36-752	Spring 2018
Advanced Probability Overview	

Lecture Notes Set 1: Course Overview, σ-Fields, and Measures

Instructor: Alessandro Rinaldo

Associated reading: Sec 1.1-1.4 of Ash and Doléans-Dade; Sec 1.1 and A.1 of Durrett.

1 Introduction

How is this course different from your earlier probability courses? There are some problems that simply can't be handled with finite-dimensional sample spaces and random variables that are either discrete or have densities.

Example 1 *Try to express the strong law of large numbers without using an infinite-dimensional space. Oddly enough, the weak law of large numbers requires only a sequence of finite-dimensional spaces, but the strong law concerns entire infinite sequences.*

Example 2 *Consider a distribution whose cumulative distribution function (cdf) increases continuously part of the time but has some jumps. Such a distribution is neither discrete nor continuous. How do you define the mean of such a random variable? Is there a way to treat such distributions together with discrete and continuous ones in a unified manner?*

General Measures Both of the above examples are accommodated by a generalization of the theories of summation and integration. Indeed, summation becomes a special case of the more general theory of integration. It all begins with a generalization of the concept of "size" of a set.

Example 3 *One way to measure the size of a set is to count its elements. All infinite sets would have the same size (unless you distinguish different infinite cardinals).*

Example 4 *Special subsets of Euclidean spaces can be measured by length, area, volume, etc. But what about sets with lots of holes in them? For example, how large is the set of irrational numbers between 0 and 1?*

We will use measures to say how large sets are. First, we have to decide which sets we will measure.

2 Set-Theoretic Preliminaries

Universe set. Let Ω be a *universe set.* Every set will be implicitly assumed to be a subset of Ω and set theoretic operations (union, intersection and complement) are well defined only with respect to Ω.

Power set. The power set of A, denoted with 2^A is the set of all subsets of A (including the empty set and A itself).

Monotone Sequences of Sets. A sequence (finite or infinite) of sets $A_1, A_2, \ldots$ such that $A_1 \subset A_2 \subset \cdots$ is said to be *increasing.* The sequence has *limit* $A = \bigcup_n A_n$, in which case we say that A_n increases to A, written $A_n \uparrow A$. Similarly, if $A_1 \supset A_2 \supset \cdots$, the sequence is *decreasing*; it is said to decrease to its limit set $A = \bigcap_n A_n$, written $A_n \downarrow A$. A sequence of sets is *monotone* if it is either increasing or decreasing.

Exercise 1 *Let $a < b$ be real numbers and set $A_n = [a - \frac{1}{n}, b - \frac{1}{n}]$. Find $\bigcup_n A_n$. Similarly, let $B_n = (a + \frac{1}{n}, b + \frac{1}{n})$. Find $\bigcap B_n$. This shows that an infinite union of closed sets needs not be closed and an infinite (non-empty) intersection of open sets needs not be open. What about arbitrary unions of open sets and intersections of closed sets?*

DeMorgan Laws: If $A_1, A_2, \ldots$ is an arbitrary sequence, $\left(\bigcup_n A_n\right)^c = \bigcap_n A_n^c$ and $\left(\bigcap_n A_n\right)^c = \bigcup_n A_n^c$. Thus $A_n \downarrow A$ if and only if $A_n^c \uparrow A^c$.

From union of sets to union of disjoint sets. Let $A_1, A_2, \ldots$ an arbitrary sequence of sets in Ω. Then

$$\bigcup_n A_n = \bigcup_n B_n,$$

where $B_n = A_n \cap A_{n-1}^c \cap \ldots \cap A_1^c$, with $A_0 = \varnothing$, and the B_n's are disjoint.

Upper and lower limit of sequence of sets. For any arbitrary sequence $A_1, A_2, \ldots$, its *limit superior* is

$$\limsup_n A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$$

and its *limit inferior* is

$$\liminf_n A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k.$$

Thus, $\omega \in \limsup_n A_n$ iif for every n, $\omega \in A_k$ for some (in fact, infinitely many) $k \geq n$. Equivalently, $\omega \in \limsup_n A_n$ if and only if $\omega \in A_n$ for infinitely many n's, or *infinitely often.*

Conversely, $\omega \in \liminf_n A_n$ iif, for some n, $\omega \in A_k$ for all $k \geq n$. That is, $\omega \in \liminf_n A_n$ iif $\omega \in A_n$ for all but finitely many n's, or *eventually*.

The sequence $A_1, A_2, \ldots$ has a limit A iif $\limsup_n A_n = \liminf_n A_n = A$. In case of monotone (i.e. either increasing or decreasing) sequences, we recover the notion of limit introduced above.

