum_{k=1}^{n} \mu(A_k).$$

   Since all $\mu(A_n)$ are nonnegative, we now have

$$\mu(A_1 \cup \dots \cup A_n) \leq \alpha, \text{ for each } n \in \mathbb{N}, \text{ where } \alpha = \sum_{n\in\mathbb{N}} \mu(A_n).$$

   The sequence $\{B_n\}_{n\in\mathbb{N}}$ given by $B_n = \cup_{k=1}^n A_k$ is increasing, so the continuity of measure with respect to increasing sequences implies that

$$\mu(\cup_n A_n) = \mu(\cup_n B_n) = \lim_n \mu(B_n) = \lim_n \mu(A_1 \cup \dots \cup A_n) \leq \alpha. \quad \square$$

Remark 2.7. The condition $\mu(A_1) < \infty$ in the part (4) of Proposition 2.6 cannot be significantly relaxed. Indeed, let μ be the counting measure on $\mathbb{N}$, and let $A_n = \{n, n+1, \dots\}$. Then $\mu(A_n) = \infty$ and, so $\lim_n \mu(A_n) = \infty$. On the other hand, $\cap A_n = \emptyset$, so $\mu(\cap_n A_n) = 0$.

In addition to unions and intersections, one can produce other important new sets from sequences of old ones. More specifically, let $\{A_n\}_{n\in\mathbb{N}}$ be a sequence of subsets of S. The subset $\liminf_n A_n$ of S, defined by

$$\liminf_n A_n = \cup_n B_n, \text{ where } B_n = \cap_{k\geq n} A_k,$$

is called the **limit inferior** of the sequence A_n. It is also denoted by $\underline{\lim}_n A_n$ or $\{A_n, \text{ ev.}\}$ (*ev.* stands for *eventually*[2]).

Similarly, the subset $\limsup_n A_n$ of S, defined by

$$\limsup_n A_n = \cap_n B_n, \text{ where } B_n = \cup_{k\geq n} A_k,$$

is called the **limit superior** of the sequence A_n. It is also denoted by $\overline{\lim}_n A_n$ or $\{A_n, \text{ i.o.}\}$ (*i.o.* stands for *infinitely often*[3]). Clearly, we have

$$\liminf_n A_n \subseteq \limsup_n A_n.$$

[2] the reason for the use of the word *eventually* is the following: $\liminf_n A_n$ is the set of all $x \in S$ which belong to A_n for *all but finitely many values* of the index n, i.e., from some value of the index n onwards.

[3] in words, $\limsup_n A_n$ is the set of all $x \in S$ which belong A_n for *infinitely many values* of n.

Problem 2.1. Let $(S, \mathcal{S}, \mu)$ be a *finite* measure space. Show that

$$\mu(\liminf A_n) \leq \liminf \mu(A_n) \leq \limsup \mu(A_n) \leq \mu(\limsup A_n),$$

for any sequence $\{A_n\}_{n\in\mathbb{N}}$ in $\mathcal{S}$. Give an example of a (single) sequence $\{A_n\}_{n\in\mathbb{N}}$ for which all inequalities above are strict.

Hint: For the second part, a measure space with finite (and small) S will do.

Proposition 2.8 (Borel-Cantelli Lemma I)**.** *Let $(S, \mathcal{S}, \mu)$ be a measure space, and let $\{A_n\}_{n\in\mathbb{N}}$ be a sequence of sets in $\mathcal{S}$ with the property that $\sum_{n\in\mathbb{N}} \mu(A_n) < \infty$. Then*

$$\mu(\limsup_n A_n) = 0.$$

Proof. Set $B_n = \cup_{k\geq n} A_k$, so that $\{B_n\}_{n\in\mathbb{N}}$ is a decreasing sequence of sets in $\mathcal{S}$ with $\limsup_n A_n = \cap_n B_n$, and so

$$\mu(\limsup A_n) \leq \mu(B_n), \text{ for each } n \in \mathbb{N}.$$

Using the subadditivity of measures of Proposition 2.6, part 5., we get

$$\mu(B_n) \leq \sum_{k=n}^{\infty} \mu(A_n). \tag{2.1}$$

Since $\sum_{n\in\mathbb{N}} \mu(A_n)$ converges, the right-hand side of (2.1) can be made arbitrarily small by choosing large enough $n \in \mathbb{N}$. $\square$

Extensions of measures and the coin-toss space

Example 1.19 has introduced the measurable space $(\{-1,1\}^{\mathbb{N}}, \mathcal{S})$, with $\mathcal{S} = \otimes_n 2^{\{-1,1\}}$ being the product σ-algebra on $\{-1,1\}^{\mathbb{N}}$. The purpose of the present section is to turn $(\{-1,1\}^{\mathbb{N}}, \mathcal{S})$ into a measure space, i.e., to define a suitable measure on it. It is easy to construct just any measure on $\{-1,1\}^{\mathbb{N}}$, but the one we are after is the one which will justify the name *coin-toss space*.

The intuition we have about tossing a fair coin infinitely many times should help us start with the definition of the coin-toss measure - denoted by μ_C - on cylinders. Since the coordinate spaces $\{-1,1\}$ are particularly simple, each product cylinder is of the form $C = \{-1,1\}^{\mathbb{N}}$ or $C = C_{n_1,\dots,n_k;b_1,\dots,b_k}$, where

$$C_{n_1,\dots,n_k;b_1,\dots,b_k} = \Big\{ s = (s_1, s_2, \dots) \in \{-1,1\}^{\mathbb{N}} : s_{n_1} = b_1, \dots, s_{n_k} = b_k \Big\},$$

for some $k \in \mathbb{N}$, and a choice of $1 \leq n_1 < n_2 < \dots < n_k \in \mathbb{N}$ of coordinates and the corresponding values $b_1, b_2, \dots, b_k \in \{-1,1\}$.

