1 Basics of measure theory

1.1 Introduction

A central theme of measure theory is the following question. How can we assign a (nonnegative) measure to subsets of some ground set Ω? In applications, the measure can have the meaning of size, content, mass, probability etc. As a function of set the measure should be *additive*. If Ω is finite or countable this is pretty straightforward, as a nonzero value $\mu(\omega)$ [(1)] can be assigned to every $\omega \in \Omega$ and then $\mu(A)$ defined for any $A \subset \Omega$ using the summation formula $\mu(A) = \sum_{\omega \in A} \mu(A)$. The problem is more involved if Ω is uncountable like $[0,1], \mathbb{R}^k$ or the infinite 'coin-tossing space' $\{0,1\}^\infty = \{0,1\} \times \{0,1\} \times \ldots$, or the space of continuous functions $C[0,1]$ etc. A fundamental example is the Lebesgue measure generalising the geometric notions of length, area and volume.

As a leading example today we shall consider the length λ defined on certain subsets in $\mathbb{R}$. The length $\lambda(I)$ of any interval $I = [a,b], (a,b], (a,b), [a,b)$ is $\lambda(I) = b-a$. For union of disjoint intervals $I_1, \ldots, I_n$ the length is

$$\lambda\left(\bigcup_{k=1}^n I_k\right) = \sum_{k=1}^n \lambda(I_k),$$

which is an instance of the property called *finite additivity*. For infinite sequence of disjoint intervals $I_1, I_2, \ldots$ the length of the union is the sum of series,

$$\lambda\left(\bigcup_{i=1}^\infty I_k\right) = \sum_{k=1}^\infty \lambda(I_k)$$

(infinite if the series diverges), which is an instance of the property called *σ-additivity*.

What is the length of the set $\mathbb{Q}$ of rational numbers? The length of a point is $\lambda(\{x\}) = 0$, and $\mathbb{Q}$ is countable, so by σ-additivity $\lambda(\mathbb{Q}) = 0$.

Now let us find the length of the *Cantor set* $C \subset [0,1]$. The Cantor set can be constructed step-by-step, at each stage obtaining some union of disjoint intervals C_k. Start with removing the middle third from $[0,1]$, thus defining $C_1 := [0,1/3] \cup [2/3,1]$. Then remove the middle third from $[0,1/3]$ and do the same with $[2/3,1]$, thus defining C_2. By induction, C_{k+1} is obtained by removing the middle third from every interval in C_k. The Cantor set is defined as the infinite intersection $C = \cap_{k=1}^\infty C_k$. Note that for $B \subset A$ we have $\lambda(B) \leq A$, because $A = B \cup (A \setminus B)$ is a disjoint union and $\lambda(A) = \lambda(B) + \lambda(B \setminus A)$. One can calculate the length of all removed intervals

$$\lambda([0,1] \setminus C) = \frac{1}{3} + \frac{1}{3}\left(1 - \frac{1}{3}\right) + \frac{1}{3}\left(\frac{2}{3}\right)^2 + \cdots = \frac{1}{3} \cdot \frac{1}{1-2/3} = 1$$

to see that $\lambda(C) = 1 - 1 = 0$. Another way to derive this is to show by induction that

$$\lambda(C_{k+1}) = \frac{2}{3}\lambda(C_k), \quad \text{hence } \lambda(C_k) = \left(\frac{2}{3}\right)^k,$$

and since $C \subset C_k$, we have

$$\lambda(C) \leq \lambda(C_k) = \left(\frac{2}{3}\right)^k, \quad k = 1, 2, \ldots$$

[(1)]This is a shorthand notation for $\mu(\{\omega\})$ in case of one-point sets.

and letting $k \to \infty$ yields $\lambda(C_k) \to 0$, so $\lambda(C) = 0$. The Cantor set is uncountable (has cardinality continuum, same as the cardinality of $[0, 1]$ or $\mathbb{R}$) and, as we have shown, has length 0.

How far can we go with ascribing the length to more complex sets $A \subset \mathbb{R}$? After the founder of measure theory Henri Lebesgue, the sets for which this can be done in a sensible way are called *Lebesgue measurable*, and the generalised length is called *the Lebesgue measure on* $\mathbb{R}$, to be discussed in the next section. Using the Axiom of Choice from the set theory it is possible to show existence of sets that are not Lebesque-mesaurable, but it is impossible to build them up from a system of intervals in some constructive manner .

