STA 711: Probability & Measure Theory
Robert L. Wolpert

2 Construction & Extension of Measures

For any finite set $\Omega = \{\omega_1, ..., \omega_n\}$, the "power set" $\mathfrak{P}(\Omega)$ is the collection of all subsets of Ω, including the empty-set $\emptyset$ and Ω itself. It has $|\mathfrak{P}| = 2^n$ elements; it can also be identified with the set of all possible *functions* $a : \Omega \to \{0, 1\}$ by the relation $A = \{\omega : a(\omega) = 1\}$. Set theorists denote the power set by $\mathfrak{P}(\Omega) = \{0, 1\}^\Omega$ or more simply by 2^Ω, even for infinite sets Ω. The function $a := \mathbf{1}_A$ equal to one if $a \in A$ and otherwise zero is the "indicator" function of A.

We will want to assign probabilities to as many subsets of Ω as possible (so we can find probabilities of a wide range of events) while actually *specifying* probabilities on as small a class of sets as possible (to minimize how much work we do). For a finite probability space Ω with $n \in \mathbb{N}$ elements, for example, we will see below that we need specify only the n probabilities $\{\mathsf{P}[\{\omega\}] : \omega \in \Omega\}$ to determine $\mathsf{P}(A)$ uniquely for all 2^n elements $A \in 2^\Omega$. Since $n \ll 2^n$ for big n, this is a bargain.

Let's consider a number of properties that classes of sets $\mathcal{A} \subset 2^\Omega$ might have. A class $\mathcal{A}$ of subsets of Ω is called a:

FIELD	if	$F_1:$	$\Omega \in \mathcal{A}$
		$F_2:$	$E \in \mathcal{A} \Rightarrow E^c \in \mathcal{A}$
		$F_3:$	$E_1, E_2 \in \mathcal{A} \Rightarrow E_1 \cup E_2 \in \mathcal{A}$.
σ-**FIELD**	if	$\sigma_1:$	$\Omega \in \mathcal{A}$
		$\sigma_2:$	$E \in \mathcal{A} \Rightarrow E^c \in \mathcal{A}$
		$\sigma_3:$	$\{E_i\} \subset \mathcal{A} \Rightarrow \cup E_i \in \mathcal{A}$.
π-**SYSTEM**	if	$\pi_1:$	$E_1, E_2 \in \mathcal{A} \Rightarrow E_1 \cap E_2 \in \mathcal{A}$.
λ-**SYSTEM**	if	$\lambda_1:$	$\Omega \in \mathcal{A}$
		$\lambda_2:$	$E \in \mathcal{A} \Rightarrow E^c \in \mathcal{A}$
		$\lambda_3:$	$\{E_i\} \subset \mathcal{A},\ E_i \cap E_j = \emptyset \Rightarrow \cup E_i \in \mathcal{A}$.

Note that if $\mathcal{A}_\alpha$ is a (F, σF, πS, resp. λS) for each α in any index set (even an uncountable one), then $\cap_\alpha \mathcal{A}_\alpha$ is also a (F, σF, πS, resp. λS) (Exercise: show that this is not true for even finite unions). Since also 2^Ω is a (F, σF, πS, resp. λS), it follows that for any collection $\mathcal{A}_0 \subset 2^\Omega$ there exists a *smallest* (F, σF, πS, resp. λS) that contains $\mathcal{A}_0$: namely, the intersection of all (F, σF, πS, resp. λS)s containing $\mathcal{A}_0$. We denote the smallest (F, σF, πS, resp. λS) containing $\mathcal{A}_0$ by $\mathcal{F}(\mathcal{A}_0)$, $\sigma(\mathcal{A}_0)$, $\pi(\mathcal{A}_0)$, and $\lambda(\mathcal{A}_0)$, respectively.

For example, if Ω is arbitrary and $\mathcal{A}_0 = \{\{\omega\} : \omega \in \Omega\}$, the singletons, then $\mathcal{F}(\mathcal{A}_0) = \sigma(\mathcal{A}_0) = 2^\Omega$ if Ω is finite. If Ω is infinite, however, then $\mathcal{F}(\mathcal{A}_0)$ is the collection of finite and co-finite sets; $\sigma(\mathcal{A}_0)$ and $\lambda(\mathcal{A}_0)$ are both the collection of countable and co-countable sets; and $\pi(\mathcal{A}_0)$ is just $\{\mathcal{A}_0 \cup \{\emptyset\}\}$.

For probability and measure theory we would like for probabilities $\mathsf{P}(A)$ to be defined on all the sets $A \subset \Omega$ that we encounter. For finite or countable Ω we can usually define $\mathsf{P}(A)$ sensibly for *all* subsets A, but for uncountable Ω this typically isn't possible (see free on-line Appendices B or C of Frank Burk's text *Lebesgue Measure and Integration: An Introduction* for a nice account). If we can't define $\mathsf{P}(A)$ on all of 2^Ω, we still need probabilities to be defined for all sets in a sigma field $\mathcal{F}$, so we can compute probabilities for countable unions and intersections. We'd like the luxury of having to *specify* measures on a much smaller collection, like a field $\mathcal{F}_0$ or a collection of sets $\mathcal{C}$ that generates a field $\mathcal{F}_0 := \mathcal{F}(\mathcal{C})$. That's our goal for the next week or so.