In the language of elementary probability, each cylinder corresponds to the event when the outcome of the n_i-th coin is $b_i \in \{-1,1\}$, for $i = 1, \dots, n$. The measure (probability) of this event can only be given by

$$\mu_C(C_{n_1,\dots,n_k;b_1,\dots,b_k}) = \underbrace{\tfrac{1}{2} \times \tfrac{1}{2} \times \dots \times \tfrac{1}{2}}_{k \text{ times}} = 2^{-k}. \tag{2.2}$$

The hard part is to extend this definition to *all* elements of $\mathcal{S}$, and not only cylinders. For example, in order to state the law of large numbers later on, we will need to be able to compute the measure of the set

$$\Big\{ s \in \{-1,1\}^{\mathbb{N}} : \lim_n \tfrac{1}{n} \sum_{k=1}^n s_k = 0 \Big\},$$

which is clearly not a cylinder.

Problem 1.9 states, however, that cylinders form an algebra and generate the σ-algebra $\mathcal{S}$. Luckily, this puts us close to the conditions of the following important theorem of Caratheodory.

Theorem 2.9 (Caratheodory's Extension Theorem)**.** *Let S be a non-empty set, let $\mathcal{A}$ be an algebra of its subsets and let $\mu : \mathcal{A} \to [0,\infty]$ be a set-function with the following properties:*

1. $\mu(\emptyset) = 0$, *and*
2. $\mu(A) = \sum_{n=1}^\infty \mu(A_n)$, *if* $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{A}$ *is a partition of* A.

Then, there exists a measure $\tilde{\mu}$ on $\sigma(\mathcal{A})$ with the property that $\mu(A) = \tilde{\mu}(A)$ for $A \in \mathcal{A}$.

Note: In words, a σ-additive measure on an algebra $\mathcal{A}$ can be extended to a σ-additive measure on the σ-algebra generated by $\mathcal{A}$. It is clear that the σ-additivity requirement of Theorem 2.9 is necessary, but it is quite surprising that it is actually sufficient.

Of Theorem 2.9. **PART I.** We start by defining a "measure-like object", called an **outer measure**, $\mu^* : 2^S \to [0,\infty]$ in the following way:

$$\mu^*(B) = \inf \Big\{ \sum_{n=1}^\infty \mu(A_n) : B \subseteq \bigcup_{n=1}^\infty A_n,\ A_n \in \mathcal{A} \text{ for all } n \in \mathbb{N} \Big\}.$$

Note: Intuitively, we try all different countable covers of B with elements of $\mathcal{A}$ and minimize the total μ.

Even though we don't expect the infimum in the definition of μ^* to be attained, μ^* has the following properties:

1. $\mu^*(\emptyset) = 0$ *(nontriviality)*,

2. for $B \subseteq C$, $\mu^*(B) \leq \mu^*(C)$ *(monotonicity)*, and

3. $\mu^*(\cup_k B_k) \leq \sum_{k=1}^{\infty} \mu^*(B_k)$ *(subadditivity)*

Parts 1. and 2. are immediately clear, while, to show 3., we pick $\varepsilon > 0$ and $k \in \mathbb{N}$ and find a countable cover $\{A_n^k\}_{n\in\mathbb{N}}$ with elements of $\mathcal{A}$ such that

$$\mu^*(B_k) \geq \sum_{n=1}^{\infty} \mu(A_n^k) + \tfrac{\varepsilon}{2^k}.$$

Using $\{A_n^k\}_{k\in\mathbb{N},n\in\mathbb{N}}$ as a candidate cover for $\cup_k B_k$, we conclude that $\mu^*(\cup_k B_k) \leq \sum_{k=1}^{\infty} \mu^*(B_k) + \varepsilon$. This being true for each $\varepsilon > 0$ implies 3.

We remark at this point that μ^* and μ coincide on $\mathcal{A}$. By using $(A, \emptyset, \emptyset, \dots)$ as a candidate countable cover of $A \in \mathcal{A}$, we can conclude that $\mu^*(A) \leq \mu(A)$, for all $A \in \mathcal{A}$. Conversely, suppose that $A \subseteq \cup_{n\in\mathbb{N}} A_n$ with $A_n \in \mathcal{A}$. Given that the elements of $\mathcal{A}$ form an algebra, we can assume that $A_n \subseteq A$, for all $n \in \mathbb{N}$ and that $\{A_n\}_{n\in\mathbb{N}}$ are pairwise disjoint, as any sequence covering A can be transformed into such a sequence without increasing $\sum_n \mu(A_n)$. The assumed countable additivity of μ on $\mathcal{A}$ now comes into play since, for partition $\{A_n\}_{n\in\mathbb{N}}$ of A into elements in $\mathcal{A}$, we necessarily have $\sum_n \mu(A_n) = \mu(A)$, and, so $\mu^*(A) \geq \mu(A)$.

Note: It is, probably, interesting to note that this is the only place in the entire proof where the countable additivity of μ on $\mathcal{A}$ is used.

PART II. The set-function μ^* is, in general, not a measure, but comes with the advantage of being defined on *all* subsets of S. The central idea of the proof is to recover countable additivity by restricting its domain a little. We say that a subset $M \subseteq S$ is **Caratheodory-measurable** or **μ^*-measurable** if

$$\mu^*(B) = \mu^*(B \cap M) + \mu^*(B \cap M^c) \text{ for all } B \subseteq S, \tag{2.3}$$

with the family of all μ^*-measurable subsets of S denoted by $\mathcal{M}^*$. We note that, by subadditivity, the equality sign in the definition of the measurability can be replaced by $\geq$; this will be used below.