A *probability measure* on Ω is a measure with $\mu(\Omega) = 1$. Subsets of Ω to which probability is assigned are called events, and notation $\mathbb{P}(A)$ will be used for probability of $A \subset \Omega$. For instance, the Lebesgue measure on $[0, 1]^k$ is a probability measure, used to model a point chosen uniformly at random from the cube.

In your probability courses you studied repeated Bernoulli trials (e.g. coin-tossing) with some success probability p. For infinite series of trials a suitable sample space to model possible outcomes is

$$\Omega = \{(\omega_1, \omega_2, \dots) : \omega_i = 0 \text{ or } 1, \text{for } i = 1, 2, \dots\} = \{0, 1\}^\infty,$$

so one outcome is an infinite sequence like $(0, 1, 1, 0, \dots)$. Identifying 1 with a 'head' the event A 'first two tosses are heads' is $A = \{\omega \in \Omega : \omega_1 = \omega_2 = 1\}$ with $\mathbb{P}(A) = p^2$. More complex events are required to formulate theorems of probability theory like the Law of Large Numbers, hence the same question arises: what is the reserve of events A to make sense of $\mathbb{P}(A)$?

1.2 Definition of measure

The idea is that a measure is an additive function of a set. Therefore the domain of definition of a measure should be a system of sets closed under the operations of taking union, and also intersection and complementation. In this context 'closed' means that applying operations $\cap, \cup, ^c$ to a countable selection of sets from the system will yield another set from the system. We write $A^c = \Omega \setminus A$ for the complement.

Definition 1.1. *A σ-algebra $\mathcal{F}$ on a set Ω is a family of subsets of Ω with the following properties:*

(i) $\Omega \in \mathcal{F}$,

(ii) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$,

(iii) $A_j \in \mathcal{F}, j \in \mathbb{N}, \Rightarrow \bigcup_{j=1}^\infty A_j \in \mathcal{F}$.

Conditions (i), (ii), (iii) is a minimal set of axioms defining σ-algebra. Using these other properties are derived. So $\varnothing \in \mathcal{F}$ by (i), (ii). Then $A_1, A_2 \in \mathcal{F} \Rightarrow A_1 \cup A_2 \in \mathcal{F}$ because we can set $A_j = \varnothing$ for $j \geq 2$ in (ii). Using complementation rules, $A_j \in \mathcal{F}, j \in \mathbb{N}, \Rightarrow \bigcap_{j=1}^\infty A_j \in \mathcal{F}$. And so on.

Note: operating with more than countably many sets from $\mathcal{F}$ may lead to outside of $\mathcal{F}$. Indeed, *every* $A \in \mathcal{P}(\Omega)$ from the power-set is a union of it individual points.

Typically, σ-algebras have too many sets to admit explicit description. However, with each collection of sets $\mathcal{S}$ we can associate a σ-algebra generated by $\mathcal{S}$, which we denote $\sigma(\mathcal{S})$. Observe that for σ-algebras $\mathcal{F}_1, \mathcal{F}_2$ also the intersection $\mathcal{F}_1 \cap \mathcal{F}_2$ is a σ-algebra. For any system of σ-algebras $(\mathcal{F}_j, j \in J)$ (possibly with uncountable index set J) also $\cap_{j \in J} \mathcal{F}_j$, is a σ-algebra. Therefore, we can specify any collection $\mathcal{S} \subset \mathcal{P}(\Omega)$ of *generators* and define $\sigma(\mathcal{S})$ to be the intersection of all σ-algebras that contain $\mathcal{S}$.

Examples

1. Consider $\mathcal{S} = \{\varnothing\}$. The generated σ-algebra is the smallest possible, $\{\varnothing, \Omega\}$.

2. Consider $\mathcal{S} = \{A_1, \dots, A_k\}$, where $A_1 \cup \cdots \cup A_k = \Omega$, A_j's are nonempty and pairwise disjoint. We speak in this situation of a partition of Ω with parts (blocks, etc) A_j. Every set in $\sigma(\mathcal{S})$ is obtained by selecting some of the A_j's and taking union, e.g. $A_2 \cup A_3 \cup A_7$ (provided $k \geq 7$). There are 2^k ways to select a subset from a set with k elements, therefore $\sigma(\mathcal{S})$ has 2^k elements.

