36-710/752 Fall 2020
Advanced Probability Overview

**Lecture 02 - Foundations of Measures**

Lecturer : Alessandro Rinaldo Scribe: Mike Stanley

# 2.1 Fields

In the last lecture, we defined a field and a $\sigma$-field. We recall those definitions here.

**Definition 1** *Let $\Omega$ be a universe set. A collection $\mathcal{F}$ of subsets of $\Omega$ is a field when the following conditions hold:*

1. $\Omega \in \mathcal{F}$
2. $A \in \mathcal{F} \implies A^c \in \mathcal{F}$ *(Note, thus $\varnothing \in \mathcal{F}$)*
3. $A_1, A_2 \in \mathcal{F} \implies A_1 \cup A_2 \in \mathcal{F}$ *(i.e. $\mathcal{F}$ is closed under finite unions and intersections)*

**Definition 2** *A field $\mathcal{F}$ is a $\sigma$-field if for every sequence $\{A_n\}_{n\in\mathbb{N}}$ in $\mathcal{F}$, $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{F}$ (i.e. closed with respect to countable unions and intersections).*

## 2.1.1 An example of a field without countable additivity

Suppose $\Omega = \mathbb{R}$. Let $\mathcal{U}$ be the collection of unions of finitely many disjoint sets of the form $(a, b], (-\infty, b], (a, \infty), (-\infty, \infty)$ where $-\infty < a \leq b < \infty$. $\mathcal{U}$ is a field but is not a $\sigma$-field because it is not closed with respect to countable unions. To see this, consider the set $(a, b)$. $(a, b) \notin \mathcal{U}$, but we can define $(a, b)$ as a union of countable elements in $\mathcal{U}$, namely, $(a, b) = \bigcup_{n\in\mathbb{N}} (a, b - \frac{1}{n}]$.

## 2.1.2 Creation of $\sigma$-fields

Let $A \subset \Omega$. The smallest $\sigma$-field containing $A$ is $\mathcal{F} = \{\varnothing, \Omega, A, A^c\}$. This follows directly from the definition of a $\sigma$-field, namely inclusion of universe set and closure under complements. Of course, there are also many other $\sigma$-fields that contain $A$ (e.g. $2^\Omega$, this is, in fact the largest), but we are interested in the smallest possible $\sigma$-field.

**Definition 3** *More generally, suppose $\mathcal{C}$ is a collection of subsets of $\Omega$. $\sigma(\mathcal{C})$ is used to denote the smallest $\sigma$-field generated by $\mathcal{C}$, also called the generated $\sigma$-field of $\mathcal{C}$.*

Intuitively, in order to define a probability of an event (i.e. a subset of some universe set), we need to be able to map every subset of the universe set to the reals, or more specifically $[0, 1]$. So, it would be nice to find the smallest $\sigma$-field that contains all of the sets to which we want to assign probabilities. This goal can be achieved with the above notion of a generated $\sigma$-field, in which $\mathcal{C}$ is our collection of subsets. A more particular construct of this type is the Borel $\sigma$-field.

**Definition 4** *Let $\Omega$ be a topological space (i.e. a collection of points with neighborhoods around each point). Let $\mathcal{C}$ be the collection of open sets of $\Omega$. $\sigma(\mathcal{C})$, i.e. the $\sigma$-field generated by the collection of open sets, is known as the Borel $\sigma$-field.*

For example, suppose $\Omega = \mathbb{R}$. Then, the collection of open sets $\mathcal{C} = \{(a, b) : -\infty < a, b < \infty\}$ can be used to generate the Borel $\sigma$-field over $\mathbb{R}$, and is often with $\mathcal{B}$.

## 2.2 Measures

A probability space is a particular form of a measure space. Before defining these terms, first note that $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty\} \cup \{\infty\}$ is known as the *extended reals.*

**Definition 5** *Let $(\Omega, \mathcal{F})$ be a measurable space, where $\mathcal{F}$ is a $\sigma$-field of $\Omega$. A function $\mu : \mathcal{F} \to \bar{\mathbb{R}}_+$ is a measure if*

1. $\mu(\varnothing) = 0$
2. *For every sequence $\{A_n\}_{n \in \mathbb{N}}$ of mutually disjoint measurable sets*

$$\mu(\bigcup_n A_n) = \sum_n \mu(A_n) \tag{2.1}$$

   *This property is known as "countable additivity"*

*Then, a measure space is a tuple $(\Omega, \mathcal{F}, \mu)$.*

A measure on a field $\mathcal{F}'$ is a function $\mu : \mathcal{F}' \to \bar{\mathbb{R}}_+$ that satisfies the conditions (1) and (2) in the above definition. Additionally, it should be noted that measures can be finite ($\mu(\Omega) < \infty$) or infinite ($\mu(\Omega) = \infty$). As we see in the next definition, a probability measure is simply a measure such that $\mu(\Omega) = 1$.

**Definition 6** *A probability measure is a measure such that* $\mu(\Omega) = 1$*. If* $\mathcal{P}$ *is a probability measure,* $(\Omega, \mathcal{F}, \mathcal{P})$ *is a probability space.*

In more standard statistical parlance, $\mathcal{F}$ may be thought of as a collection of events over the universe set $\Omega$. Hence the probability of some event $A \in \mathcal{F}$ is defined with the probability measure, $\mathcal{P}(A)$.

Let us consider two examples of probability measures.