The first thing we need to establish about $\mathcal{M}^*$ is that it is an algebra and that μ^* is a finitely-additive measure on $\mathcal{M}^*$. Clearly $\emptyset \in \mathcal{A}^*$ and the complement axiom follows directly from the symmetry in (2.3). Only the closure under finite unions needs some discussion, and, by induction, we only need to consider two-element unions; for that, we pick $M, N \in \mathcal{M}^*$, and introduce the following notation

$$M_{00} = M^c \cap N^c,\ M_{01} = M^c \cap N,\ M_{10} = M \cap N^c\ M_{11} = M \cap N. \tag{2.4}$$

By the measurability of M and N, for any $B \subseteq S$, we have

$$\begin{aligned}\mu^*(B) &= \mu^*(B \cap N^c) + \mu^*(B \cap N) \\ &= \mu^*(M_{00} \cap B) + \mu^*(M_{10} \cap B) + \mu^*(M_{01} \cap B) + \mu^*(M_{11} \cap B)\end{aligned}$$

On the other hand, $M_{01} \cup M_{10} \cup M_{11} = M \cup N$, so that, by subadditivity and (2.4), we have

$$\begin{aligned}&\mu^*\big((M \cup N)^c \cap B\big) + \mu^*\big((M \cup N) \cap B\big) \\ &\quad = \mu^*(M_{00} \cap B) + \mu^*\Big((M_{01} \cap B) \cup (M_{10} \cap B) \cup (M_{11} \cap B)\Big) \\ &\quad \leq \mu^*(M_{00} \cap B) + \mu^*(M_{10} \cap B) + \mu^*(M_{01} \cap B) + \mu^*(M_{11} \cap B) \\ &\quad = \mu^*(B),\end{aligned}$$

which implies that that $M \cup N \in \mathcal{M}^*$. When $M \cap N = \emptyset$ an application of measurability of N to $B = M \cup N$ yields the finite additivity of μ^* on $\mathcal{M}^*$:

$$\mu^*(M \cup N) = \mu^*((M \cup N) \cap N) + \mu^*((M \cup N) \cap N^c) = \mu^*(N) + \mu^*(M).$$

PART III. We now turn to the closure of $\mathcal{M}^*$ under countable unions and the σ-additive property of μ^*. Since $\mathcal{M}^*$ already known to be an algebra, it will be enough to show that it is closed under pairwise-disjoint unions, i.e., that $M \in \mathcal{M}^*$ whenever $\{M_n\}_{n \in \mathbb{N}}$ are pairwise disjoint elements in $\mathcal{M}^*$ with $M = \cup_n M_n$.

For $n \in \mathbb{N}$, we set $L_n = \cup_{k=1}^n M_k$ so that, for $B \subseteq S$, we have

$$\begin{aligned}\mu^*(B) &= \mu^*(B \cap L_n) + \mu^*(B \cap L_n^c) \\ &\geq \textstyle\sum_{k=1}^n \mu^*(B \cap M_k) + \mu^*(B \cap M^c),\end{aligned}$$

so that

$$\begin{aligned}\mu^*(B) &\geq \mu^*(B \cap M^c) + \textstyle\sum_{k \in \mathbb{N}} \mu^*(B \cap M_k) \\ &\geq \mu^*(B \cap M^c) + \mu^*(\cup_k (B \cap M_k)) + \mu^*(B \cap M^c) \\ &= \mu^*(B \cap M) + \mu^*(B \cap M^c).\end{aligned}$$

Since all the inequalities above need to be equalities, we immediately conclude that, with $B = S$,

$$\mu^*(M) = \sum_k \mu^*(M_k),$$

i.e., that μ^* is a countably-additive measure on $\mathcal{M}^*$. Since $\mathcal{A} \subseteq \mathcal{M}^*$, we have $\mathcal{M}^* \supseteq \sigma(\mathcal{A})$ and $\tilde{\mu} = \mu^*|_{\sigma(\mathcal{A})}$ is the required σ-additive extension of μ. □

Back to the coin-toss space. In order to apply Theorem 2.9 in our situation, we need to check that μ is indeed a countably-additive measure on the algebra $\mathcal{A}$ of all cylinders. The following problem will help pinpoint the hard part of the argument:

Problem 2.2. Let $\mathcal{A}$ be an algebra on a non-empty set S, and let $\mu : \mathcal{A} \to [0, \infty]$ be a finite ($\mu(S) < \infty$) and finitely-additive set function on $\mathcal{S}$ with the following, additional, property:

$$\lim_n \mu(A_n) = 0, \text{ whenever } A_n \searrow \emptyset. \tag{2.5}$$

Then μ satisfies the conditions of Theorem 2.9.

The part about finite additivity is easy (perhaps a bit messy) and we leave it to the reader:

Problem 2.3. Show that the set-function μ_C, defined by (2.2) on the product cyllinders and extended by additivity to the algebra $\mathcal{A}$ of cylinders, is finitely additive.

Lemma 2.10 (Conditions of Caratheodory's theorem)**.** *The set-function μ_C, defined by* (2.2)*, and extenede by additivity to the the algebra $\mathcal{A}$ of cylinders, has the property* (2.5)*.*

Proof. By Problem 1.10, cylinders are closed sets, and so $\{A_n\}_{n \in \mathbb{N}}$ is a sequence of closed sets whose intersection is empty. The same problem states that $\{-1, 1\}^{\mathbb{N}}$ is compact, so, by the finite-intersection property[4], we have $A_{n_1} \cap \dots A_{n_k} = \emptyset$, for some finite collection $n_1, \dots, n_k$ of indices. Since $\{A_n\}_{n \in \mathbb{N}}$ is decreasing, we must have $A_n = \emptyset$, for all $n \geq n_k$, and, consequently, $\lim_n \mu(A_n) = 0$. □

[4] The *finite-intersection property* refers to the following fact, familiar from real analysis: If a family of closed sets of a compact topological space has empty intersection, then it admits a *finite* subfamily with an empty intersection.