3. Consider $\Omega = \{0,1\}^\infty$. For each k and $(\epsilon_1, \dots, \epsilon_k) \in \{0,1\}^k$ let $A(\epsilon_1, \dots, \epsilon_k) = \{\omega \in \Omega : \omega_1 = \epsilon_1, \dots, \omega_k = \epsilon_k\}$. Let $\mathcal{F}_k$ be generated by the partition with parts $A(\epsilon_1, \dots, \epsilon_k)$, where k is fixed; so the cardinality of $\mathcal{F}_k$ is 2^k. Observe that $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots$ is an increasing sequence of σ-algebras, we call such sequence *filtration*. In the coin-tossing interpretation, the event $A(1,0,1,1)$ occurs when the first outcomes are $1,0,1,1$. So $\mathcal{F}_k$ incorporates the information contained in the first k coin-tosses. As more trials are observed, we get more information.

   Now, let $\mathcal{F} = \sigma(\cup_{k=1}^\infty \mathcal{F}_k)$, which is the σ-algebra generated by *all* $A(\epsilon_1, \dots, \epsilon_k)$'s, that is with k and ϵ_j's freely chosen. Think of $\mathcal{F}$ as complete information gathered after infinitely many trials.

   This $\mathcal{F}$ is rich enough to state the laws of probability. For example, the event

$$A = \{\omega \in \Omega : \lim_{k \to \infty} (\omega_1 + \cdots + \omega_k)/k = 1/2\}$$

   is in $\mathcal{F}$, but does not belong to $\mathcal{F}_k$ for some k. Indeed, we can only compute the long-run frequency of heads as infinitely many coin tosses have been observed. If $p = 1/2$ (the coin is fair), then $\mathbb{P}(A) = 1/2$, but $\mathbb{P}(A) = 0$ for $p \neq 1/2$. Indeed, recall the Law of Large Numbers.

4. Define the *Borel σ-algebra* on $\mathbb{R}$, denoted $\mathcal{B}(\mathbb{R})$, as the σ-algebra generated by the set of semi-open intervals $\{(a,b] : -\infty < a < b \leq \infty]\}$. Elements of $\mathcal{B}(\mathbb{R})$ are called *Borel-measurable* or *Borel sets*. Borel σ-algebra $\mathcal{B}(\mathbb{R})$ is a universum of sets sufficient for all practical purposes.

   There are many other ways to select the set of generators: we can take for $\mathcal{S}$ all open sets, or all closed sets. A 'spare' collection of generators $\mathcal{S}$ for the Borel σ-algebra is the set of half-lines $\{(-\infty, x] : x \in \mathbb{R}\}$. This can be further reduced to the countable collection of half-lines $\{(-\infty, x] : x \in \mathbb{Q}\}$.

Definition 1.2. *Let $(\Omega, \mathcal{F})$ be a measurable space. A measure on Ω is a function $\mu : \mathcal{F} \in [0, \infty]$ such that $\mu(\varnothing) = 0$ and the σ-additivity property holds:*

$$\mu\left(\bigcup_{i=1}^\infty A_i\right) = \sum_{i=1}^\infty \mu(A_i), \tag{1}$$

for disjoint sets $A_i \in \mathcal{F}, i \in \mathbb{N}$. The triple $(\Omega, \mathcal{F}, \mu)$ is referred to as a measure space.

By the definition $\mu(A)$ is nonnegative, and the value ∞ is allowed. If $\mu(\Omega) < \infty$ we say that μ is a finite measure. If $\mu(\Omega) = 1$ we call μ probability measure, and often use notation $\mathbb{P}$. In the probability context we call measurable sets $A \in \mathcal{F}$ events, to which probability $\mathbb{P}(A)$ is assigned.

Examples For fixed $x \in \Omega$, *Dirac measure* is

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

Normally it is assumed that the one-point set $\{x\}$ is measurable ($\{x\} \in \mathcal{F}$), in that case x is called an atom. We sometimes say that the support of the Dirac measure is the atom x.

Choose $x_1, x_2, \dots$ from Ω and let $y_1, y_2, \dots$ be positive numbers. A *discrete* (aka *atomic*) measure is defined as

$$\mu(A) = \sum_{i=1}^\infty y_i \delta_{x_i}(A), \quad A \in \mathcal{F}.$$

Plainly, mass y_i sits in point x_i, so to compute the measure of set A you calculate the total mass of atoms in this set. If Ω is countable, e.g. $\Omega = \mathbb{N}$ then every measure on $(\Omega, \mathcal{P}(\Omega))$ is discrete.

In the last example we implicitly used the following simple fact: for measures $\mu_1, \mu_2, \dots$ on $(\Omega, \mathcal{F})$ and nonnegative reals $y_1, y_2, \dots,$ the linear combination $\sum_{i=1}^{\infty} y_i \mu_i$ is also a measure on $(\Omega, \mathcal{F})$.