**Example 1** *Let* $\Omega$ *be countable, i.e.* $\Omega = \{\omega_1, \omega_2, \dots\}$*,* $\mathcal{F} = 2^{\Omega}$*, and* $\{p_i\}_{i \in \mathbb{N}}$ *be such that* $p_i \in [0,1]$ *and* $\sum_i p_i = 1$*. Then the function* $\mathcal{P} : 2^{\Omega} \to [0,1]$ *given by* $\mathcal{P}(A) = \sum_{i:\omega_i \in A} p_i$ *is a probability measure. This is clearly true since* $\mathcal{P}(\varnothing) = 0$ *and for mutually disjoint sets* $A_1, A_2 \in 2^{\Omega}$*,* $\mathcal{P}(A_1 \cup A_2) = \sum_{i:\omega_i \in A_1 \cup A_2} p_i = \sum_{i:\omega_i \in A_1} p_i + \sum_{i:\omega_i \in A_2} p_i = \mathcal{P}(A_1) + \mathcal{P}(A_2)$*.*

**Example 2** *Let* $\Omega$ $\mathbb{R}$ *and* $\mathcal{F} = \mathcal{B}^1$ *(i.e. the Borel* $\sigma$*-field on* $\mathbb{R}$*). Define*

$$\mathcal{P}((-\infty, a]) = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} \exp^{\frac{-x^2}{2}} dx \tag{2.2}$$

*for all* $(-\infty, a]$*,* $a \in \mathbb{R}$*. This also defines a probability measure on* $\mathcal{B}^1$ *since* $\mathcal{P}(\varnothing) = 0$*, and for countable disjoint sets* $A_1, A_2, \dots \in \mathcal{B}^1$*,* $\mathcal{P}(\bigcup_n A_n) = \sum_n \mathcal{P}(A_n)$*, but properties of integrals.*

To make these definitions slightly less abstract, let us consider two more specific types of measures.

**Definition 7** *Take* $\Omega$ *to be any set and let* $\mathcal{F} = 2^{\Omega}$*. For any* $A \in \mathcal{F}$*, the counting measure is defined as* $\mu(A) = \mid A \mid$*.*

We may also wish to put a countability criterion onto our measure. The following measure is a regularity condition.

**Definition 8** *Let* $(\Omega, \mathcal{F}, \mu)$ *be a measure space.* $\mu$ *is said to be* $\sigma$*-finite if there exists a countable collection of measurable sets* $\{A_1, A_2, \dots\}$ *such that* $\mu(A_n) < \infty$ *and* $\bigcup_n A_n = \Omega$*.*

### 2.2.1 Properties of Measures

Assume throughout a measure space $(\Omega, \mathcal{F}, \mu)$.

**Claim 3** *If $A \subseteq B$, then $\mu(A) \leq \mu(B)$.*

**Proof:** Note that we can write $B = A \cup (B \cap A^c)$. Thus:

$$\begin{aligned}\mu(B) &= \mu(A \cup (B \cap A^c)) \\ &= \mu(A) + \mu(B \cap A^c) \\ &\geq \mu(A)\end{aligned}$$

The second line follows from the additivity property of measures of disjoint sets. ■

**Claim 4** *More generally, if $\{A_n\}$ is a sequence of measurable sets, then*

$$\mu(\bigcup_n A_n) \leq \sum_n \mu(A_n) \tag{2.3}$$

**Proof:** Define a sequence $\{B_n\}$ of measurable sets as follows. $B_1 = A_1$. For all $n \geq 2$, $B_n = A_n - \bigcup_{i=1}^{n-1} B_i = A_n \cap A_{n-1}^c \cap \cdots \cap A_1^c$. This imples that $\bigcap_n B_n = \bigcap_n A_n$, and $\{B_n\}$ is a sequence of disjoint sets. Thus,

$$\begin{aligned}\mu(\bigcap_n A_n) &= \mu(\bigcap_n B_n) \\ &= \sum_n \mu(B_n) \\ &\leq \sum_n \mu(A_n)\end{aligned}$$

where the last line follows because $B_n \subseteq A_n$. ■

Note, if $\mu$ is a probability measure, the countable additivity property is often known as the *union bound.*

We also have the following two interesting properties of measures:

1. if $\mu(A_n) = 0$ for all $n$, then $\mu(\bigcap_n A_n) = 0$
2. if $\mu(A_n) = 1$ for all $n$, then $\mu(\bigcap_n A_n) = 1$

Ultimately, we will be interested in looking at mathematical properties over measurable spaces. So, it stands to reason that there should be a notion for a property that holds on all elements of $\mathcal{F}$ that have non-zero measure.

**Definition 9** *Suppose that a certain property hold for all* $\omega \in A^c$ *where* $\mu(A) = 0$*. Then, we say that the property holds almost everywhere, abbreviated as* a.e$[\mu]$*. If* $mu = \mathcal{P}$ *(a probability measure), we say instead that the property holds almost surely, abbreviated as* a.s$[\mathcal{P}]$*.*

Similarly, in talking about identifying the events on which there is non-zero measure, we present the following definition.

**Definition 10** *Given a probability space* $(\Omega, \mathcal{F}, \mathcal{P})$*, the suppose of* $\mathcal{P}$ *is the smallest closed set* $S \subset \Omega$ *such that* $\mathcal{P}(S) = 1$*. If* $A \subseteq S^c$*, then* $\mathcal{P}(A) = 0$*.*

## 2.3 Next lecture preview

In the next lecture, we will discuss continuity of measures. Recall from analysis that if $f : \mathbb{R} \to \mathbb{R}$ is a continuous function on its domain, then

$$f(x^*) = \lim_{x_n \to x^*} f(X_n) \tag{2.4}$$

In the next lecture, we will extend this notion to measures.