Proposition 2.11 (Existence of the coin-toss measure)**.** *There exists a measure μ_C on $(\{-1, 1\}^{\mathbb{N}}, \mathcal{S})$ with the property that (2.2) holds for all cylinders.*

Proof. Thanks to Lemma 2.10, Theorem 2.9 can now be used. □

In order to prove uniqueness, we will need the celebrated π-λ Theorem of Eugene Dynkin:

Theorem 2.12 (Dynkin's "π-λ" Theorem)**.** *Let $\mathcal{P}$ be a π-system on a non-empty set S, and let Λ be a λ-system which contains $\mathcal{P}$. Then Λ also contains the σ-algebra $\sigma(\mathcal{P})$ generated by $\mathcal{P}$.*

Proof. Using the result of part 4. of Problem 1.1, we only need to prove that $\lambda(\mathcal{P})$ (where $\lambda(\mathcal{P})$ denotes the λ-system generated by $\mathcal{P}$) is a π-system. For $A \subseteq S$, let $\mathcal{G}_A$ denote the family of all subsets of S whose intersections with A are in $\lambda(\mathcal{P})$:

$$\mathcal{G}_A = \{C \subseteq S : C \cap A \in \lambda(\mathcal{P})\}.$$

Claim: $\mathcal{G}_A$ is a λ-system for $A \in \lambda(\mathcal{P})$.

- Since $A \in \lambda(\mathcal{P})$, clearly $S \in \mathcal{G}_A$.
- For an increasing family $\{C_n\}_{n\in\mathbb{N}}$ in $\mathcal{G}_A$ we have $(\cup_n C_n) \cap A = \cup_n(C_n \cap A)$. Each $C_n \cap A$ is in Λ, and the family $\{C_n \cap A\}_{n\in\mathbb{N}}$ is increasing, so $(\cup_n C_n) \cap A \in \Lambda$.
- Finally, for $C_1, C_2 \in \mathcal{G}$ with $C_1 \subseteq C_2$, we have

  $$(C_2 \setminus C_1) \cap A = (C_2 \cap A) \setminus (C_1 \cap A) \in \Lambda,$$

  because $C_1 \cap A \subseteq C_2 \cap A$.

$\mathcal{P}$ is a π-system, so for any $A \in \mathcal{P}$, we have $\mathcal{P} \subseteq \mathcal{G}_A$. Therefore, $\lambda(\mathcal{P}) \subseteq \mathcal{G}_A$, because $\mathcal{G}_A$ is a λ-system. In other words, for $A \in \mathcal{P}$ and $B \in \lambda(\mathcal{P})$, we have $A \cap B \in \lambda(\mathcal{P})$.

That means, however, that $\mathcal{P} \subseteq \mathcal{G}_B$, for any $B \in \lambda(\mathcal{P})$. Using the fact that $\mathcal{G}_B$ is a λ-system we must also have $\lambda(\mathcal{P}) \subseteq \mathcal{G}_B$, for any $B \in \lambda(\mathcal{P})$, i.e., $A \cap B \in \lambda(\mathcal{P})$, for all $A, B \in \lambda(\mathcal{P})$, which shows that $\lambda(\mathcal{P})$ is π-system. □

Proposition 2.13 (Measures which agree on a π-system)**.** *Let $(S, \mathcal{S})$ be a measurable space, and let $\mathcal{P}$ be a π-system which generates $\mathcal{S}$. Suppose that μ_1 and μ_2 are two measures on $\mathcal{S}$ with the property that $\mu_1(S) = \mu_2(S) < \infty$ and*

$$\mu_1(A) = \mu_2(A), \text{ for all } A \in \mathcal{P}.$$

Then $\mu_1 = \mu_2$, i.e., $\mu_1(A) = \mu_2(A)$, for all $A \in \mathcal{S}$.

Proof. Let $\mathcal{L}$ be the family of all subsets A of $\mathcal{S}$ for which $\mu_1(A) = \mu_2(A)$. Clearly $\mathcal{P} \subseteq \mathcal{L}$, but $\mathcal{L}$ is, potentially, bigger. In fact, it follows easily from the elementary properties of measures (see Proposition 2.6) and the fact that $\mu_1(S) = \mu_2(S) < \infty$ that it necessarily has the structure of a λ-system[5]. By Theorem 2.12 (the π-λ Theorem), $\mathcal{L}$ contains the σ-algebra generated by $\mathcal{P}$, i.e., $\mathcal{S} \subseteq \mathcal{L}$. On the other hand, by definition, $\mathcal{L} \subseteq \mathcal{S}$ and so $\mu_1 = \mu_2$. □

[5] It seems that the structure of a λ-system is defined so that it would exactly describe the structure of the family of all sets on which two measures (with the same total mass) agree. The structure of the π-system corresponds to the minimal assumption that allows Proposition 2.13 to hold.

Proposition 2.14 (Uniqueness of the coin-toss measure)**.** *The measure μ_C is the unique measure on $(\{-1,1\}^{\mathbb{N}}, \mathcal{S})$ with the property that (2.2) holds for all cylinders.*

Proof. The existence is the content of Proposition 2.11. To prove uniqueness, it suffices to note that algebras are π-systems and use Proposition 2.13. □

Problem 2.4. Define $D_1, D_2 \subseteq \{-1,1\}^{\mathbb{N}}$ by

1. $D_1 = \{s \in \{-1,1\}^{\mathbb{N}} : \limsup_n s_n = 1\}$,
2. $D_2 = \{s \in \{-1,1\}^{\mathbb{N}} : \exists\, N \in \mathbb{N},\, s_N = s_{N+1} = s_{N+2}\}$.

Show that $D_1, D_2 \in \mathcal{S}$ and compute $\mu(D_1), \mu(D_2)$.

Our next task is to probe the structure of the σ-algebra $\mathcal{S}$ on $\{-1,1\}^{\mathbb{N}}$ a little bit more and show that $\mathcal{S} \neq 2^{\{-1,1\}^{\mathbb{N}}}$. It is interesting that such a result (which deals exclusively with the structure of $\mathcal{S}$) requires a use of a measure in its proof.