There are useful properties implied σ-additivity. Let $A_i \in \mathcal{F}, i \in \mathbb{N}$.

1. Increasing tower of sets, monotonicity:

$$A_1 \subset A_2 \subset \dots \Rightarrow \mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{i\to\infty} \mu(A_i).$$

   To see this, apply the σ-additivity property to the union of disjoint sets $A_{i+1} \setminus A_i$. Note that $\mu(A_i)$ is nondecreasing in i in this case.

2. Decreasing tower of sets, monotonicity:

$$A_1 \supset A_2 \supset \dots \Rightarrow \mu\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{i\to\infty} \mu(A_i).$$

   This is obtained from the above increasing case by passing to complements.

3. Subadditivity:

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \le \sum_{i=1}^{\infty} \mu(A_i).$$

   The sets A_i can be arbitrary here.

1.3 Construction of measures by extension

Having introduced the general concept of measure, we wish to return to our principal example. We have the length $\lambda(A)$ defined for intervals and some other sets of relatively simple nature. Is it possible to have λ well defined for all Borel sets, consistently with the naive idea of length?

This is the fundamental problem of extension, which we shall treat in a general setting. A system of sets $\mathcal{A} \subset \mathcal{P}(\Omega)$ is called *algebra* if it satisfies conditions (i),(ii) from Definition 1.1, but (iii) is replaced by the finite additivity condition

$$A, B \in \mathcal{F} \Rightarrow A \cup B \in \mathcal{F}.$$

A function $\mu_0 : \mathcal{A} \to [0, \infty]$ is called a *pre-measure* if it satisfies (1) whenever $\cup_{i=1}^{\infty} A_i \in \mathcal{A}$.

The difference between pre-measure and measure is that a pre-measure is defined on algebra, which need not be closed under countable unions of sets.

These concepts are best seen on our main example, the set $\mathbb{R}$. Let $\mathcal{S}$ be the set of intervals $(a, b]$, this is a generator of the Borel σ-algebra. Let $\mathcal{A}$ be the collection of sets $A \subset \mathbb{R}$ representable as finite unions of disjoint intervals,

$$A = \bigcup_{i=1}^{k} (a_i, b_i],$$

one may check that $\mathcal{A}$ is an algebra. We have the length defined on $\mathcal{A}$ by the formula

$$\lambda(A) = \sum_{i=1}^{k} (b_i - a_i).$$

Note that a countable union of disjoint intervals *may* belong to $\mathcal{A}$, for example $(0, 1/2] \cup (1/2, 3/4] \cup (3/4, 7/8] \cup \cdots = (0, 1]$. The length λ (which is a pre-measure for a time being) is σ-additive on $\mathcal{A}$.

The next is the measure extension theorem due to Carathéodory.

Theorem 1.3. *Suppose μ_0 is a pre-measure on $(\Omega, \mathcal{A})$, where $\mathcal{A}$ is an algebra. Then there is a measure on $(\Omega, \sigma(\mathcal{A}))$ such that*

$$\mu(A) = \mu_0(A) \quad \text{for } A \in \mathcal{A}.$$

Moreover, this measure μ is unique if there exists a sequence of sets $B_1 \subset B_2 \dots$ such that $\cup_{j=1}^{\infty} B_j = \Omega$, $B_j \in \mathcal{A}$ and $\mu_0(B_j) < \infty$ for all $j \in \mathbb{N}$.

If $\cup_{j=1}^{\infty} B_j = \Omega$, for some $B_j \in \mathcal{F}, j \in \mathbb{N}$, such that $\mu(B_j) < \infty$ for all $j \in \mathbb{N}$, we call measure μ *σ-finite.* Carathéodory's Theorem entails that a σ-finite measure on $(\Omega, \mathcal{F})$ is uniquely determined by its values on some algebra of generators.

By Carathéodory's Theorem , the length λ defined initially on intervals has a unique extension to the Borel σ-algebra.

Example. Let us look how to define probability as a measure on $\Omega = \{0, 1\}^{\infty}$, to give a rigorous meaning to the notion of 'infinitely many independent Bernoulli trials with success probability p'.

