Probability and Random Processes

Lecture 3

- General measure theory

Measure

- For the generalization of "length" (to Lebesgue measure), we required
  - length$(A) \geq 0$ for all A
  - length$(A) = b - a$ if A is an interval with endpoints $a \leq b$
  - length$(B) =$ length$(B_1) +$ length(B_2) if $B = B_1 \cup B_2$ and $B_1 \cap B_2 = \emptyset$
- How to generalize to "measure" in a much more general setting?

- An arbitrary set Ω, and a measure μ on sets from Ω
- As before, require
  - $\mu(A) \geq 0$ for all A
  - $\mu(\emptyset) = 0$
  - $\mu(B) = \mu(B_1) + \mu(B_2)$ if $B = B_1 \cup B_2$ and $B_1 \cap B_2 = \emptyset$ (finite additivity)
- To prove limit theorems and similar, we also need countable additivity,
  - for a sequence of sets $\{B_n\}$ with $B_i \cap B_j = \emptyset$ if $i \neq j$

$$\mu\left(\bigcup_n B_n\right) = \sum_n \mu(B_n)$$

- Ω is the universal set ($\mathbb{R}$ in the case of Lebesgue measure), as before we cannot expect that μ can act on all subsets of Ω
  - but sets of the kind $\bigcup_n B_n$ need to be in the domain of μ

$\Rightarrow$ A σ-algebra is a class $\mathcal{A}$ of sets in Ω such that
  - $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$
  - $A_n \in \mathcal{A}, n = 1, 2, 3, \ldots, \Rightarrow \cup_n A_n \in \mathcal{A}$
- On $\mathbb{R}$, the set $\mathcal{L}$ of Lebesgue measurable sets is a σ-algebra

- A set Ω, a σ-algebra $\mathcal{A}$ of subsets. A measure μ is a function $\mu : \mathcal{A} \to \mathbb{R}^*$ such that
  - $\mu(A) \geq 0$ for all A
  - $\mu(\emptyset) = 0$
  - $\{B_n\}$ with $B_i \cap B_j = \emptyset$ if $i \neq j \Rightarrow$

$$\mu\left(\bigcup_n B_n\right) = \sum_n \mu(B_n)$$

- $(\Omega, \mathcal{A})$ is a measurable space, and $(\Omega, \mathcal{A}, \mu)$ is a measure space; the sets in $\mathcal{A}$ are called $\mathcal{A}$-measurable

Borel Measurable Sets and Functions

- σ-algebra generated by a class $\mathcal{A}$ of sets = smallest σ-algebra that contains $\mathcal{A}$
- Collection of Borel sets $\mathcal{B} = \sigma$-algebra generated by the class of open sets:
  - In general (beyond $\mathbb{R}$) the "open sets" are picked from a topology
  - For the real line, $\mathcal{B}$ is the smallest σ-algebra that contains the open intervals
- The sets in $\mathcal{B}$ are called Borel measurable
- On the real line, $\mathcal{B} \subset \mathcal{L}$, that is, Borel measurable $\Rightarrow$ Lebesgue measurable (but not vice versa)

- A function $f : \mathcal{U} \to \mathcal{V}$ is Borel measurable if the inverse image of any open set in $\mathcal{V}$ is Borel measurable in $\mathcal{U}$,
  - again, "open set" and "Borel measurable" can be general (defined by topologies on $\mathcal{U}$ and $\mathcal{V}$)
- On the real line, let $\mathcal{F} = \{\,\text{Borel measurable functions}\,\}$
  - $f \in \mathcal{F}$ if the inverse image of any open set (or interval) $\subset \mathcal{B}$
  - "Borel measurable set" more general than "open set" $\Rightarrow$ $\mathcal{F}$ contains the continuous functions
  - $f \in \mathcal{F} \Rightarrow f$ Lebesgue measurable (but not vice versa)
  - $\mathcal{F}$ is closed under pointwise limits
  - $\mathcal{F} =$ smallest class of functions that contains the continuous functions and their pointwise limits,
    - c.f. the class $\{\,\text{continuous functions}\,\}$ which is closed under uniform but not pointwise convergence

Some Measure Spaces

- $(\mathbb{R}, \mathcal{L}, \lambda)$ and $(\mathbb{R}, \mathcal{B}, \lambda_{|\mathcal{B}})$
- $\Omega \neq \emptyset$ and $\mathcal{A} = \mathcal{P}(\Omega) =$ set of all subsets. For $A \in \mathcal{A}$, let

$$\mu(A) = \begin{cases} N(A) & A \text{ finite} \\ \infty & A \text{ infinite} \end{cases}$$

  where $N(A) =$ number of elements in A; then μ is called counting measure
- $\Omega \neq \emptyset$ and $\mathcal{A} = \mathcal{P}(\Omega)$, for $A \in \mathcal{A}$ the measure

$$\delta_x(A) = \begin{cases} 1 & x \in A \\ 0 & \text{o.w.} \end{cases}$$

  is called Dirac measure concentrated at x

- A measure space $(\Omega, \mathcal{A}, \mu)$ with the additional condition $\mu(\Omega) = 1$ is called a probability space and μ is a probability measure
- A measure space $(\Omega, \mathcal{A}, \mu)$. The measure μ is finite if $\mu(\Omega) < \infty$.
- Given $(\Omega, \mathcal{A})$ a measure μ is σ-finite if there is a sequence $\{A_i\}$, $A_i \in \mathcal{A}$, such that $\cup_i A_i = \Omega$ and $\mu(A_i) < \infty$