Example 2.15 (A non-measurable subset of $\{-1,1\}^{\mathbb{N}}$ (*))**.** Since σ-algebras are closed under countable set operations, and since the product σ-algebra $\mathcal{S}$ for the coin-toss space $\{-1,1\}^{\mathbb{N}}$ is generated by sets obtained by restricting finite collections of coordinates, one is tempted to think that $\mathcal{S}$ contains *all* subsets of $\{-1,1\}^{\mathbb{N}}$. That is not the case. We will use the axiom of choice, together with the fact that a measure μ_C can be defined on the whole of $\{-1,1\}^{\mathbb{N}}$, to show to "construct" an example of a non-measurable set.

Let us start by constructing a relation $\sim$ on $\{-1,1\}^{\mathbb{N}}$ in the following[6] way: we set $s^1 \sim s^2$ if and only if there exists $n \in \mathbb{N}$ such that $s^1_k = s^2_k$, for $k \geq n$ (here, as always, $s^i = (s^i_1, s^i_2, \dots)$, $i = 1,2$). It is easy to check that $\sim$ is an equivalence relation and that it splits $\{-1,1\}^{\mathbb{N}}$ into disjoint equivalence classes. One of the many equivalent forms of the axiom of choice states that there exists a subset N of $\{-1,1\}^{\mathbb{N}}$ which contains exactly one element from each of the equivalence classes.

[6] In words, s^1 and s^2 are related if they only differ in a finite number of coordinates.

Let us suppose that N is an element in $\mathcal{S}$ and see if we can reach a contradiction. For each nonempty $\boldsymbol{n} = \{n_1, \dots, n_k\} \in 2^{\mathbb{N}}_{fin}$, where $2^{\mathbb{N}}_{fin}$ denotes the family of all finite subsets of $\mathbb{N}$, let us define the mapping $T_{\boldsymbol{n}} : \{-1,1\}^{\mathbb{N}} \to \{-1,1\}^{\mathbb{N}}$ in the following[7] manner:

[7] $T_{\boldsymbol{n}}$ flips the signs of the elements of its argument on the positions corresponding to $\boldsymbol{n}$.

$$T_{\emptyset} = \mathrm{Id} \text{ and } (T_{\boldsymbol{n}}(s))_l = \begin{cases} -s_l, & l \in \boldsymbol{n}, \\ s_l, & l \notin \boldsymbol{n}, \end{cases} \text{ for } \boldsymbol{n} \in \mathbb{N}.$$

Since $\boldsymbol{n}$ is finite, $T_{\boldsymbol{n}}$ preserves the $\sim$-equivalence class of each element. Consequently (and using the fact that N contains exactly one element from each equivalence class) the sets N and $T_{\boldsymbol{n}}(N) = \{T_{\boldsymbol{n}}(s) : s \in N\}$ are disjoint. Similarly and more generally, the sets $T_{\boldsymbol{n}}(N)$ and $T_{\boldsymbol{n}'}(N)$ are also disjoint whenever $\boldsymbol{n} \neq \boldsymbol{n}'$. On the other hand, each $s \in \{-1,1\}^{\mathbb{N}}$ is equivalent to some $\hat{s} \in N$, i.e., it can be obtained from $\hat{s}$ by flipping a finite number of coordinates. Therefore, the family

$$\mathcal{N} = \{T_{\boldsymbol{n}}(N) : \boldsymbol{n} \in 2^{\mathbb{N}}_{fin}\}$$

forms a partition of $\{-1,1\}^{\mathbb{N}}$.

The mapping $T_{\boldsymbol{n}}$ has several other nice properties. First of all, it is immediate that it is involutory, i.e., $T_{\boldsymbol{n}} \circ T_{\boldsymbol{n}} = \mathrm{Id}$. To show that it is $(\mathcal{S}, \mathcal{S})$-measurable, we need to prove that its composition with each projection map $\pi_k : \mathcal{S} \to \{-1, 1\}$ is measurable. This follows immediately from the fact that for $k \in \mathbb{N}$

$$(\pi_k \circ T_{\boldsymbol{n}})^{-1}(\{1\}) = \begin{cases} C_{k;1}, & k \notin \boldsymbol{n}, \\ C_{k;-1}, & k \in \boldsymbol{n}, \end{cases}$$

where, for $b \in \{-1, 1\}$, we recall that $C_{k;b} = \{\boldsymbol{s} \in \{-1,1\}^{\mathbb{N}} : s_k = b\}$ is a product cylinder. If we combine the involutivity and measurability of $T_{\boldsymbol{n}}$, we immediately conclude that $T_{\boldsymbol{n}}(A) \in \mathcal{S}$ for each $A \in \mathcal{S}$. In particular, $\mathcal{N} \subseteq \mathcal{S}$.

In addition to preserving measurability, the map $T_{\boldsymbol{n}}$ also preserves the measure[8] the in μ_C, i.e., $\mu_C(T_{\boldsymbol{n}}(A)) = \mu_C(A)$, for all $A \in \mathcal{S}$. To prove that, let us pick $\boldsymbol{n} \in F$ and consider the set-function $\mu_{\boldsymbol{n}} : \mathcal{S} \to [0, 1]$ given by

$$\mu_{\boldsymbol{n}}(A) = \mu_C(T_{\boldsymbol{n}}(A)).$$

It is a simple matter to show that $\mu_{\boldsymbol{n}}$ is, in fact, a measure on $(S, \mathcal{S})$ with $\mu_{\boldsymbol{n}}(S) = 1$. Moreover, thanks to the simple form (2.2) of the action of the measure μ_C on cylinders, it is clear that $\mu_{\boldsymbol{n}} = \mu_C$ on the algebra of all cylinders. It suffices to invoke Proposition 2.13 to conclude that $\mu_{\boldsymbol{n}} = \mu_C$ on the entire $\mathcal{S}$, i.e., that $T_{\boldsymbol{n}}$ preserves μ_C.