Exercise 2 *It is instructive to compare the notion of limit superior and inferior to the analogous notion for sequences, where for a sequence of real numbers* $x_1, x_2, \ldots$ *recall that* $\limsup_n x_n = \inf_n \sup_{k \geq n} x_k$ *and* $\liminf_n x_n = \sup_n \inf_{k \geq n} x_k$. *For a set A, let* $I_A \colon \Omega \to \{0,1\}$ *be its indicator function:*

$$I_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A, \\ 0 & \text{if } \omega \notin A. \end{cases}$$

Then $I_{\limsup_n A_n} = \limsup_n I_{A_n}$ *and* $I_{\liminf_n A_n} = \liminf_n I_{A_n}$.

Countable vs uncountable sets. A set A is finite if $|A| < \infty$, where $|A|$ (or $\mathrm{card}(A)$ or $\#A$) is the number of elements or *cardinality* of A, and infinite otherwise. A better distinction, which is very important in measure theory, is between sets that are *countable* versus sets that are *uncountable*. A set A is countable if there exists a function $\phi \colon A \to \mathbb{N}$ mapping the elements of A into the naturals that is injective. It is just a mathematical way of expressing the fact that the set of natural numbers is "large enough" compared to A so that *all* elements of A can be labeled using the naturals (or a subset thereof). A countable set may be finite or infinite. An uncountable set is always infinite. A finite set if always countable.

Exercise 3 *Show that the cartesian product of two countable sets is countable. Conclude that the cartesian product of finitely many countable sets is countable. Use this to show that countable unions of countable sets if countable. On the other hand, the countable cartesian product of countable sets is not countable. To see this, show that even the set of infinite binary sequences is not countable.*

Exercise 4 *Use the result in the previous exercise to show that the power set of an infinite countable set is not countable.*

Example 5 (The unit interval is uncountable.) *We have seen that the set of infinite binary sequences is uncountable. The claim therefore will follow if we can show that each number in* $(0,1]$ *can be expressed as an infinite binary sequence. Let T a mapping of the interval* $\Omega = (0,1]$ *into itself given by*

$$T\omega = \begin{cases} 2\omega & \text{if } \; 0 < \omega \leq 1/2, \\ 2\omega - 1 & \text{if } \; 1/2 < \omega \leq 1. \end{cases}$$

Now define d_1 on Ω by

$$d_1(\omega) = \begin{cases} 0 & \textit{if} \quad 0 < \omega \leq 1/2, \\ 1 & \textit{if} \quad 1/2 < \omega \leq 1, \end{cases}$$

and for any integer $i > 1$ and $\omega \in \Omega$, set $d_i(\omega) = d_1(T^{i-1}\omega)$. Then, it can be shown that, for all $n \geq 1$,

$$\sum_{i=1}^{n} \frac{d_i(\omega)}{2^i} < \omega \leq \sum_{i=1}^{n} \frac{d_i(\omega)}{2^i} + \frac{1}{2^n}, \quad \forall \omega \in \Omega. \tag{5}$$

As as result,

$$\omega = \sum_{i=1}^{\infty} \frac{d_i(\omega)}{2^i}, \quad \forall \omega \in \Omega.$$

This gives the dyadic representation of each ω in $(0,1]$ as a binary sequence $(d_i(\omega), i = 1, 2, \ldots)$. Notice that if $d_i(\omega) = 0$ for all $i > n$, then $\omega = \sum_{i=1}^{n} \frac{d_i(\omega)}{2^i}$, contradicting the strict inequality in (5). Thus, the binary representation of each $\omega \in \Omega$ does not terminate in 0's (equivalently, it contains an infinite number of 1's). Thus, we have shown that Ω can be represented as the set if infinite binary sequences that do not terminate in 0's. Since $(0,1]$ is in one-to-one correspondence with any interval on the real line of the form $(a, b]$ with $a < b$, we conclude that any interval on the real line has uncountably many points.

3 σ-fields

Definition 1 (fields and σ-fields) *Let Ω be a set. A collection $\mathcal{F}$ of subsets of Ω is called a* field *if it satisfies*

- $\Omega \in \mathcal{F}$,
- *for each $A \in \mathcal{F}$, $A^C \in \mathcal{F}$,*
- *for all $A_1, A_2 \in \mathcal{F}$, $A_1 \cup A_2 \in \mathcal{F}$.*

A field $\mathcal{F}$ is a σ-field *if, in addition, it satisfies*

- *for every sequence $\{A_k\}_{k=1}^{\infty}$ in $\mathcal{F}$, $\bigcup_{k=1}^{\infty} A_k \in \mathcal{F}$.*

We will define measures on fields and σ-field's.

Definition 2 (Measurable Space) *A set Ω together with a σ-field $\mathcal{F}$ is called a* measurable space $(\Omega, \mathcal{F})$, *and the elements of $\mathcal{F}$ are called* measurable sets .

Example 6 (Intervals on $\mathbb{R}^1$) *Let $\Omega = I\!R$ and define $\mathcal{U}$ to be the collection of all unions of finitely many disjoint intervals of the form $(a, b]$ or $(-\infty, b]$ or (a, ∞) or $(-\infty, \infty)$, together with $\varnothing$. Then $\mathcal{U}$ is a field.*

Example 7 (Power set) *Let Ω be an arbitrary set. The collection of all subsets of Ω is a σ-field, in fact the largest σ-field containing Ω. It is denoted 2^Ω and is called the* power set of Ω.