To do this we need to know that we can always *extend* a probability assignment μ_0 defined on a field $\mathcal{F}_0$ to *exactly one* measure μ on the sigma field $\mathcal{F} = \sigma(\mathcal{F}_0)$— *i.e.*, that (a) there exists *at least one* such extention, and that (b) any two must agree on all of $\mathcal{F}$.

It turns out to be easier to show that μ_0 extends uniquely to the λ-system $\lambda(\mathcal{A}_0)$ than it is to show unique extension to the sigma field $\sigma(\mathcal{A}_0)$; luckily, when $\mathcal{A}_0$ is a *field* (or even just a π-system), these are the same. This will follow from:

2.1 Dynkin's Theorem

Theorem (Dynkin's π-λ Theorem). Let $\mathcal{P}$ be a π-system; then $\lambda(\mathcal{P}) = \sigma(\mathcal{P})$.
Proof sketch:

The proof is in two parts. First we show that $\lambda(\mathcal{P})$ is not only a λ-system, it's also a π-system; then, show that if a collection $\mathcal{L} \subset 2^\Omega$ is both a λ-system and a π-system, then it's a σ-algebra too.

Set $\mathcal{L} := \lambda(\mathcal{P})$; we must show that $\mathcal{L}$ is a σ-algebra. The first two properties are immediate:

$\sigma_1:$	$\Omega \in \mathcal{L}:$	This is λ_1.
$\sigma_2:$	$A \in \mathcal{L} \;\Rightarrow\; A^c \in \mathcal{L}:$	This is λ_2.
$\sigma_3:$	$\{A_i\} \subset \mathcal{L} \;\Rightarrow\; \cup A_i \in \mathcal{L}:$	We must show this, below.

For the third, we must turn an arbitrary countable union $\cup A_i$ into a *disjoint* one.

Let $\{A_i\} \subset \mathcal{L}$, and for $n \in \mathbb{N}$ let B_n be "what's new in A_n," *i.e.*, define

$$B_n := A_n \cap \Big(\bigcup_{i<n} B_i\Big)^c = A_n \cap \bigcap_{i<n} B_i^c. \tag{1}$$

The $\{B_n\}$ are disjoint (since each B_n is in B_i^c for each $i < n$) and, since $\cup_{i\le n} A_i = \cup_{i\le n} B_i$ for every $n \in \mathbb{N}$, the $\{B_n\}$ have the same union as $\{A_n\}$. To complete the proof we must show that $B_n \in \mathcal{L}$, for then $\cup A_n = \cup B_n \in \mathcal{L}$ by λ_3. By induction, suppose $B_i \in \mathcal{L}$ for each $i < n$. The union $\big(\cup_{i<n} B_i\big)$ in (1) lies in $\mathcal{L}$ by λ_3; its complement lies in $\mathcal{L}$ by λ_2; and A_n lies in $\mathcal{L}$ by assumption. To complete the induction step the only remaining obstacle is to show that $\mathcal{L}$ contains finite *intersections*, *i.e.*, that it is a π-system.

$\mathcal{L}$ is a π-system

Fix any $A \in \mathcal{P}$ and set $\mathcal{A} := \{B \in \mathcal{L} : \ A \cap B \in \mathcal{L}\}$. Let's show: **$\mathcal{A}$ is a λ-system**
There are three things to show for all $\{B_i\} \subset \mathcal{A}$:

λ_1 : $\Omega \in \mathcal{A}$: $A \cap \Omega = A \in \mathcal{P} \subset \mathcal{L}$.
λ_2 : $B \in \mathcal{A} \ \Rightarrow \ B^c \in \mathcal{A}$: $A \cap B^c = A \cap (A \cap B)^c = [A^c \cup (A \cap B)]^c \in \mathcal{L}$ by λ_2, λ_3.
λ_3 : $B_i \cap B_j = \emptyset \Rightarrow \cup B_i \in \mathcal{A}$: $A \cap (\cup B_i) = \cup (A \cap B_i) \in \mathcal{A}$ by λ_3.

Also $\mathcal{P} \subset \mathcal{A}$ by π_1, so $\mathcal{A}$ is a λ-system containing $\mathcal{P}$ and hence containing $\mathcal{L} = \lambda(\mathcal{P})$.

We have just shown that $A \cap B \in \mathcal{L}$ for every $A \in \mathcal{P}$ and $B \in \mathcal{L}$. So, for every $B \in \mathcal{L}$, the class

$$\mathcal{B} = \{A \in \mathcal{L} : \ A \cap B \in \mathcal{L}\}$$

contains each $A \in \mathcal{P}$. Also $\Omega \in \mathcal{B}$ (by λ_1) and $\mathcal{B}$ is closed under complements (as before) and disjoint unions, so $\mathcal{B}$ is a λ-system containing $\mathcal{P}$ and hence containing $\mathcal{L}$.