The above properties of the maps $T_{\boldsymbol{n}}$, $\boldsymbol{n} \in F$ can imply the following: $\mathcal{N}$ is a partition of $\mathcal{S}$ into countably many measurable subsets of equal measure. Such a partition $\{N_1, N_2, \dots\}$ cannot exist, however. Indeed if it did, one of the following two cases would occur:

1. $\mu(N_1) = 0$. In that case
$$\mu(S) = \mu(\cup_k N_k) = \sum_n \mu(N_k) = \sum_n 0 = 0 \neq 1 = \mu(S).$$
2. $\mu(N_1) = \alpha > 0$. In that case
$$\mu(S) = \mu(\cup_k N_k) = \sum_n \mu(N_k) = \sum_n \alpha = \infty \neq 1 = \mu(S).$$

Therefore, the set N cannot be measurable[9] in $\mathcal{S}$.

The Lebesgue measure

As we shall see, the coin-toss space can be used as a sort of a universal measure space in probability theory. We use it here to construct the Lebesgue measure on $[0, 1]$. We start with the notion somewhat dual to the already introduced notion of the pull-back in Definition 1.8. We leave it as an exercise for the reader to show that the set function $f_*\mu$ from Definition 2.16 is indeed a measure.

[8] Actually, we say that a map f from a measure space $(S, \mathcal{S}, \mu_S)$ to the measure space $(T, \mathcal{T}, \mu_T)$ is **measure preserving** if it is measurable and $\mu_S(f^{-1}(A)) = \mu_T(A)$, for all $A \in \mathcal{T}$. The involutivity of the map $T_{\boldsymbol{n}}$ implies that this general definition agrees with our usage in this example.

[9] Somewhat heavier set-theoretic machinery can be used to prove that most of the subsets of S are not in $\mathcal{S}$, in the sense that the cardinality of the set $\mathcal{S}$ is strictly smaller than the cardinality of the set 2^S of all subsets of S

Definition 2.16 (Push-forwards). Let $(S, \mathcal{S}, \mu)$ be a measure space and let $(T, \mathcal{T})$ be a measurable space. The measure $f_*\mu$ on $(T, \mathcal{T})$, defined by

$$f_*\mu(B) = \mu(f^{-1}(B)), \text{ for } B \in \mathcal{T},$$

is called the **push-forward** of the measure μ by f.

Let $f : \{-1, 1\}^{\mathbb{N}} \to [0, 1]$ be the mapping given by

$$f(\boldsymbol{s}) = \sum_{k=1}^{\infty} \left(\tfrac{1+s_k}{2}\right) 2^{-k}, \ \boldsymbol{s} \in \{-1, 1\}^{\mathbb{N}}.$$

The idea is to use f to establish a correspondence between all real numbers in $[0, 1]$ and their expansions in the binary system, with the coding $-1 \mapsto 0$ and $1 \mapsto 1$. It is interesting to note that f is not one-to-one[10] as it, for example, maps $\boldsymbol{s}_1 = (1, -1, -1, \dots)$ and $\boldsymbol{s}_2 = (-1, 1, 1, \dots)$ into the same value - namely $\frac{1}{2}$. Let us show, first, that the map f is continuous in the metric d defined by part (1.2) of Problem 1.9. Indeed, we pick $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ in $\{-1, 1\}^{\mathbb{N}}$ and remember that for $d(\boldsymbol{s}_1, \boldsymbol{s}_2) \leq 2^{-n}$, the first $n-1$ coordinates of $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ coincide. Therefore,

$$|f(\boldsymbol{s}_1) - f(\boldsymbol{s}_2)| \leq \sum_{k=n}^{\infty} 2^{-k} = 2^{-n+1} = 2d(\boldsymbol{s}_1, \boldsymbol{s}_2).$$

Hence, the map f is Lipschitz and, therefore, continuous.

The continuity of f (together with the fact that $\mathcal{S}$ is the Borel σ-algebra for the topology induced by the metric d) implies that $f : (\{-1, 1\}^{\mathbb{N}}, \mathcal{S}) \to ([0, 1], \mathcal{B}([0, 1]))$ is a measurable mapping. Therefore, the push-forward $\lambda = f_*(\mu)$ is well defined on $([0, 1], \mathcal{B}([0, 1]))$, and we call it the **Lebesgue measure** on $[0, 1]$.

[10] The reason for this is, poetically speaking, that $[0, 1]$ is not the Cantor set.

Proposition 2.17 (Intuitive properties of the Lebesgue measure). *The Lebesgue measure λ on $([0, 1], \mathcal{B}([0, 1]))$ satisfies*

$$\lambda([a, b)) = b - a, \ \lambda(\{a\}) = 0 \textit{ for } 0 \leq a < b \leq 1. \tag{2.6}$$

Proof.

1. Consider a, b of the form $b = \frac{k}{2^n}$ and $b = \frac{k+1}{2^n}$, for $n \in \mathbb{N}$ and $k < 2^n$. For such a, b we have $f^{-1}([a, b)) = C_{1,\dots,n;c_1,c_2,\dots,c_n}$, where $\overline{c_1 c_2 \dots c_n}$ is the base-2 expansion of k (after the "recoding" $-1 \mapsto 0$, $1 \mapsto 1$). By the very definition of λ and the form (2.2) of the action of the coin-toss measure μ_C on cylinders, we have

$$\lambda\Big([a, b)\Big) = \mu_C\Big(f^{-1}([a, b))\Big) = \mu_C(C_{1,\dots,n;c_1,c_2,\dots,c_n}) = 2^{-n} = \tfrac{k+1}{2^n} - \tfrac{k}{2^n}.$$