Example 8 (Trivial σ-field) *Let Ω be an arbitrary set. Let $\mathcal{F} = \{\Omega, \varnothing\}$. This is the* trivial σ-field.

Exercise 6 *Let $\mathcal{F}_1$, $\mathcal{F}_2, \ldots$ be classes of sets in a common space Ω such that $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ for each n. Show that if each $\mathcal{F}_n$ is a field, then $\cup_{n=1}^\infty \mathcal{F}_n$ is also a field.*

If each $\mathcal{F}_n$ is a σ-field, then is $\cup_{n=1}^\infty \mathcal{F}_n$ also necessarily a σ-field? Think about the following case: Ω is the set of nonnegative integers and $\mathcal{F}_n$ is the σ-field of all subsets of $\{0, 1, \ldots, n\}$ and their complements.

Generated σ-fields A field is closed under finite set theoretic operations whereas a σ-field is closed under countable set theoretic operations. In a problem dealing with probabilities, one usually deals with a small class of subsets $\mathcal{A}$, for example the class of subintervals of $(0, 1]$. It is possible that if we perform countable operations on such a class $\mathcal{A}$ of sets, we might end up operating on sets outside the class $\mathcal{A}$. Hence, we would like to define a class denoted by $\sigma(\mathcal{A})$ in which we can safely perform countable set-theoretic operations. This class $\sigma(\mathcal{A})$ is called the σ-field generated by $\mathcal{A}$, and it is defined as the intersection of all the σ-fields containing $\mathcal{A}$ (exercise: show that this is a σ-field). $\sigma(\mathcal{A})$ is the smallest σ-field containing $\mathcal{A}$.

Example 9 *Let $\mathcal{C} = \{A\}$ for some nonempty A that is not itself Ω. Then $\sigma(\mathcal{C}) = \{\varnothing, A, A^C, \Omega\}$.*

Example 10 *Let $\Omega = I\!R$ and let $\mathcal{C}$ be the collection of all intervals of the form $(a, b]$. Then the field generated by $\mathcal{C}$ is $\mathcal{U}$ from Example 6 while $\sigma(\mathcal{C})$ is larger.*

Example 11 (Borel σ-field) *Let Ω be a topological space and let $\mathcal{C}$ be the collection of open sets. Then $\sigma(\mathcal{C})$ is called the* Borel σ-field. *If $\Omega = I\!R$, the Borel σ-field is the same as $\sigma(\mathcal{C})$ in Example 10. The Borel σ-field of subsets of $I\!R^k$ is denoted $\mathcal{B}^k$.*

Exercise 7 *Give some examples of classes of sets $\mathcal{C}$ such that $\sigma(\mathcal{C}) = \mathcal{B}^1$.*

Exercise 8 *Are there subsets of $I\!R$ which are not in $\mathcal{B}^1$?*

4 Measures

Notation 12 (Extended Reals) *The* extended reals *is the set of all real numbers together with* ∞ *and* $-\infty$*. We shall denote this set* $\overline{\mathbb{R}}$*. The* positive extended reals*, denoted* $\overline{\mathbb{R}}^{+}$ *is* $(0, \infty]$*, and the* nonnegative extended reals*, denoted* $\overline{\mathbb{R}}^{+0}$ *is* $[0, \infty]$*.*

Definition 3 *Let* $(\Omega, \mathcal{F})$ *be a measurable space. Let* $\mu : \mathcal{F} \to \overline{\mathbb{R}}^{+0}$ *satisfy*

- $\mu(\varnothing) = 0$,
- *for every sequence* $\{A_k\}_{k=1}^{\infty}$ *of mutually disjoint elements of* $\mathcal{F}$*,* $\mu(\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} \mu(A_k)$*.*

Then μ *is called a* measure *on* $(\Omega, \mathcal{F})$ *and* $(\Omega, \mathcal{F}, \mu)$ *is a* measure space*. If* $\mathcal{F}$ *is merely a field, then a* μ *that satisfies the above two conditions whenever* $\bigcup_{k=1}^{\infty} A_k \in \mathcal{F}$ *is called a* measure on the field $\mathcal{F}$*.*

Example 13 *Let* Ω *be arbitrary with* $\mathcal{F}$ *the trivial* σ*-field. Define* $\mu(\varnothing) = 0$ *and* $\mu(\Omega) = c$ *for arbitrary* $c > 0$ *(with* $c = \infty$ *possible).*

Example 14 (Counting measure) *Let* Ω *be arbitrary and* $\mathcal{F} = 2^{\Omega}$*. For each finite subset* A *of* Ω*, define* $\mu(A)$ *to be the number of elements of* A*. Let* $\mu(A) = \infty$ *for all infinite subsets. This is called* counting measure *on* Ω*.*

Definition 4 (Probability measure) *Let* $(\Omega, \mathcal{F}, P)$ *be a measure space. If* $P(\Omega) = 1$*, then* P *is called a* probability*,* $(\Omega, \mathcal{F}, P)$ *is a* probability space*, and elements of* $\mathcal{F}$ *are called* events*.*

Sometimes, if the name of the probability P is understood or is not even mentioned, we will denote $P(E)$ by $\Pr(E)$ for events E.

