Lecture 2

Recap: A measure space consists of:

- A set S ("universe")
- A σ-algebra Σ of subsets
  - $\emptyset \in \Sigma$
  - $A \in \Sigma \Rightarrow A^c \in \Sigma$
  - $A_i \in \Sigma \ \forall i \in \mathbb{N} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \Sigma$
- A measure $\mu : \Sigma \to [0, \infty]$
  - $\mu(\emptyset) = 0$
  - $\mu\left(\bigcup_{i=0}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$ for all p.w. disjoint $A_i \in \Sigma$.

We will mostly consider probability spaces (i.e. $\mu(S) = 1$)

Generated σ-algebras.

Given any subset $\mathcal{A} \subseteq \mathcal{P}(S)$, the σ-algebra generated by $\mathcal{A}$, denoted $\sigma(\mathcal{A})$ or $\langle \mathcal{A} \rangle$ is the smallest σ-algebra containing $\mathcal{A}$.

Formally, $\sigma(\mathcal{A}) = \bigcap_{\substack{\Sigma : \sigma \text{ alg} \\ \text{with } \mathcal{A} \subseteq \Sigma}} \Sigma$.

This is a σ-algebra:

1) $\emptyset \in \Sigma$ for all σ-algebras, so $\emptyset \in \bigcap \Sigma$. ✓

2) If $A \in \sigma(\mathcal{A})$ then $A \in \Sigma$ for all such σ-alg.
Then $A^c \in \Sigma$ for all such Σ and $A^c \in \bigcap \Sigma$. ✓

3) If $A_n \in \sigma(\mathcal{A})$ for all n, then $A_n \in \Sigma$ for all n and Σ. But then $\bigcup A_n \in \Sigma$ for all such Σ and $\bigcup A_n \in \sigma(\mathcal{A})$. ✓

Example (most important):
The Borel σ-algebra $\mathcal{B}(S) = \sigma(\{A \subseteq S : A \text{ is open}\})$.
is generated by the open subsets of S.

$\mathcal{B}(\mathbb{R}) = \sigma(\{U \subseteq \mathbb{R} : U \text{ open}\})$

There are many other subsets $\mathcal{A} \subseteq \mathcal{P}(\mathbb{R})$ that generate $\mathcal{B}(\mathbb{R})$. For example

$\mathcal{B}(\mathbb{R}) = \sigma(\{(a,b) : a \leq b\})$ or

$\mathcal{B}(\mathbb{R}) = \sigma(\{[-\infty, a) : a \in \mathbb{R}\})$

$$\mathcal{B}(\mathbb{R}) = \sigma(\{F \subseteq \mathbb{R} : F \text{ closed}\})$$

$$= \sigma(\{(q_1, q_2) : q_1 < q_2 \,;\, q_1, q_2 \in \mathbb{Q}\})$$

(countable!)

(why?)

Example (finite)

Take $S = \{1, 2, .., 10\}$ and $\mathcal{A} = \{\{1,2\}, \{5\}\}$

Then $\sigma(\mathcal{A}) = \{\emptyset, \{1,2\}, \{5\}, \{3,4,5,6,7,8,9,10\},$
$\{1,2,3,4,6,7,8,9,10\}, \{1,2,5\}, \{3,4,6,7,8,9,10\},$
$\{1,2,3,4,5,6,7,8,9,10\}\}$

We can write this as $\{\emptyset, A, B, C, A\cup B, A\cup C, B\cup C, A\cup B\cup C\}$

where $A = \{1,2\}$, $B = \{5\}$, $C = (A\cup B)^c = \{3,4,6,7,8,9,10\}$.

The notion of generated σ-algebras is useful as per this theorem

Th$^{\underline{m}}$ Suppose $\mathcal{A}$ is a π system (i.e. closed under finite intersections). Suppose further that μ_1, μ_2 are measures on $(S, \sigma(\mathcal{A}))$ such that

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

In other words, μ is uniquely determined by any π-system $\mathcal{A} \subseteq \mathcal{P}(S)$.

Example: $\mathcal{B}(\mathbb{R}) = \mathcal{B}(\{(-\infty, a] : a \in \mathbb{R}\})$

Hence every probability measure $\mathbb{P}$ on $\mathbb{R}$ is determined by the values of $\mathbb{P}((-\infty, a])$, i.e. its cumulative distribution function F.

The following important theorem tells us that measures can be constructed from "small" collections of subsets.