This completes the proof that $A \cap B \in \mathcal{L}$ for every $A, B \in \mathcal{L}$, *i.e.*, that $\mathcal{L}$ is a π-system. It also completes the induction step in (1) to verify that

$$\{A_i\} \subset \mathcal{L} \ \Rightarrow \ \cup A_i \in \mathcal{L},$$

completing the proof that $\lambda(\mathcal{P})$ is a σ-algebra and hence $\sigma(\mathcal{P}) \subset \lambda(\mathcal{P})$. But $\sigma(\mathcal{P})$ is also a λ-system and so $\lambda(\mathcal{P}) \subset \sigma(\mathcal{P})$, completing the proof of Dynkin's π-λ theorem that $\sigma(\mathcal{P}) = \lambda(\mathcal{P})$. □

How can this help us to extend a probability assignment (or "pre-measure") μ_0 from a π-system $\mathcal{P}$ (for example, a field) to the σ-field $\mathcal{F} = \sigma(\mathcal{P})$ it generates? First, note that λ-systems are just perfect for uniqueness:

Proposition 1 *Let P and Q be two probability measures on a space $(\Omega, \mathcal{F})$. The class*

$$\mathcal{L} = \{A \in \mathcal{F} : \ P(A) = Q(A)\}$$

is a λ-system.

Can you prove that? By Dynkin's $\pi-\lambda$ theorem, there is at most one extension of a "pre-measure" P_0 from any π-system $\mathcal{P}$ to the σ-algebra $\sigma(\mathcal{P}) = \lambda(\mathcal{P})$ it generates, because if P and Q were two different ones, the collection of events on which they agree would be a λ-system containing $\mathcal{P}$ and hence containing $\lambda(\mathcal{P}) = \sigma(\mathcal{P})$. Let's look at examples:

1. $\mathcal{P} := \big\{\{a\}\big\}$ on $\Omega = \{a, b, c\}$. To illustrate that uniqueness of extensions can fail, consider a probability assignment μ on the π-system $\mathcal{P}$ that assigns probability $\mu(\{a\}) = 1/2$. For any number $0 \le p \le \frac{1}{2}$ there exists a distinct extension μ_p of μ to the σ-algebra $\mathcal{F} = 2^\Omega$ that assigns probabilities $\mu_p(\{b\}) = p$, $\mu_p(\{c\}) = (\frac{1}{2} - p)$. For $p \ne q$, the collection of events L for which $\mu_p(L) = \mu_q(L)$ is $\mathcal{L} = \{\emptyset, \{a\}, \{b, c\}, \Omega\}$, a λ-system (and σ-algebra) strictly smaller than $\mathcal{F}$.

2. $\mathcal{P} := \{ \{\omega\} : \omega \in \Omega \} \cup \{\emptyset\}$: Given any finite or countable set $\Omega = \{\omega_i\}$ and positive numbers $\{p_i \ge 0\}$ with unit sum $\sum_i p_i = 1$, define μ_0 on $\mathcal{P}$ by setting $\mu_0(\{\omega_i\}) = p_i$ and $\mu_0(\emptyset) = 0$. Then by countable additivity the only possible probability measure on 2^Ω that extends μ_0 is $\mu(A) := \sum[p_i : \ \omega_i \in A]$. Every probability measure on 2^Ω for any finite or countable set Ω is of this form.

3. $\mathcal{P} := \{ (-\infty, b], \ b \in \mathbb{Q} \}$ on $\Omega = (-\infty, \infty)$. The *field* generated by $\mathcal{P}$ consists of finite disjoint unions of left-open rational intervals $(a, b]$, including semi-infinite intervals of the form $(-\infty, b]$ and (a, ∞), and $\Omega = (-\infty, \infty)$. The sigma field $\sigma(\mathcal{A})$ is *not* just countable unions of such sets; it is the "Borel" σ-algebra $\mathcal{B}(\mathbb{R})$ generated by the open sets in the real line and includes all open and closed sets, the Cantor set, and many others. It can be constructed explicitly by transfinite induction (!), and hence includes only $c = \#(\mathbb{R})$ elements (while the power set $2^{\mathbb{R}}$ contains $2^c > c$), but it is not easily described. It is *not* every possible subset of $\mathbb{R}$, but it includes every set of real numbers we'll need in this course.

   A "Distribution Function" (or "DF") is a right-continuous non-decreasing function on $\mathbb{R}$ with limits $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to+\infty} F(x) = 1$ at $\pm\infty$. For any DF $F(x)$, we can define a pre-pm μ_0 on $\mathcal{P}$ by setting $\mu_0((-\infty, b]) := F(b)$. If $F = F_d$ is purely discontinuous this just assigns probability $p_i = F(x_i) - F(x_i-)$ to each x_i where $F(x)$ jumps; if $F(x) = F_{ac} = \int_{-\infty}^x f(t)\,dt$ is absolutely continuous this just assigns probability $\mu(A) = \int_A f(t)\,dt$ to A (and in fact this is the usual *definition* of that integral!)

2.2 Extension 1: π-System to Field

In this section we show that any finitely-additive pre-measure defined on a π-system $\mathcal{P}$ can be extended uniquely to the field $\mathcal{F}(\mathcal{P})$ it generates. Our text instead begins with a pre-measure P_0 defined on a "semi-algebra" (a π-system $\mathcal{S}$ such that if $A \in \mathcal{S}$ then A^c is a finite disjoint union of elements of $\mathcal{S}$), but that isn't necessary— here we show that any π-system will do, since P_0 on a π-system $\mathcal{P}$ can always be extended uniquely to the field $\mathcal{F}(\mathcal{P})$ it generates. I think this simplifies the argument and the result.