Infinite measures pose a few unique problems. Some infinite measures are just like finite ones.

Definition 5 (σ-finite measure) *Let* $(\Omega, \mathcal{F}, \mu)$ *be a measure space, and let* $\mathcal{C} \subseteq \mathcal{F}$*. Suppose that there exists a sequence* $\{A_n\}_{n=1}^{\infty}$ *of elements of* $\mathcal{C}$ *such that* $\mu(A_n) < \infty$ *for all* n *and* $\Omega = \bigcup_{n=1}^{\infty} A_n$*. Then we say that* μ *is* σ-finite on $\mathcal{C}$*. If* μ *is* σ*-finite on* $\mathcal{F}$*, we merely say that* μ *is* σ-finite*.*

Example 15 *Let* $\Omega = \mathbb{Z}$ *with* $\mathcal{F} = 2^{\Omega}$ *and* μ *being counting measure. This measure is* σ*-finite. Counting measure on an uncountable space is not* σ*-finite.*

Exercise 9 *Prove the claims in Example 15.*

4.1 Basic properties of measures

There are several useful properties of measures that are worth knowing.

First, measures are countably subadditive in the sense that

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu(A_n), \tag{10}$$

for arbitrary sequences $\{A_n\}_{n=1}^{\infty}$. The proof of this uses a standard trick for dealing with countable sequences of sets. Let $B_1 = A_1$ and let $B_n = A_n \setminus \bigcup_{i=1}^{n-1} B_i$ for $n > 1$. The B_n's are disjoint and have the same finite and countable unions as the A_n's. The proof of Equation 10 relies on the additional fact that $\mu(B_n) \leq \mu(A_n)$ for all n.

Next, if $\mu(A_n) = 0$ for all n, it follows that $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = 0$. This gets used a lot in proofs. Similarly, if μ is a probability and $\mu(A_n) = 1$ for all n, then $\mu\left(\bigcap_{n=1}^{\infty} A_n\right) = 1$.

Definition 6 (Almost sure/almost everywhere) *Suppose that some statement about elements of Ω holds for all $\omega \in A^C$ where $\mu(A) = 0$. Then we say that the statement holds* almost everywhere*, denoted a.e. $[\mu]$. If P is a probability, then almost everywhere is often replaced by* almost surely*, denoted a.s. $[P]$.*

Example 16 *Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of functions from Ω to $I\!R$. To say that X_n converges to X a.s. $[P]$ (denoted $X_n \stackrel{\text{a.s.}}{\to} X$) means that there is a set A with $P(A) = 0$ and $\lim_{n\to\infty} X_n(\omega) = X(\omega)$ for all $\omega \in A^C$.*

Proposition 11 (Linearity) *If $\mu_1, \mu_2, \ldots$ are all measures on $(\Omega, \mathcal{F})$ and if $\{a_n\}_{n=1}^{\infty}$ is a sequence of positive numbers, then $\sum_{n=1}^{\infty} a_n \mu_n$ is a measure on $(\Omega, \mathcal{F})$.*

Exercise 12 *Prove Proposition 11.*

4.2 Monotone sequences of sets and limits of measure

Definition 7 (Monotone sequences of sets) *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. A sequence $\{A_n\}_{n=1}^{\infty}$ of elements of $\mathcal{F}$ is called* monotone increasing *if $A_n \subseteq A_{n+1}$ for each n. It is* monotone decreasing *if $A_n \supseteq A_{n+1}$ for each n.*

Lemma 17 *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Let $\{A_n\}_{n=1}^{\infty}$ be a monotone sequence of elements of $\mathcal{F}$. Then $\lim_{n\to\infty} \mu(A_n) = \mu\left(\lim_{n\to\infty} A_n\right)$ if either of the following hold:*

- *the sequence is increasing,*

- *the sequence is decreasing and $\mu(A_k) < \infty$ for some k.*

If $\{A_n\}_{n=1}^{\infty}$ is any sequence of measurable sets and μ is finite, then

$$\mu\left(\liminf_n A_n\right) \leq \liminf_n \mu(A_n) \leq \limsup_n \mu(A_n) \leq \mu(\limsup_n A_n).$$

In particular, if $\lim_n A_n = A$ exists, then $\lim_n \mu(A_n) = \mu(A)$.