<u>Caratheodory's extension theorem</u>:

If Σ_0 is an algebra and $\mu_0 : \Sigma \to [0,\infty]$ is a σ-additive function, there exists a unique measure μ on $\Sigma = \sigma(\Sigma_0)$ s.t. $\mu(A) = \mu_0(A)$ for all $A \in \Sigma_0$. In other words $\mu|_{\Sigma_0} = \mu_0$.

Important Consequence: The Lebesgue measure is unique. We can construct $\mathcal{L}$ on $\mathcal{B}(\mathbb{R})$ by defining

$$\mathcal{L}_0\big((a_1,b_1)\cup(a_2,b_2)\cup\ldots\cup(a_n,b_n)\big) = (b_1-a_1)+\ldots+(b_n-a_n)$$

for all $a_1 < b_1 \leq a_2 < b_2 \leq \ldots \leq a_n < b_n$

and noting that sets of that form are an algebra.

Probability spaces

universe / σ-algebra called "events" / probability measure

Probability spaces $(\Omega, \mathcal{E}, \mathbb{P})$ are measure spaces where the measure $\mathbb{P}$ is a probability measure, $\mathbb{P}(\Omega)=1$.

Example: $\Omega = \{1,2,3,4,5,6\}$, $\mathcal{E} = \mathcal{P}(\Omega)$,

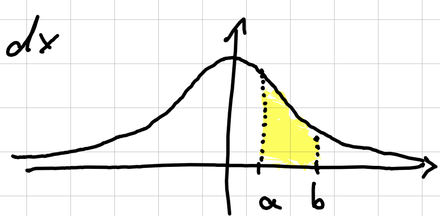
$\mathbb{P}(E) = \#E/6$ for all $E \in \mathcal{E}$

Formal model for rolling a die

Example: $\Omega = \mathbb{R}$, $\mathcal{E} = \mathcal{B}(\mathbb{R})$, $\mathbb{P}$ determined

by $\mathbb{P}((a,b)) = \int_a^b \frac{1}{2\pi} e^{-x^2/2}\,dx$

The normal distribution.



Almost sure events.

We say that an event $E \in \mathcal{E}$ happens almost surely if $\mathbb{P}(E)=1$. (Equivalently, $\mathbb{P}(E^c)=0$)

Example: Consider a uniformly random number X on the interval $[0,1]$. For every fixed $y \in [0,1]$ we have

$$\mathbb{P}(X=y) = \mathbb{P}(\{y\}) = 0$$

$$\mathbb{P}(X \neq y) = \mathbb{P}([0,1] \setminus \{y\}) = 1.$$

In words: $X \neq y$ almost surely.

Prop: If $A_1, A_2, \ldots$ are almost sure events, then so is $\bigcap_{i=1}^{\infty} A_i$:

$$\mathbb{P}(A_i) = 1 \quad \forall i \in \mathbb{N} \quad \Rightarrow \quad \mathbb{P}\Big(\bigcap_{i \in \mathbb{N}} A_i\Big) = 1.$$

Proof: By assumption, $\mathbb{P}(A_i) = 1$ for all $i \in \mathbb{N}$ and so $\mathbb{P}(A_i^c) = 0$.

[Formally, this is because $\Omega = A_i \cup A_i^c$ is disjoint and so $\mathbb{P}(\Omega) = \mathbb{P}(A_i) + \mathbb{P}(A_i^c) \Leftrightarrow 1 = \mathbb{P}(A_i) + 1$]

So $\mathbb{P}\Big(\bigcup_{i \in \mathbb{N}} A_i^c\Big) \leq \sum_{i \in \mathbb{N}} \mathbb{P}(A_i^c) = 0$.

But $\bigcup_{i \in \mathbb{N}} A_i^c = \Big(\bigcap_{i \in \mathbb{N}} A_i\Big)^c$ (de Morgan's law) and so

$$\mathbb{P}\Big(\bigcup_{i\in\mathbb{N}} A_i^c\Big) = \mathbb{P}\Big(\big(\bigcap_{i\in\mathbb{N}} A_i\big)^c\Big) = 0$$

Hence $\mathbb{P}\big(\bigcap_{i\in\mathbb{N}} A_i\big) = 1.$ $\square$

Important: This only works for countable intersections!