Let μ_0 be a pre-pm defined on a π-system $\mathcal{P}$, and let $\mathcal{F}_0 := \mathcal{F}(\mathcal{P})$ be the field generated by $\mathcal{P}$. For example, if we have an assignment of μ_0 to all sets in

$$\mathcal{P} = \{(0, b] \ : 0 \le b \le 1\}$$

in the unit inverval $\Omega = (0, 1]$, say, $\mu_0\big((0, b]\big) := F(b)$ for some increasing function $F : \Omega \to \mathbb{R}_+$. Then by additivity we must have

$$\mu_0\big((a, b]\big) = \mu_0\big((0, b]\big) - \mu_0\big((0, a]\big) = F(b) - F(a)$$

for $0 \le a \le b \le 1$, and, for the disjoint union of such intervals,

$$\mu_0\Big(\bigcup_{j=1}^J (a_j, b_j]\Big) = \sum_{j=1}^J \big[F(b_j) - F(a_j)\big]$$

for $0 \le a_1 \le b_1 \le a_2 \le \cdots \le b_J \le 1$. But the field $\mathcal{F}_0 := \mathcal{F}(\mathcal{P})$ consists precisely of sets of that form, so μ_0 has a unique extension to $\mathcal{F}_0$. Similarly, any pre-pm μ_0 defined only on rectangles $(0,b] \times (0,d]$ in the unit square with the origin as the south-west corner (given perhaps by a function $F(b,d) = \mu_0\big((0,b] \times (0,d]\big)$ on $\Omega := (0,1]^2$) has a unique extension to the disjoint union of all rectangles $(a,b] \times (c,d] \in \Omega$; can you find an explicit expression for $\mu_0\big((a,b] \times (c,d]\big)$? Hint: First find $\mu_0\big((a,b] \times (0,d]\big)$.

In a week-2 homework exercise you will show that for *any* π-system $\mathcal{P}$ the field $\mathcal{F}_0 := \mathcal{F}(\mathcal{P})$ consists precisely of sets of the form

$$\mathcal{F}_0 = \big\{B : \ B = \cup_{i=1}^m B_i, \quad B_i = \cap_{j=1}^{n_i} A_{ij} \text{ for some } m \in \mathbb{N},\ \{n_i\} \subset \mathbb{N}\big\}$$

with each $A_{ij} \in \mathcal{P}$ or $A_{ij}^c \in \mathcal{P}$, and with the m sets $\{B_i\}$ disjoint. By induction on the number of A_{ij} with $A_{ij}^c \in \mathcal{P}$ and finite additivity, you can show that μ_0 is well defined on each B_i; by finite additivity again, it is a well-defined pre-pm on all of $\mathcal{F}_0$.

2.3 Extension 2: Field to σ-Algebra

Let μ_0 be a pre-pm defined on a field $\mathcal{F}_0$, *i.e.*, a function $\mu_0 : \ \mathcal{F}_0 \to \mathbb{R}$ that satisfies the conditions:

1. $A \in \mathcal{F}_0 \Rightarrow \mu_0(A) \ge 0$
2. $\mu_0(\Omega) = 1$
3. $\{A_i\} \subset \mathcal{F}_0$ and $\cup A_i \in \mathcal{F}_0$ and $A_i \cap A_j = \emptyset \Rightarrow \mu_0(\cup A_i) = \sum \mu_0(A_i)$.

Define two new set functions μ^* and μ_* on *all* subsets of Ω, *i.e.*, on 2^Ω, by:[1]

$$\mu^*(E) := \inf\Big[\sum_{i=0}^{\infty} \mu_0(F_i) : \ E \subset \bigcup_{i=0}^{\infty} F_i,\ F_i \in \mathcal{F}_0\Big] \qquad \mu_*(E) := 1 - \mu^*(E^c)$$

On reflection it's clear that $\mu_*(E) \le \mu^*(E)$ (or, equivalently, that $\mu^*(E) + \mu^*(E^c) \ge 1$) for each set $E \in 2^\Omega$, and $\mu_*(E) = \mu_0(E) = \mu^*(E)$ for each set $E \in \mathcal{F}_0$. Thus there is an obvious well-defined extension of μ_0 to a set function μ defined on the μ-*completion*

$$\begin{aligned}\overline{\mathcal{F}}^\mu &= \{E \in 2^\Omega : \ \mu_*(E) = \mu^*(E)\} \\ &= \{E \in 2^\Omega : \ \mu^*(E) + \mu^*(E^c) = 1\}.\end{aligned}$$

It remains to show three things:

1. The extension μ is nonnegative on $\overline{\mathcal{F}}^\mu$, with $\mu(\Omega) = 1$, and is countably additive. Showing that $\mu(\cup E_n) \le \sum \mu(E_n)$ is a simple $\epsilon/2^n$ argument; how can you show $\mu(\cup E_n) \ge \sum \mu(E_n)$ for disjoint $\{E_n\}$?

[1] Why do we need infinitely-many F_is? Why not just $\inf\,[\mu_0(F) : \ E \subset F]$? See "Examples" below.