Proof: Define $A_\infty = \lim_{n\to\infty} A_n$. In the first case, write $B_1 = A_1$ and $B_n = A_n \setminus A_{n-1}$ for $n > 1$. Then $A_n = \bigcup_{k=1}^{n} B_k$ for all n (including $n = \infty$). Then $\mu(A_n) = \sum_{k=1}^{n} \mu(B_k)$, and

$$\mu\left(\lim_{n\to\infty} A_n\right) = \mu(A_\infty) = \sum_{k=1}^{\infty} \mu(B_k) = \lim_{n\to\infty} \sum_{k=1}^{n} \mu(B_k) = \lim_{n\to\infty} \mu(A_n).$$

In the second case, write $B_n = A_n \setminus A_{n+1}$ for all $n \geq k$. Then, for all $n > k$,

$$\begin{aligned} A_k \setminus A_n &= \bigcup_{i=k}^{n-1} B_i, \\ A_k \setminus A_\infty &= \bigcup_{i=k}^{\infty} B_i. \end{aligned}$$

By the first case,

$$\lim_{n\to\infty} \mu(A_k \setminus A_n) = \mu\left(\bigcup_{i=k}^{\infty} B_i\right) = \mu(A_k \setminus A_\infty).$$

Because $A_n \subseteq A_k$ for all $n > k$ and $A_\infty \subseteq A_k$, it follows that

$$\begin{aligned} \mu(A_k \setminus A_n) &= \mu(A_k) - \mu(A_n), \\ \mu(A_k \setminus A_\infty) &= \mu(A_k) - \mu(A_\infty). \end{aligned}$$

It now follows that $\lim_{n\to\infty} \mu(A_n) = \mu(A_\infty)$.

As for the second claim, for each $n \geq 1$ let $B_n = \cap_{k=n}^{\infty} A_n$ and $C_n \cup_{k=n}^{\infty} A_n$. Then, $B_n \to \liminf_n A_n$ and $C_n \to \limsup_n A_n$. Thus, for each n

$$\mu(A_n) \geq \mu(B_n),$$

which implies that

$$\liminf_n \mu(A_n) \geq \liminf_n \mu(B_n) = \lim_n \mu(B_n) = \mu(\liminf_n A_n).$$

Similarly, the fact that $\mu(A_n) \leq \mu(C_n)$ for all n implies that and

$$\limsup_n \mu(A_n) \leq \limsup_n \mu(C_n) = \lim_n \mu(C_n) = \mu(\limsup_n A_n).$$

The claims easily follow. ∎

Exercise 13 *Construct a simple counterexample to show that the condition $\mu(A_k) < \infty$ is required in the second claim of Lemma 17.*

4.3 Uniqueness of Measures

There is a popular method for proving uniqueness theorems about measures. The idea is to define a function μ on a convenient class $\mathcal{C}$ of sets and then prove that there can be at most one extension of μ to $\sigma(\mathcal{C})$.

Example 18 *Suppose it is given that for any $a \in \mathbb{R}$,*

$$P((-\infty, a]) = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} \exp\bigl(-u^2/2\bigr)\, du.$$

Does that uniquely define a **unique** *probability measure on the class of Borel subsets of the line, $\mathcal{B}^1$?*

Definition 8 (π-system and λ-system) *A collection $\mathcal{A}$ of subsets of Ω is a* π-system *if, for all $A_1, A_2 \in \mathcal{A}$, $A_1 \cap A_2 \in \mathcal{A}$. A class $\mathcal{C}$ is a* λ-system *if*

- $\Omega \in \mathcal{C}$,
- *for each $A \in \mathcal{C}$, $A^C \in \mathcal{C}$,*
- *for each sequence $\{A_n\}_{n=1}^{\infty}$ of disjoint elements of $\mathcal{C}$, $\bigcup_{n=1}^{\infty} A_n \in \mathcal{C}$.*

Example 19 *The collection of all intervals of the form $(-\infty, a]$ is a π-system of subsets of $\mathbb{R}$. So too is the collection of all intervals of the form $(a, b]$ (together with $\varnothing$). The collection of all sets of the form $\{(x, y) : x \leq a, y \leq b\}$ is a π-system of subsets of $\mathbb{R}^2$. So too is the collection of all rectangles with sides parallel to the coordinate axes.*

Some simple results about π-systems and λ-systems are the following.

Proposition 14 *If Ω is a set and $\mathcal{C}$ is both a π-system and a λ-system, then $\mathcal{C}$ is a σ-field.*

Proposition 15 *Let Ω be a set and let Λ be a λ-system of subsets. If $A \in \Lambda$ and $A \cap B \in \Lambda$ then $A \cap B^C \in \Lambda$.*

Exercise 16 *Prove Propositions 14 and 15.*

Lemma 20 ($\pi - \lambda$ theorem) *Let Ω be a set and let Π be a π-system and let Λ be a λ-system that contains Π. Then $\sigma(\Pi) \subseteq \Lambda$.*

Proof: Define $\lambda(\Pi)$ to be the smallest λ-system containing Π. For each $A \subseteq \Omega$, define $\mathcal{G}_A$ to be the collection of all sets $B \subseteq \Omega$ such that $A \cap B \in \lambda(\Pi)$.