Fix p and for each $A(\epsilon_1, \dots, \epsilon_k)$ set

$$\mathbb{P}(A(\epsilon_1, \dots, \epsilon_k)) = p^t (1-p)^{k-t}, \quad \text{where} \quad t = \epsilon_1 + \cdots + \epsilon_k. \tag{2}$$

The union $\mathcal{A} = \cup_{k=1}^{\infty} \mathcal{F}_k$ is an algebra, and $\mathbb{P}$ is a pre-measure on $(\Omega, \mathcal{A})$. By Carathéodory's theorem there is a probability measure consistent with (2) and defined on $\mathcal{F} = \sigma(\mathcal{A})$. This probability measure is unique because $\mathbb{P}(\Omega) = 1$ is finite. The Law of Large numbers says that the event

$$A = \{\omega \in \Omega : \lim_{k \to \infty} (\omega_1 + \cdots + \omega_k)/k = z\}$$

has probability $\mathbb{P}(A) = 1$ if $z = p$, and $\mathbb{P}(A) = 0$ if $z \neq p$.

Construction via distribution function. Many measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ can be defined in terms of generalised distribution functions. Let $F : \mathbb{R} \to (0, \infty)$ be a nondecreasing right-continuous function with left limits and $\lim_{x \to -\infty} F(x) = 0$. We define the measure of halfline $(-\infty, x]$ to be

$$\mu(-\infty, x] = F(x). \tag{3}$$

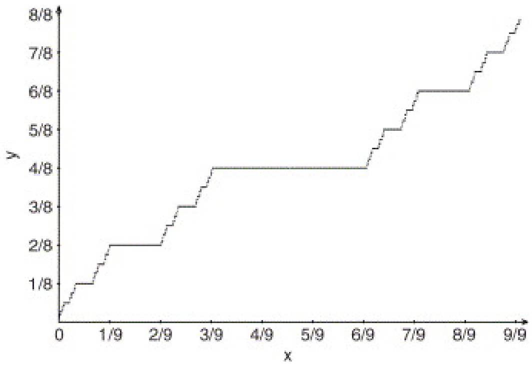
This is extended to intervals by $\mu(a, b] = F(b) - F(a)$ and is extendible to all Borel sets in a unique way by Carathéodory's theorem. If $\lim_{x \to \infty} F(x) = 1$ the measure μ is a probability measure, and F its cumulative distribution function. This method is very general, and allows one to construct both discrete distributions (e.g. supported by $\mathbb{N}$) and probability distributions with densities. The correspondence defined by (3) is invertible, in the sense that for every μ with $\mu(-\infty, x] < \infty, x \in \mathbb{R}$ the function F defined by this formula has the above properties (nondecreasing, etc).

If F has a jump at x, then the corresponding μ has an atom at x of mass $\mu(\{x\}) = F(x) - \lim_{k \to \infty} F(x - 1/k)$. If F has a density, in the sense that

$$F(x) = \int_{-\infty}^{x} f(z) dz \tag{4}$$

then the measure of each point $\{x\}$ is zero, in which case we say that the measure is non-atomic (or diffuse). Conversely, if F is continuous then the associated measure is non-atomic, but this does not mean that the measure has a density!

Cantor distribution function (see the picture) is an example of a probability measure which is non-atomic, but has no density to represent F as integral (4). Under this measure, the Cantor set has full probability $\mu(C) = 1$ although its Lebesgue measure is $\lambda(C) = 0$; in this sense the Cantor distribution is singular.



1.4 Lebesgue measure and Lebesgue measurable sets

The Lebesgue measure on the line has natural generalisation to Euclidean spaces $\mathbb{R}^k$. For a rectangular parallelepiped $A = [a_1, b_1] \times \cdots \times [a_k, b_k]$ its Lebesgue measure is defined as the k-dimensional volume

$$\lambda^{(k)}(A) = \prod_{i=1}^{k}(b_i - a_i).$$

The σ-algebra of Borel sets $\mathcal{B}(\mathbb{R}^k)$ in k dimensions is the σ-algebra generated by open sets in $\mathbb{R}^k$. Like in $\mathbb{R}$, there is a more spare systems of generators generalising the half-lines in one dimension

$$\mathcal{S} = \{(-\infty, x_1] \times \cdots \times (-\infty, x_k] : (x_1, \dots, x_k) \in \mathbb{R}^k\}.$$

There is a larger than $\mathcal{B}(\mathbb{R}^k)$ σ-algebra of sets, to which the Lebesgue measure can be extended. If A is a Borel set with $\lambda^{(k)}(A) = 0$ and $B \subset A$ it is reasonable to asiign to B measure 0. The σ-algebra generated by $\mathcal{B}(\mathbb{R}^k)$ and such subsets B is the σ-algebra of *Lebesgue*-measurable sets. This operation of adding subsets of zero-measure sets is called *completion*, that is the σ-algebra of Lebesgue-measurable sets is complete.