2. The σ-field $\mathcal{F} := \sigma(\mathcal{F}_0)$ is contained in $\overline{\mathcal{F}}^\mu$ (just show that $\overline{\mathcal{F}}^\mu$ is a σF containing $\mathcal{F}_0$)
3. The extension to $\mathcal{F}$ is unique (show that for any two extensions μ_1 and μ_2, $\{E \in \mathcal{F} : \mu_1(E) = \mu_2(E)\}$ is a λ-S containing the π-S $\mathcal{F}_0$).

See Section (3) on page 9 of these notes for details, or Resnick (1999) §2.4 or Billingsley (1995), pp. 38–41. Warning: the appealing idea of defining $\mu_*(E)$ by approximating E from inside *doesn't work*— consider the inner Borel measure of the irrationals in $(0, 1]$ with $\mathcal{F}_0 = \{\cup_i (a_i, b_i]\}$. What's the μ-completion for a discrete measure μ on $\mathbb{R}$?

2.4 Completions

It is possible that the σ-algebra $\mathcal{F}$ generated by $\mathcal{F}_0$ will not be "complete", in the sense that there may exist null sets N (*i.e.*, events $N \in \mathcal{F}$ with $\mu(N) = 0$) that have *subsets* $E \subset N$ that are not events, *i.e.*, $E \notin \mathcal{F}$. The "μ-completion" $\overline{\mathcal{F}}^\mu$ of $\mathcal{F}$ is the smallest μ-complete σ-algebra containing $\mathcal{F}$, and is the largest σ-algebra to which μ may be extended unambiguously. Four characterizations of the μ-completion $\overline{\mathcal{F}}^\mu$ of a σ-field $\mathcal{F}$ for a probability (or σ-finite) measure μ on $\mathcal{F}$ are sometimes useful:

$$\begin{aligned}
\overline{\mathcal{F}}^\mu &:= \{E \in 2^\Omega : \ \mu_*(E) = \mu^*(E)\} \\
&= \{A \cup B : \ A \in \mathcal{F}, \ B \subset N \in \mathcal{F}, \ \mu(N) = 0\} \\
&= \{E \in 2^\Omega : \ \exists A, B \in \mathcal{F}, \ \text{s.t.} \ A \subset E \subset B, \ \mu(B \backslash A) = 0\} \\
&= \{E \in 2^\Omega : \ \exists A, N \in \mathcal{F}, \ \text{s.t.} \ A \Delta E \subset N, \ \mu(N) = 0\}.
\end{aligned}$$

The σ-algebra $\mathcal{F}$ will be our main focus, and not its completion $\overline{\mathcal{F}}^\mu$. One reason is that $\overline{\mathcal{F}}^\mu$ depends on μ while $\mathcal{F}$ is intrinsic. For example, the ν-completion of the Borel sets $\mathcal{B}$ on the unit interval $\Omega = (0, 1]$ for any discrete probability measure ν is $\overline{\mathcal{B}}^\nu = 2^\Omega$. The completion of the Borel sets on Ω for Lebesgue measure μ is the "Lebesgue sets" $\overline{\mathcal{B}}^\mu$, which (under the axiom of choice) satisfy the strict inclusions $\mathcal{B} \subsetneq \overline{\mathcal{B}}^\mu \subsetneq 2^\Omega$.

2.5 Examples

Lebesgue Measure of the Dyadic Rationals: $\lambda(\mathbb{Q}_2)$ =?

Consider the unit interval $\Omega = (0, 1]$ and the π-system $\mathcal{P}$ consisting of intervals $(0, q]$ for dyadic rational numbers $q \in \mathbb{Q}_2 := \{i/2^n : \ i, n \in \mathbb{N}_0 := \{0, 1, 2, ...\}, \ i \le 2^n\}$. The field $\mathcal{F}(\mathcal{P})$ generated by $\mathcal{P}$ consists of all finite unions $\cup(a_i, b_i]$ of half-open intervals with dyadic rational end-points $0 \le a_i \le b_i \le 1$. One can show (Resnick does so in §2.5.1) that the set function $\mu_0\big((0, q]\big) := q$ on $\mathcal{P}$ extends to a countibly additive set function μ on $\mathcal{F}$. What is the outer measure $\mu^*(\mathbb{Q}_2)$? Note here that Ω contains real numbers, not just rationals.

Any *finite* cover $\cup_{i \le n} F_i$ of $\mathbb{Q}_2$ with elements of $\mathcal{F}$ would also cover $\Omega = (0, 1]$ and so would have $\sum \mu(F_i) \ge 1$; does it follow that $\mu^*(\mathbb{Q}_2) \ge 1$????

Well, no. Since $\mathbb{Q}_2$ is countable, we can enumerate it as $\{q_n : n \in \mathbb{N}\}$ and for any rational $\epsilon > 0$ we can cover $\mathbb{Q}_2$ with the infinite union $\cup F_n$ where $F_n = (a_n, b_n]$ with $b_n = q_n$ and $a_n = \max(0, q_n - \epsilon/2^n)$, with total length

$$\sum \mu(F_n) = \sum_n [q_n - \max(0, q_n - \epsilon/2^n)] = \sum_n \min(q_n, \epsilon/2^n) \le \sum_n \epsilon/2^n = \epsilon.$$

Since $\mu(\mathbb{Q}_2) \le \epsilon$ for every $\epsilon > 0$, necessarily $\mu(\mathbb{Q}_2) = 0$. This example illustrates why we need infinite covers in the definition of μ^*.