First, we show that $\mathcal{G}_A$ is a λ-system for each $A \in \lambda(\Pi)$. To see this, note that $A \cap \Omega \in \lambda(\Pi)$, so $\Omega \in \mathcal{G}_A$. If $B \in \mathcal{G}_A$, then $A \cap B \in \lambda(\Pi)$, and Proposition 15 says that $A \cap B^C \in \lambda(\Pi)$, so $B^C \in \mathcal{G}_A$. Finally, $\{B_n\}_{n=1}^{\infty} \in \mathcal{G}_A$ with the B_n disjoint implies that $A \cap B_n \in \lambda(\Pi)$ with $A \cap B_n$ disjoint, so their union is in $\lambda(\Pi)$. But their union is $A \cap (\bigcup_{n=1}^{\infty} B_n)$. So $\bigcup_{n=1}^{\infty} B_n \in \mathcal{G}_A$.

Next, we show that $\lambda(\Pi) \subseteq \mathcal{G}_C$ for every $C \in \lambda(\Pi)$. Let $A, B \in \Pi$, and notice that $A \cap B \in \Pi$, so $B \in \mathcal{G}_A$. Since $\mathcal{G}_A$ is a λ-system containing Π, it must contain $\lambda(\Pi)$. It follows that $A \cap C \in \lambda(\Pi)$ for all $C \in \lambda(\Pi)$. If $C \in \lambda(\Pi)$, it then follows that $A \in \mathcal{G}_C$. So, $\Pi \subseteq \mathcal{G}_C$ for all $C \in \lambda(\Pi)$. Since $\mathcal{G}_C$ is a λ-system containing Π, it must contain $\lambda(\Pi)$.

Finally, if $A, B \in \lambda(\Pi)$, we just proved that $B \in \mathcal{G}_A$, so $A \cap B \in \lambda(\Pi)$ and hence $\lambda(\Pi)$ is also a π-system. By Proposition 14, $\lambda(\Pi)$ is a σ-field containing Π and hence must contain $\sigma(\Pi)$. Since $\lambda(\Pi) \subseteq \Lambda$, the proof is complete. ■

The uniqueness theorem is the following.

Theorem 21 (Uniqueness theorem) *Suppose that μ_1 and μ_2 are measures on $(\Omega, \mathcal{F})$ and $\mathcal{F} = \sigma(\Pi)$, for a π-system Π. If μ_1 and μ_2 are both σ-finite on Π and they agree on Π, then they agree on $\mathcal{F}$.*

Proof: First, let $C \in \Pi$ be such that $\mu_1(C) = \mu_2(C) < \infty$, and define $\mathcal{G}_C$ to be the collection of all $B \in \mathcal{F}$ such that $\mu_1(B \cap C) = \mu_2(B \cap C)$. It is easy to see that $\mathcal{G}_C$ is a λ-system that contains Π, hence it equals $\mathcal{F}$ by Lemma 20. (For example, if $B \in \mathcal{G}_C$,

$$\mu_1(B^C \cap C) = \mu_1(C) - \mu_1(B \cap C) = \mu_2(C) - \mu_2(B \cap C) = \mu_2(B^C \cap C),$$

so $B^C \in \mathcal{G}_C$.)

Since μ_1 and μ_2 are σ-finite, there exists a sequence $\{C_n\}_{n=1}^{\infty} \in \Pi$ such that $\mu_1(C_n) = \mu_2(C_n) < \infty$, and $\Omega = \bigcup_{n=1}^{\infty} C_n$. (Since Π is only a π-system, we cannot assume that the C_n are disjoint.) For each $A \in \mathcal{F}$,

$$\mu_j(A) = \lim_{n\to\infty} \mu_j \left(\bigcup_{i=1}^{n} [C_i \cap A] \right) \text{ for } j = 1, 2.$$

Since $\mu_j\left(\bigcup_{i=1}^{n} [C_i \cap A]\right)$ can be written as a linear combination of values of μ_j at sets of the form $A \cap C$, where $C \in \Pi$ is the intersection of finitely many of $C_1, \ldots, C_n$, it follows from $A \in \mathcal{G}_C$ that $\mu_1\left(\bigcup_{i=1}^{n} [C_i \cap A]\right) = \mu_2\left(\bigcup_{i=1}^{n} [C_i \cap A]\right)$ for all n, hence $\mu_1(A) = \mu_2(A)$. ■

Exercise 17 *Return to Example 18. You should now be able to answer the question posed there.*

Exercise 18 *Suppose that $\Omega = \{a, b, c, d, e\}$ and I tell you the value of $P(\{a, b\})$ and $P(\{b, c\})$. For which subset of Ω do I need to define $P(\cdot)$ in order to have a unique extension of P to a σ-field of subsets of Ω?*

5 Lebesgue Measure and Caratheodory's Extension Theorem

Let F be a cdf (nondecreasing, right-continuous, limits equal 0 and 1 at $-\infty$ and ∞ respectively). Let $\mathcal{U}$ be the field in Example 6 (unions of finitely many disjoint intervals). Define $\mu : \mathcal{U} \to [0,1]$ by $\mu(A) = \sum_{k=1}^{n} F(b_k) - F(a_k)$ when $A = \bigcup_{k=1}^{n}(a_k, b_k]$ and $\{(a_k, b_k]\}$ are disjoint. This set-function is well-defined and finitely additive.