   Therefore, (2.6) holds for a, b of the form $b = \frac{k}{2^n}$ and $b = \frac{l}{2^n}$, for $n \in \mathbb{N}$, $k < 2^n$ and $l = k + 1$. Using (finite) additivity of λ, we immediately conclude that (2.6) holds for all k, l, i.e., that it holds for all dyadic rationals. A general $a \in (0,1]$ can be approximated by an increasing sequence $\{q_n\}_{n\in\mathbb{N}}$ of dyadic rationals from the left, and the continuity of measures with respect to decreasing sequences implies that

$$\lambda\Big([a,p)\Big) = \lambda\Big(\cap_n\, [q_n,p)\Big) = \lim_n \lambda\Big([q_n,p)\Big) = \lim_n (p - q_n) = (p-a),$$

whenever $a \in (0,1]$ and p is a dyadic rational. In order to remove the dyadicity requirement from the right limit, we approximate it from the left by a sequence $\{p_n\}_{n\in\mathbb{N}}$ of dyadic rationals with $p_n > a$, and use the continuity with respect to increasing sequences to get, for $a < b \in (0,1)$,

$$\lambda\Big([a,b)\Big) = \lambda\Big(\cup_n\, [a,p_n)\Big) = \lim_n \lambda\Big([a,p_n)\Big) = \lim_n (p_n - a) = (b-a).$$

$\square$

The Lebesgue measure has another important property:

Problem 2.5. Show that the Lebesgue measure is **translation invariant**. More precisely, for $B \in \mathcal{B}([0,1])$ and $x \in [0,1)$, we have

Hint: Use Proposition 2.13 for the second part.

1. $B +_1 x = \{b + x \pmod 1 \,:\, b \in B\}$ is in $\mathcal{B}([0,1])$ and
2. $\lambda(B +_1 x) = \lambda(B)$,

where, for $a \in [0,2)$, we define

Geometrically, the set $x +_1 B$ is obtained from B by translating it to the right by x and then shifting the part that is "sticking out" by 1 to the left.

$$a \pmod 1 = \begin{cases} a, & a \leq 1, \\ a-1, & a > 1. \end{cases}$$

Finally, the notion of the Lebesgue measure is just as useful on the entire $\mathbb{R}$, as on its compact subset $[0,1]$. For a general $B \in \mathcal{B}(\mathbb{R})$, we can define the Lebesgue measure of B by measuring its intersections with all intervals of the form $[n, n+1)$, and adding them together, i.e.,

$$\lambda(B) = \sum_{n=-\infty}^{\infty} \lambda\Big(\big(B \cap [n,n+1)\big) - n\Big).$$

Note how we are overloading the notation and using the letter λ for both the Lebesgue measure on $[0,1]$ and the Lebesgue measure on $\mathbb{R}$.

It is a quite tedious, but does not require any new tools, to show that many of the properties of λ on $[0,1]$ transfer to λ on $\mathbb{R}$:

Problem 2.6. Let λ be the Lebesgue measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Show that

1. $\lambda([a,b)) = b - a,\ \lambda(\{a\}) = 0$ for $a < b$,

2. λ is σ-finite but not finite,

3. $\lambda(B + x) = \lambda(B)$, for all $B \in \mathcal{B}(\mathbb{R})$ and $x \in \mathbb{R}$, where $B + x = \{b + x : b \in B\}$.

Additional Problems

Problem 2.7 (Local separation by constants). Let $(S, \mathcal{S}, \mu)$ be a measure space and let $f, g \in \mathcal{L}^0(S, \mathcal{S}, \mu)$ satisfy $\mu\big(\{x \in S : f(x) < g(x)\}\big) > 0$. Prove or construct a counterexample for the following statement:

> "There exist constants $a, b \in \mathbb{R}$ such that
> $\mu\big(\{x \in S : f(x) \leq a < b \leq g(x)\}\big) > 0$."

Problem 2.8 (A pseudometric on sets). Let $(S, \mathcal{S}, \mu)$ be a finite measure space. For $A, B \in \mathcal{S}$ define

$$d(A, B) = \mu(A \vartriangle B),$$

where $\vartriangle$ denotes the symmetric difference: $A \vartriangle B = (A \setminus B) \cup (B \setminus A)$. Show that d is a pseudometric on $\mathcal{S}$, and for $A \in \mathcal{S}$ describe the set of all $B \in \mathcal{S}$ with $d(A, B) = 0$.

Note: Let X be a nonempty set. A function $d : X \times X \to [0, \infty)$ is called a **pseudometric** if

1. $d(x, y) + d(y, x) \geq d(x, z)$, for all $x, y, z \in X$,
2. $d(x, y) = d(y, x)$, for all $x, y \in X$, and
3. $d(x, x) = 0$, for all $x \in X$.

Note how the only difference between a metric and a pseudometric is that for a metric $d(x, y) = 0$ implies $x = y$, while no such requirement is imposed on a pseudometric.

Problem 2.9 (Complete measure spaces). A measure space $(S, \mathcal{S}, \mu)$ is called **complete** if all subsets of null sets are themselves in $\mathcal{S}$. For a (possibly incomplete) measure space $(S, \mathcal{S}, \mu)$ we define the **completion** $(S, \mathcal{S}^*, \mu^*)$ in the following way:

$$\mathcal{S}^* = \{A \cup N^* : A \in \mathcal{S} \text{ and } N^* \subseteq N \text{ for some } N \in \mathcal{S} \text{ with } \mu(N) = 0\}.$$

For $B \in \mathcal{S}^*$ with representation $B = A \cup N^*$ we set $\mu^*(B) = \mu(A)$.

1. Show that $\mathcal{S}^*$ is a σ-algebra.

2. Show that the definition $\mu^*(B) = \mu(A)$ above does not depend on the choice of the decomposition $B = A \cup N^*$, i.e., that $\mu(\hat{A}) = \mu(A)$ if $B = \hat{A} \cup \hat{N}^*$ is another decomposition of B into a set $\hat{A}$ in $\mathcal{S}$ and a subset $\hat{N}$ of a null set in $\mathcal{S}$.

3. Show that μ^* is a measure on $(S, \mathcal{S}^*)$ and that $(S, \mathcal{S}^*, \mu^*)$ is a complete measure space with the property that $\mu^*(A) = \mu(A)$, for $A \in \mathcal{S}$.