Uniform Distribution on $\mathbb{N}$?

Let $\Omega = \mathbb{N}$ be the natural numbers $\{1, 2, 3, ...\}$, E and E^c the even and odd ones respectively, and set

$$\begin{aligned} F &:= \cup_{k=0}^{\infty} \{2^{2k}, ..., 2^{2k+1} - 1\} \\ &= \{1, \quad 4, ..., 7, \quad 16, ..., 31, \quad 64, ..., 127, \quad 256, ..., 511, \quad ...\}. \end{aligned}$$

Notice that:

1. For $n = 2^{2k} - 1$, the ratio $P_n(F) = \#[F \cap \{1, ..., n\}]/n$ is exactly $P_n(F) = 1/3$, while for $n = 2^{2k+1} - 1$ it is $P_n(F) = 2/3$. Thus $P_n(F)$ cannot possibly converge as $n \to \infty$.

2. The even portion $A := F \cap E$ of F and odd portion $B := F^c \cap E^c$ of F^c both have relative frequencies ranging from 1/6 to 1/3, which also cannot converge. In fact, $A = F \cap E$ is exactly the same as the set $2 \times (F^c)$, while $B = F^c \cap E^c$ is exactly the same as the set $2 \times F + 1$.

3. $C := (A \cup B)$ however DOES have an asymptotic frequency— in fact, $|P_n(C) - \frac{1}{2}| \le \frac{1}{n}$ for all $n \in \mathbb{N}$, so $P_n(C) \to 1/2$ as $n \to \infty$.

4. Thus E and C both have well-defined asymptotic frequencies (each is 1/2), but $A = E \cap C$ does not.

Thus, the collection of sets S for which $\lim_{n\to\infty} P_n(S)$ converges is not even a *field*, let alone a σ-field, and there does not exist a uniform probability distribution on the integers.

Uniform Distribution on $(0,1]^n$

Earlier (Example 3 on page 4) we constructed a measure μ on the σ-algebra $\mathcal{F} = \sigma(\mathcal{F}_0)$ generated by a field $\mathcal{F}_0$ of subsets of the real line $\Omega = \mathbb{R}$ based on a DF $F(x)$. The same approach works more generally, starting with a set assignment μ_0 on any field $\mathcal{F}_0$ or, slightly more generally, on any π-system. Any set function $\mu_0 : \mathcal{A} \to \mathbb{R}$ satisfying (1) $\mu_0(A) \geq 0$, (2) $\mu_0(\Omega) = 1$, and (3) $\mu_0(\cup A_j) = \sum \mu_0(A_j)$ if $A_j \in \mathcal{A}$, $A_i \cap A_j = \emptyset$, *and* $\cup A_j \in \mathcal{A}$, has a unique extension to a probability measure μ on $\sigma(\mathcal{A})$.

In particular this lets us construct Lebesgue measure $m(dx)$ on the unit cube in $\mathbb{R}^n$, so we can explore some of its properties. We constructed *probability* measures P on $(\Omega, \mathcal{F})$ that satisfy the three rules

1. $\mathsf{P}(A) \geq 0$ for each $A \in \mathcal{F}$;
2. For disjoint $\{A_i \in \mathcal{F}\}$, $\mathsf{P}\big(\cup_i A_i\big) = \sum_i \mathsf{P}(A_i)$;
3. $\mathsf{P}(\Omega) = 1$.

Extensions

The same approach would let us construct **finite positive measures** μ on a set Ω and σ-algebra $\mathcal{F}$ by replacing condition 3. with "$\mu(\Omega) < \infty$". By piecing these together we can construct **sigma finite positive measures** where we replace condition 3. with "$\Omega = \cup_i A_i$ with each $A_i \in \mathcal{F}$ and $\mu(A_i) < \infty$." In particular, we can construct Lebesgue measure $m(dx)$ on all of $\mathbb{R}^n$.

The m-completion $\overline{\mathcal{F}}^m$ of the Borel σ-algebra $\mathcal{F}$ is called the "Lebesgue σ-algebra" on $\mathbb{R}^n$; it contains $\mathcal{F}$ and has the property of *completeness*, *i.e.*, that $N \in \overline{\mathcal{F}}^m$ and $m(N) = 0$ imply that $A \in \overline{\mathcal{F}}^m$ and $m(A) = 0$ for every $A \subseteq N$. The question of whether or not it coincides with 2^Ω is delicate (it depends on the Axiom of Choice) and won't concern us in this course, but you can find more with google (for example, your search should discover Appendices B or C of Frank Burk's text *Lebesgue Measure and Integration: An Introduction*). You can also ask me outide pf class if you're interested.