Is μ countably additive as probabilities are supposed to be? That is, if $A = \bigcup_{i=1}^{\infty} A_i$ where the A_i's are disjoint, each A_i is a union of finitely many disjoint intervals, and A itself is the union of finitely many disjoint intervals $(a_k, b_k]$ for $k = 1, \ldots, n$, does $\mu(A) = \sum_{i=1}^{\infty} \mu(A_i)$? First, take the collection of intervals that go into all of the A_i's and split them, if necessary, so that each is a subset of at most one of the $(a_k, b_k]$ intervals. Then apply the following result to each $(a_k, b_k]$.

Lemma 22 *Let* $(a, b] = \bigcup_{k=1}^{\infty}(c_k, d_k]$ *with the* $(c_k, d_k]$*'s disjoint. Then* $F(b) - F(a) = \sum_{k=1}^{\infty} F(d_k) - F(c_k)$.

Proof: Since $(a, b] \supseteq \bigcup_{k=1}^{n}(c_k, d_k]$ for all n, it follows that $F(b) - F(a) \geq \sum_{k=1}^{n} F(d_k) - F(c_k)$ (because (c_k, d_k)'s are disjoint), hence $F(b) - F(a) \geq \sum_{k=1}^{\infty} F(d_k) - F(c_k)$. We need to prove the opposite inequality.

Suppose first that both a and b are finite. Let $\epsilon > 0$. For each k, there is $e_k > d_k$ such that

$$F(d_k) \leq F(e_k) \leq F(d_k) + \frac{\epsilon}{2^k}.$$

Also, there is $f > a$ such that $F(a) \geq F(f) - \epsilon$. Now, the interval $[f, b]$ is compact and $[f, b] \subseteq \bigcup_{k=1}^{\infty}(c_k, e_k)$. So there are finitely many (c_k, e_k)'s (suppose they are the first n) such that $[f, b] \subseteq \bigcup_{k=1}^{n}(c_k, e_k)$. Now,

$$F(b) - F(a) \leq F(b) - F(f) + \epsilon \leq \epsilon + \sum_{k=1}^{n} F(e_k) - F(c_k) \leq 2\epsilon + \sum_{k=1}^{n} F(d_k) - F(c_k).$$

Here we have to work with finitely many (c_k, e_k)'s because we do not yet have countable sub-additivity. It follows that $F(b) - F(a) \leq 2\epsilon + \sum_{k=1}^{\infty} F(d_k) - F(c_k)$. Since this is true for all $\epsilon > 0$, it is true for $\epsilon = 0$.

If $-\infty = a < b < \infty$, let $g > -\infty$ be such that $F(g) < \epsilon$. The above argument shows that

$$F(b) - F(g) \leq \sum_{k=1}^{\infty} F(d_k \vee g) - F(c_k \vee g) \leq \sum_{k=1}^{\infty} F(d_k) - F(c_k).$$

Since $\lim_{g \to -\infty} F(g) = 0$, it follows that $F(b) \leq \sum_{k=1}^{\infty} F(d_k) - F(c_k)$. Similar arguments work when $a < b = \infty$ and $-\infty = a < b = \infty$. ∎

In Lemma 22 you can replace F by an arbitrary nondecreasing right-continuous function with only a bit more effort. (See the supplement following at the end of this lecture.)

The function μ defined in terms of a nondecreasing right-continuous function is a measure on the field $\mathcal{U}$. There is an extension theorem that gives conditions under which a measure on a field can be extended to a measure on the generated σ-field. Furthermore, the extension is unique.

Example 23 (Lebesgue measure) *Start with the function $F(x) = x$, form the measure μ on the field $\mathcal{U}$ and extend it to the Borel σ-field. The result is called* Lebesgue measure, *and it extends the concept of "length" from intervals to more general sets.*

Example 24 *Every distribution function for a random variable has a corresponding probability measure on the real line.*

Theorem 25 (Caratheodory Extension) *Let μ be a σ-finite measure on the field $\mathcal{C}$ of subsets of Ω. Then μ has a unique extension to a measure on $\sigma(\mathcal{C})$.*

Exercise 19 *In this exercise, we prove Theorem 25. Note that the uniqueness of the extension is a direct consequence of Theorem 21. We only need to prove the existence.*