Complete Measure

- On the real line: $A \in \mathcal{L}$, $\lambda(A) = 0$ and $B \subset A \Rightarrow B \in \mathcal{L}$, however $A \in \mathcal{B}$, $\lambda_{|\mathcal{B}}(A) = 0$ and $B \subset A$ does in general not imply $B \in \mathcal{B}$
  - $\Rightarrow$ $(\mathbb{R}, \mathcal{B}, \lambda_{|\mathcal{B}})$ is not *complete*
- A measure space is complete if subsets of sets of measure zero are measurable
- For $(\Omega, \mathcal{A}, \mu)$, let $\bar{\mathcal{A}} =$ collection of all sets of the form $B \cup A$ where $B \in \mathcal{A}$ and $A \subset C$ for some $C \in \mathcal{A}$ with $\mu(C) = 0$. For such a set $\bar{A} = B \cup A \in \bar{\mathcal{A}}$ define $\bar{\mu}(\bar{A}) = \mu(B)$, then
  - $(\Omega, \bar{\mathcal{A}}, \bar{\mu})$ is the completion of $(\Omega, \mathcal{A}, \mu)$
  - $\mathcal{A} \subset \bar{\mathcal{A}}$, $\bar{\mu}_{|\mathcal{A}} = \mu$ and $\bar{\mu}(\bar{A}) = \inf\{\mu(A) : \bar{A} \subset A \in \mathcal{A}\}$
  - if $(\Omega, \mathcal{F}, \nu)$ is complete, $\mathcal{A} \subset \mathcal{F}$ and $\nu_{|\mathcal{A}} = \mu$ then $\bar{\mathcal{A}} \subset \mathcal{F}$ and $\nu_{|\bar{\mathcal{A}}} = \bar{\mu}$, i.e. completion gives the smallest complete space
  - the completion is unique if $(\Omega, \mathcal{A}, \mu)$ is σ-finite
- $(\mathbb{R}, \mathcal{L}, \lambda)$ is the completion of $(\mathbb{R}, \mathcal{B}, \lambda_{|\mathcal{B}})$

Measurable Functions

- A measure space $(\Omega, \mathcal{A}, \mu)$
- A function $f : \Omega \to \mathbb{R}$ is $\mathcal{A}$-measurable if $f^{-1}(O) \in \mathcal{A}$ for all open sets $O \subset \mathbb{R}$
  - If $(\Omega, \mathcal{A}, \mu)$ is a probability space ($\mu(\Omega) = 1$), then measurable functions are called random variables
- If $(\Omega, \mathcal{A}, \mu)$ is complete, f is measurable and $g = f$ μ-a.e., then g is measurable
- Given two measurable spaces $(\Omega, \mathcal{A})$ and $(\Lambda, \mathcal{S})$, a mapping $T : \Omega \to \Lambda$ is a measurable transformation if $T^{-1}(S) \in \mathcal{A}$ for each $S \in \mathcal{S}$ (note: the sets in $\mathcal{S}$ are not necessarily "open")
- For T from $(\Omega, \mathcal{A})$ to $(\Lambda, \mathcal{S})$,
  - the class $T^{-1}(\mathcal{S})$ is a σ-algebra $\subset \mathcal{A}$
  - the class $\{L \subset \Lambda : T^{-1}(L) \in \mathcal{A}\}$ is a σ-algebra $\subset \mathcal{S}$
  - if $\mathcal{S} = \sigma(\mathcal{C})$ for some $\mathcal{C}$, then T is measurable (from $\mathcal{A}$ to $\mathcal{S}$) iff $T^{-1}(\mathcal{C}) \subset \mathcal{A}$

Almost Everywhere

- A measure space $(\Omega, \mathcal{A}, \mu)$; if a statement is true for elements from all sets in $\mathcal{A}$ except for those in $N \in \mathcal{A}$ with $\mu(N) = 0$, then the statement is true μ-almost everywhere (μ-a.e.)
  - . . . almost always, for almost all, almost surely, with probability one, almost certainly. . .
- A sequence $\{f_n\}$ of $\mathcal{A}$-measurable functions converges μ-a.e. to the function f if $\lim_{n\to\infty} f_n(x) = f$ pointwise for all x except for x in a set $E \in \mathcal{A}$ with $\mu(E) = 0$

Convergence in Measure

- A measure space $(\Omega, \mathcal{A}, \mu)$ and a sequence $\{f_n\}$ of $\mathcal{A}$-measurable functions. The sequence converges in measure to the function f if

$$\lim_{n\to\infty} \mu(\{x : |f(x) - f_n(x)| \geq \varepsilon\}) = 0$$

  for each $\varepsilon > 0$
- If μ is finite then convergence μ-a.e. implies convergence in measure, but not vice versa
- If $(\Omega, \mathcal{A}, \mu)$ is a probability space and $\{f_n\}$ are random variables, then convergence in measure is called convergence in probability, and convergence μ-a.e. is with probability one