Note: Unfortunately, the same notation μ^* is often used for the completion of the measure μ and the outer measure associated with μ as in the proof of Theorem 2.9. Fortunately, it can be shown that these two object coincide on the domain of the completion.

Problem 2.10 (The Cantor set). The **Cantor set** is defined as the collection of all real numbers x in $[0, 1]$ with the representation

$$x = \sum_{n=1}^{\infty} c_n 3^{-n}, \text{ where } c_n \in \{0, 2\}.$$

Show that it is Borel-measurable and compute its Lebesgue measure.

Problem 2.11 (The uniform measure on a circle). Let S^1 be the unit circle, and let $f : [0,1) \to S^1$ be the "winding map"

$$f(x) = \Big(\cos(2\pi x), \sin(2\pi x)\Big), \ x \in [0,1).$$

1. Show that the map f is $(\mathcal{B}([0,1)), \mathcal{S}^1)$-measurable, where $\mathcal{S}^1$ denotes the Borel σ-algebra on S^1 (with the topology inherited from $\mathbb{R}^2$).

2. For $\alpha \in (0, 2\pi)$, let R_α denote the (counter-clockwise) rotation of $\mathbb{R}^2$ with center $(0,0)$ and angle α. Show that $R_\alpha(A) = \{R_\alpha(x) : x \in A\}$ is in $\mathcal{S}^1$ if and only if $A \in \mathcal{S}^1$.

3. Let μ^1 be the push-forward of the Lebesgue measure λ by the map f. Show that μ^1 is rotation-invariant, i.e., that $\mu^1(A) = \mu^1\big(R_\alpha(A)\big)$.

Note: The measure μ^1 is called the **uniform measure** (or the **uniform distribution**) on S^1.

Problem 2.12 (Asymptotic densities). We say that the subset A of $\mathbb{N}$ **admits asymptotic density** if the limit

$$d(A) = \lim_n \frac{\#(A \cap \{1, 2, \dots, n\})}{n},$$

exists (remember that # denotes the number of elements of a set). Let $\mathcal{D}$ be the collection of all subsets of $\mathbb{N}$ which admit asymptotic density.

1. Is $\mathcal{D}$ an algebra? A σ-algebra?

2. Is the map $A \mapsto d(A)$ finitely-additive on $\mathcal{D}$? A measure?

Problem 2.13 (A subset of the coin-toss space). An element in $\{-1,1\}^{\mathbb{N}}$ (i.e., a sequence $s = (s_1, s_2, \dots)$ where $s_n \in \{-1, 1\}$ for all $n \in \mathbb{N}$) is said to be **eventually periodic** if there exists $N_0, K \in \mathbb{N}$ such that $s_n = s_{n+K}$ for all $n \geq N_0$. Let $P \subseteq \{-1,1\}^{\mathbb{N}}$ be the collection of all eventually-period sequences. Show that P is measurable in the product σ-algebra $\mathcal{S}$ and compute $\mu_C(P)$.

Problem 2.14 (Regular measures). The measure space $(S, \mathcal{S}, \mu)$, where (S, d) is a metric space and $\mathcal{S}$ is a σ-algebra on S which contains the Borel σ-algebra $\mathcal{B}(d)$ on S is called **regular** if for each $A \in \mathcal{S}$ and each $\varepsilon > 0$ there exist a closed set C and an open set O such that $C \subseteq A \subseteq O$ and $\mu(O \setminus C) < \varepsilon$.

1. Suppose that $(S, \mathcal{S}, \mu)$ is a regular measure space, and that the measure space $(S, \mathcal{B}(d), \mu|_{\mathcal{B}(d)})$ is obtained from $(S, \mathcal{S}, \mu)$ by restricting the measure μ onto the σ-algebra of Borel sets. Show that $\mathcal{S} \subseteq \mathcal{B}(d)^*$, where $\big(S, \mathcal{B}(d)^*, (\mu|_{\mathcal{B}(d)})^*\big)$ is the completion (in the sense of Problem 2.9) of $(S, \mathcal{B}(d), \mu|_{\mathcal{B}(d)})$

2. Suppose that (S, d) is a metric space and that μ is a finite measure on $\mathcal{B}(d)$. Show that $(S, \mathcal{B}(d), \mu)$ is a regular measure space.

Hint: Consider a collection $\mathcal{A}$ of subsets A of S such that for each $\varepsilon > 0$ there exists a closed set C and an open set O with $C \subseteq A \subseteq O$ and $\mu(O \setminus C) < \varepsilon$. Argue that $\mathcal{A}$ is a σ-algebra. Then show that each closed set can be written as an intersection of open sets; use (but prove, first) the fact that the map

$$x \mapsto d(x, C) = \inf\{d(x, y) : y \in C\},$$

is continuous on S for any nonempty $C \subseteq S$.

3. Show that $(S, \mathcal{B}(d), \mu)$ is regular if μ is not necessarily finite, but has the property that $\mu(A) < \infty$ whenever $A \in \mathcal{B}(d)$ is bounded, i.e., when $\sup\{d(x, y) : x, y \in A\} < \infty$.

   Hint: Pick a point $x_0 \in S$ and, for $n \in \mathbb{N}$, define the family $\{R_n\}_{n \in \mathbb{N}}$ of subsets of S as follows:

   $$R_1 = \{x \in S : d(x, x_0) < 2\}, \text{ and}$$
   $$R_n = \{x \in S : n - 1 < d(x, x_0) < n + 1\}, \text{ for } n > 1,$$

   as well as a sequence $\{\mu_n\}_{n \in \mathbb{N}}$ of set functions on $\mathcal{B}(d)$, given by $\mu_n(A) = \mu(A \cap R_n)$, for $A \in \mathcal{B}(d)$. Under the right circumstances, even countable unions of closed sets are closed.

4. Conclude that the Lebesgue measure on $\left(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)\right)$ is regular.