First, for each $B \in 2^\Omega$, define

$$\mu^*(B) = \inf \sum_{i=1}^{\infty} \mu(A_i), \tag{20}$$

where the inf *is taken over all $\{A_i\}_{i=1}^{\infty}$ such that $B \subseteq \bigcup_{i=1}^{\infty} A_i$ and $A_i \in \mathcal{C}$ for all i. Since $\mathcal{C}$ is a field, we can assume that the A_i's are mutually disjoint without changing the value of $\mu^*(B)$. Let*

$$\mathcal{A} = \{B \in 2^\Omega : \mu^*(C) = \mu^*(C \cap B) + \mu^*(C \cap B^C), \text{ for all } C \in 2^\Omega\}.$$

Now take the following steps:

1. *Show that μ^* extends μ, i.e. that $\mu^*(A) = \mu(A)$ for each $A \in \mathcal{C}$.*
2. *Show that μ^* is monotone and subadditive.*
3. *Show that $\mathcal{C} \subseteq \mathcal{A}$.*
4. *Show that $\mathcal{A}$ is a field.*
5. *Show that μ^* is finitely additive on $\mathcal{A}$.*
6. *Show that $\mathcal{A}$ is a σ-field.*
7. *Show that μ^* is countably additive on $\mathcal{A}$.*

5.1 Extension to $\mathbb{R}^k$

The Borel σ-field on $\mathbb{R}^k$, denoted with $\mathcal{B}^k$, is generated by the class of hyper-rectangles of the form

$$A = \left\{x \in \mathbb{R}^k \colon a_i < x_i \leq b_i, i = 1, \ldots, k\right\} = (a_1, b_1] \times \ldots \times (a_k, b_k].$$

Above, $-\infty \leq a_i < b_i$ for all $i = 1, \ldots, k$.

Let $F \colon \mathbb{R}^k \to \mathbb{R}$ be a function. Typically F takes values in $[0, 1]$ or a bounded interval but this is not necessary. For each hyper-rectangle A let

$$\Delta_A F = \sum_{v \in V_A} \operatorname{sgn}(v) F(v),$$

where $V_A = \{a_1, b_1\} \times \ldots \times \{a_k, b_k\}$ is the set of vertices of A and, for any $v \in V_A$, $\operatorname{sgn}(v)$ is -1 or 1 depending on whether v has an odd or even number of a's.

Assume that F is a distribution function, i.e. that it satisfies the properties:

1. F is right continuous: if $x_n \downarrow x$ in $\mathbb{R}^k$ (meaning that $x_1 \geq x_2 \geq \ldots x_k \geq \ldots \to x$) then $F(x_n) \downarrow F(x)$;
2. F is non-decreasing: $\Delta_A F \geq 0$ for all hyper-rectangles A.

Carathedory extension's theorem allows for the general following construction of measure on $(\mathbb{R}^k, \mathcal{B}^k)$ from distribution functions.

Theorem 26 *Let F be a distribution function in $\mathbb{R}^k$ and for a hyper-rectangle A, set $\mu(A) = \Delta_A F$. Then, μ has a unique extension to a measure on $\mathcal{B}^k$.*

Example 27 *The Lebesgue measure on $(\mathbb{R}^k, \mathcal{B}^k)$ is the measure corresponding to the distribution function*

$$F(x) = \prod_{i=1}^k x_i, \quad x \in \mathbb{R}^k.$$

It is can be seen that F is right-continuous (as it is continuous) and non-decreasing, since, for any hyper-rectangle $A = (a_1, b_1] \times \ldots \times (a_k, b_k]$,

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

In particular, the Lebesgue measure of any Borel set A (not just a hyper-rectangle) coincides with its volume.

Supplement: Measures from Increasing Functions

Lemma 22 deals only with functions F that are cdf's. Suppose that F is an unbounded nondecreasing function that is continuous from the right. If $-\infty < a < b < \infty$, then the proof of Lemma 22 still applies. Suppose that $(-\infty, b] = \bigcup_{k=1}^{\infty}(c_k, d_k]$ with $b < \infty$ nd all $(c_k, d_k]$ disjoint. Suppose that $\lim_{x\to-\infty} F(x) = -\infty$. We want to show that $\sum_{k=1}^{\infty} F(d_k) - F(c_k) = \infty$. If one $c_k = -\infty$, the proof is immediate, so assume that all $c_k > -\infty$. Then there must be a subsequence $\{k_j\}_{j=1}^{\infty}$ such that $\lim_{j\to\infty} c_{k_j} = -\infty$. For each j, let $\{(c'_{j,n}, d'_{j,n}]\}_{n=1}^{\infty}$ be the subsequence of intervals that cover $(c_{k_j}, b]$. For each j, the proof of Lemma 22 applies to show that

$$F(b) - F(c_{k_j}) = \sum_{n=1}^{\infty} F(d'_{j,n}) - F(c'_{j,n}). \tag{21}$$

As $j \to \infty$, the left side of Equation 21 goes to ∞ while the right side eventually includes every interval in the original collection.

A similar proof works for an interval of the form (a, ∞) when $\lim_{x\to\infty} F(x) = \infty$. A combination of the two works for $(-\infty, \infty)$.