3 Countable Additivity of Outer Measure μ^* on $\overline{\mathcal{F}}^\mu$

Let μ_0 be countably additive on a field $\mathcal{F}_0$ on a space Ω and, for *all* subsets $E \subseteq \Omega$, define the *outer measure* μ^* and *inner measure* μ_* by

$$\mu^*(E) := \inf\Big[\sum_{i=0}^{\infty} \mu_0(F_i) : E \subset \bigcup_{i=0}^{\infty} F_i,\ \{F_i\} \subset \mathcal{F}_0\Big] \qquad \mu_*(E) := 1 - \mu^*(E^c)$$

and the μ-*completion* of $\mathcal{F}_0$,

$$\overline{\mathcal{F}}^\mu = \{E \in 2^\Omega : \mu_*(E) = \mu^*(E)\} = \{E \in 2^\Omega : \mu^*(E) + \mu^*(E^c) = 1\}$$

on which we define $\mu(E) := \mu^*(E) = \mu_*(E)$. Evidently μ "extends" μ_0 in the sense that $\mathcal{F}_0 \subset \overline{\mathcal{F}}^\mu$ and, for any $A \in \mathcal{F}_0$, we have $\mu_0(A) = \mu(A)$. It is also clear that μ is (1) nonnegative on $\overline{\mathcal{F}}^\mu$ and (2) satisfies $\mu(\Omega) = 1$; here we verify that (3) μ is countably additive on $\overline{\mathcal{F}}^\mu$.

Let $\{E_n\} \subset \overline{\mathcal{F}}^\mu$ be disjoint, and set $E := \cup_n E_n$. We will show that $\mu(E) = \sum \mu(E_n)$ in two steps. First, the easy direction:

1. $\mu(E) \le \sum \mu(E_n)$

Fix $\epsilon > 0$ and, for each n, find $\{F_{ni}\} \subset \mathcal{F}_0$ with $E_n \subset \cup_i F_{ni}$ and

$$\mu^*(E_n) \le \sum_i \mu_0(F_{ni}) < \mu^*(E_n) + 2^{-n}\epsilon \tag{2}$$

Then $E := \cup_n E_n \subset \cup_{n,i} F_{ni}$ and

$$\mu^*(E) \le \sum_{n,i} \mu_0(F_{ni}) < \sum_n \mu^*(E_n) + \epsilon$$

verifying $\mu^*(E) \le \sum_n \mu(E_n)$.

2. $\mu(E) \ge \sum \mu(E_n)$

Still $\{E_n\} \subset \overline{\mathcal{F}}^\mu$ are disjoint, and $E := \cup_n E_n$. Fix $\epsilon > 0$ and $N \in \mathbb{N}$ (suggestion: work through the case $N = 2$ first, and draw pictures). For each $n \le N$ find $\{F_{nj}\} \subset \mathcal{F}_0$ with $E_n^c \subset \cup_j F_{nj}$ and

$$\mu^*(E_n^c) \le \sum_j \mu_0(F_{nj}) < \mu^*(E_n^c) + \epsilon/N \tag{3}$$

and, similarly, find $\{G_j\} \subset \mathcal{F}_0$ with $E \subset \cup_j G_j$ and

$$\mu^*(E) \le \sum_j \mu_0(G_j) < \mu^*(E) + \epsilon. \tag{4}$$

For each fixed n, $\cup_j F_{nj}$ covers every point outside E_n at least once, so $\cup_{n,j} F_{nj}$ covers every point outside $\cup_{n=1}^N E_n$ at least N times, and every point in Ω at least $(N-1)$ times. Since $\cup_j G_j$ covers every point inside $\cup_{n=1}^N E_n \subset E$ once, the union $\big(\cup_{n,j} F_{nj}\big) \bigcup \big(\cup_j G_j\big)$ covers every point in Ω at least N times and, since $\mu^*(\Omega) = 1$, we have

$$\begin{aligned}
N &\le \sum_{n=1}^N \sum_j \mu_0(F_{nj}) \; + \sum_j \mu_0(G_j) \\
&\le \sum_{n=1}^N \mu^*(E_n^c) + \epsilon \;\; + \mu^*(E) + \epsilon \\
&= N - \sum_{n=1}^N \mu_*(E_n) + \mu^*(E) + 2\epsilon \\
&= N - \sum_{n=1}^N \mu^*(E_n) + \mu^*(E) + 2\epsilon
\end{aligned}$$

so

$$\mu^*(E) \ge \sum_{n=1}^N \mu^*(E_n) - 2\epsilon$$

for every $N \in \mathbb{N}$ and every $\epsilon > 0$, hence

$$\mu^*(E) \ge \sum_{n=1}^\infty \mu^*(E_n)$$

completing the proof.

4 Explicit Construction of Sigma Fields [optional]

Ordinals and Transfinite Induction

Every finite set S (say, with $n < \infty$ elements) can be *totally ordered* $a_1 \prec a_2 \prec a_3 \prec \cdots \prec a_n$ in $n!$ ways, but in some sense every one of these is the same— if $\prec_1$ and $\prec_2$ are two orderings, there exists a 1–1 order-preserving isomorphism $\varphi : (S, \prec_1) \longleftrightarrow (S, \prec_2)$. Thus *up to isomorphism* there is only one ordering for any finite set.