Using tranfinite induction, tt can be shown, that the cardinality of $\mathcal{B}(\mathbb{R})$ is continuum. Onte other hand, for Cantor set C the cardinality of the power-set $\mathcal{P}(C)$ is bigger than continuum, and each $A \subset C$ is Lebesgue-measurable. It follows that there are more Lebesgue-measurable sets than Borel sets. Hence many Lebesgue-measurable non-Borel sets exist although they do not admit a constructive description.

Exercises

1. For $A \subset \Omega$ proper subset, describe $\sigma(\{A\})$.

2. Let $\Omega = [0, 1]$. Find the σ generated by $\{[0, 1/4), (3/4, 1]\}$.

3. Show that the increasing monotonicity property is equivalent to σ-additivity.

4. Let $\mathcal{A}$ be the family of sets $A \in \mathcal{B}(\mathbb{R})$ with the property that there exists a limit

$$\mu(A) = \lim_{n\to\infty} n^{-1}\lambda(A \cap [0, n]).$$

   Show that $\mathcal{A}$ is an algebra. Is μ σ-additive on $\mathcal{A}$?

5. Consider the space of functions $x : \mathbb{R} \to \mathbb{R}$. Let $\mathcal{S} = \{x : x(t) \in B, \text{ for some } t \in \mathbb{R}, B \in \mathcal{B}(\mathbb{R})\}$. Show that all sets in $\sigma(\mathcal{S})$ have the form

$$A = \{x : (x(t_1), \dots, x(t_k)) \in D\}$$

   for some k, $t_1 < \cdots < t_k$ and $D \in \mathcal{B}(\mathbb{R}^k)$.

6. Is there a probability measure μ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ supported by $\mathbb{Q}$, i.e. such that $\mu(x) > 0$ for $x \in \mathbb{Q}$, $\sum_{x\in\mathbb{Q}} \mu(x) = 1$. Is this measure discrete or diffuse?
7. For $A \subset \mathbb{R}$ define $x + A := \{x + a, a \in A\}$. Prove translation invariance of the Lebesgue measure: $\lambda(x + A) = \lambda(A), A \in \mathcal{B}(\mathbb{R})$. Extend the property to Lebesgue-measurable sets A.
8. Explain why the distribution function is right-continuous with left limits.
9. Show that every probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ admits a representation $\mu + \nu$, where μ is a discrete measure and ν is a diffuse measure.
10. Let $\mu = \sum_{j=1}^{\infty} 2^{-j}\delta_j$. Is it a probability measure? Sketch the graph of its cumulative distribution function.
11. Let $\Omega = \{0, 1\}^{\infty}$. Using set-teoretic operations $\cup, \cap, ^c$ express the event

$$A = \{\omega \in \Omega : \lim_{k\to\infty} (\omega_1 + \cdots + \omega_k)/k = z\}$$

    in terms of events $A(\epsilon_1, \ldots, \epsilon_k)$.
12. (First half of Borel-Cantelli lemma) Let $A_j, j \in \mathbb{N}$, be events in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\sum_{j=1}^{\infty} \mathbb{P}(A_j) < \infty$. Prove that $\mathbb{P}(\cap_{n=1}^{\infty} \cup_{j=n}^{\infty} A_j) = 0$.
13. Consider $\mathcal{S} := \{\{x\} : x \in \mathbb{R}\}$. Show that for $A \in \sigma(\mathcal{S})$, either A is countable (i.e. either finite or countably infinite) or A^c is countable. Now let $\mu(x) = 1$ for every $x \in \mathbb{R}$. What are possible values of $\mu(A)$? When $\mu(A) = \infty$?
14. For Borel sets $A, B \in \mathcal{B}(\mathbb{R})$ let $d(A, B) = \lambda(A\Delta B)$. Show that $d(A, B)$ is a metric on $\mathcal{B}(\mathbb{R})$ (in particular, satisfies the triangle inequality). The metric space $(d, \mathcal{B}(\mathbb{R}))$ is not complete: some Cauchy sequences do not have a limit. Show that the completion of the metric space $(d, \mathcal{B}(\mathbb{R}))$ is the σ-algebra of Lebesgue sets.

Literature

1. S. Resnick, A probability path, Springer 2003.
2. R. Schilling, Measures, integrals and martingales, CUP 2005.