Example: Let X be uniformly random on $[0,1]$.
For any $x \in [0,1]$ we have $\mathbb{P}(X \neq x) = 1$.
Since there are countably many rational numbers

$$\mathbb{P}\Big(\bigcap_{x\in\mathbb{Q}\cap[0,1]} \{X \neq x\}\Big) = \mathbb{P}(X \text{ is irrational}) = 1.$$

But $\mathbb{P}\Big(\bigcap_{x\in[0,1]} \{X\neq x\}\Big) = \mathbb{P}(X \text{ takes a value outside } [0,1])$

$$= 0 \neq 1$$

since it is an uncountable intersection.

lim inf and lim sup ($\underline{\lim}$ & $\overline{\lim}$)
Recall that

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty} \sup_{m\geq n} x_m$$

$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty} \inf_{m\geq n} x_m$$

The limsup and liminf always exist;
The limit lim exists if liminf = limsup.

$$\limsup_{n\to\infty} x_n \geq x \iff \left(\begin{array}{l} \text{there exists a subsequence} \\ \text{of } (x_n) \text{ with limit greater} \\ \text{than } x \end{array} \right)$$

There is a similar concept for sets:
Let $E_1, E_2, ..$ be events (sets)

$$\liminf_{n\to\infty} E_n = \bigcup_{n\geq 1} \bigcap_{m \geq n} E_m$$

$$\limsup_{n\to\infty} E_n = \bigcap_{n\geq 1} \bigcup_{m\geq n} E_m$$

an increasing seq of sets

a decreasing seq. of sets

$\liminf\limits_{n\to\infty} E_n$ contains all elements that eventually occur in all E_m, $m \geq n$.
($\iff$ occur in all but finitely many E_m)

$\limsup\limits_{n\to\infty} E_n$ contains all elements that occur in infinitely many E_m.

We have $\liminf\limits_n E_n \subseteq \limsup\limits_n E_n$.

Fatou's Lemma:

$$\mathbb{P}(\liminf_n E_n) \leq \liminf \mathbb{P}(E_n).$$

Proof: Write $F_n = \bigcap_{m \geq n} E_m$.

Then $\liminf_n E_n = \bigcup_{n \in \mathbb{N}} F_n$.

Since $F_n \subseteq E_m$ for all $m \geq n$, we have

$\mathbb{P}(F_n) \leq \mathbb{P}(E_m)$ for all $m \geq n$ and so

$$\mathbb{P}(F_n) \leq \inf_{m \geq n} \mathbb{P}(E_m). \quad (*)$$

F_n is an increasing seq. of sets and so

$\lim_n \mathbb{P}(F_n)$ exists and equals

$\mathbb{P}(\bigcup_{n \in \mathbb{N}} F_n) = \mathbb{P}(\liminf_{n \in \mathbb{N}} E_n)$. Since,

$\lim_n \mathbb{P}(F_n) \leq \lim_n \inf \mathbb{P}(E_m)$ by $(*)$

we get $\mathbb{P}(\liminf_n E_n) \leq \liminf_n \mathbb{P}(E_n)$

as required $\square$

Similarly, $\mathbb{P}(\limsup_n E_n) \geq \limsup \mathbb{P}(E_n)$.

(Reverse Fatou Lemma)

Borel-Cantelli Lemma

Let $E_1, E_2, \dots,$ be a seq of events with the property that $\sum_{i \in \mathbb{N}} \mathbb{P}(E_i) < \infty$. Then,

$$\mathbb{P}(\limsup_{n \to \infty} E_n) = \mathbb{P}(\text{infinitely many } E_n \text{ occur}) = 0. \quad (\dagger)$$

<u>Proof</u>: Recall that $\limsup_{n \to \infty} E_n = \bigcap_{n \in \mathbb{N}} \underbrace{\bigcup_{m \geq n} E_m}_{=: G_n}$,

where G_n is a decreasing seq. of sets.

Hence $\limsup_{n \to \infty} E_n \subseteq G_m$ for all $m \in \mathbb{N}$

and $\mathbb{P}(\limsup_n E_n) \leq \mathbb{P}(G_m) \leq \sum_{k \geq m} \mathbb{P}(E_k)$

for all $m \in \mathbb{N}$. But since $\sum_{k \geq 1} \mathbb{P}(E_k) < \infty$,

we must have $s_m = \sum_{k=m}^{\infty} \mathbb{P}(E_k) \to 0$ as $m \to \infty$.

Hence $\mathbb{P}(\limsup_n E_n) \leq s_m$ for all m and $(\dagger)$ follows. $\square$