For countably infinite sets there are many different orderings. The obvious one is $a_1 \prec a_2 \prec a_3 \prec \cdots$, ordered just like the positive integers $\mathbb{N}$; this ordering is called ω, the first *limit ordinal.* But we could pick any element (say, $b_1 \in S$) and order the remainder of S in the usual way, but declare $a_n \prec b_1$ for every $n \in \mathbb{N}$; one element is "bigger" (in the ordering) than all the others. This is *not* isomorphic to ω, and it is called $\omega + 1$, the *successor* to ω. If we set aside two elements (say, $b_1 \prec b_2$) to follow all the others we have $\omega + 2$, and similarly we have $\omega + n$ for each $n \in \mathbb{N}$. The limit of all these is $\omega + \omega$, or 2ω... it is the ordering we would get if we lexicographically ordered the set $\{(i, j) : i = 1, 2\ j \in \mathbb{N}\}$ of the first two rows of integers in the first quadrant, declaring $(1, j) \prec (2, k)$ for every j, k and otherwise $(i, j) \prec (i, k)$ if $j < k$.

We would get the successor to this, $2\omega + 1$, by extending the lexicographical ordering as we add $(3, 1)$ to S; in an obvious way we get $2\omega + n$ for every $n \in \mathbb{N}$ and eventually the limit ordinals 3ω, 4ω, *etc.*, and the successor ordinals $m\omega + n$. The limit of all these is $\omega\omega$ or ω^2, the lexicographical ordering of the entire first quadrant of integers (i, j). It too has successors $\omega^2 + n$ (graphically you can think about integer triplets (i, j, k)), and limits like $\omega^2 + \omega$ and ω^3 and ω^ω (which turns out to be the same as 2^ω).

In general an ordinal is a *successor* ordinal if it has a maximal element, and otherwise is a *limit* ordinal. Every ordinal α has a successor $\alpha + 1$, and every set of ordinals $\{\alpha_n\}$ has a limit (least upper bound) λ. Let Ω be the first *un*countable ordinal.

Proofs and constructions by *transfinite induction* typically have one step at each successor ordinal, and another at each limit ordinal. The *Borel sets* can be defined by transfinite construction as follows.

Let $\mathcal{F}_0$ be the class of open subsets of some topological space $\mathcal{X}$ (perhaps the real numbers $\mathcal{X} = \mathbb{R}$, for example).

Succ: For any ordinal α, let $\mathcal{F}_{\alpha+1}$ be the class of countable unions of sets $E_n \in \mathcal{F}_\alpha$ and their complements $E_m : E_m^c \in \mathcal{F}_\alpha$.

Lim: For any limit ordinal λ, let $\mathcal{F}_\lambda = \cup_{\alpha \prec \lambda} \mathcal{F}_\alpha$.

Together these define a nested family $\mathcal{F}_\alpha$ for all ordinals, limit and successor, with $\alpha \prec \beta \Rightarrow \mathcal{F}_\alpha \subset \mathcal{F}_\beta$. The sigma field *generated by* $\mathcal{F}_0$ is $\mathcal{F}_\Omega$, where Ω is the first uncountable ordinal. It remains to prove that:

1. $\mathcal{F}_0 \subset \mathcal{F}_\Omega$, *i.e.*, $\mathcal{F}_\Omega$ contains the open sets (including $\mathcal{X}$ itself);
2. $E \in \mathcal{F}_\Omega \Longrightarrow E^c \in \mathcal{F}_\Omega$, *i.e.*, $\mathcal{F}_\Omega$ is closed under complements;
3. $\{E_n\} \subset \mathcal{F}_\Omega \Longrightarrow \cup_{n=1}^\infty E_n \in \mathcal{F}_\Omega$, *i.e.*, $\mathcal{F}_\Omega$ is closed under countable unions;
4. $\mathcal{F}_\Omega \subset \mathcal{G}$ for any sigma field $\mathcal{G}$ containing $\mathcal{F}_0$.

Item 1. is trivial since $\mathcal{F}_\Omega = \cup_{\alpha \prec \Omega} \mathcal{F}_\alpha$, and in particular contains $\mathcal{F}_0$. Item 2. follows by noting that $E \in \mathcal{F}_\alpha \Longrightarrow E^c \in \mathcal{F}_{\alpha+1}$. Item 3 follows by noting that $E_n \in \mathcal{F}_\Omega \Longrightarrow E_n \in \mathcal{F}_{\alpha_n}$ for some $\alpha_n \prec \Omega$, and $\beta := \sup_{n<\infty} \alpha_n$ is an ordinal satisfying $\alpha_n \preceq \beta \prec \Omega$. Hence $E_n \in \mathcal{F}_\beta$ for all n and $\cup_{n=1}^\infty E_n \in \mathcal{F}_{\beta+1} \subset \mathcal{F}_\Omega$. Verifying the minimality condition Item 4 is left as an exercise.

It isn't immediately obvious from the construction that we couldn't have stopped earlier—for example, that $\mathcal{F}_2$ or $\mathcal{F}_\omega$ isn't already the Borel sets, unchanging as we allow successively more intersections and unions. In fact that does happen if the original space $\mathcal{X}$ is countable or finite; in the case of $\mathbb{R}$, however, one can show that $\mathcal{F}_\alpha \neq \mathcal{F}_{\alpha+1}$ for every $\alpha \prec \Omega$.

Do you think this explicit construction is clearer or more complicated than the completion argument used in the text?