EECS502 Stochastic Processes Lecture 1 - 01/19/2021

Review of Probability and Random Processes

Lecturer: Vijay G. Subramanian *Scribes: V. Subramanian*

1 Introduction

Textbooks for the course are [1, 2] .

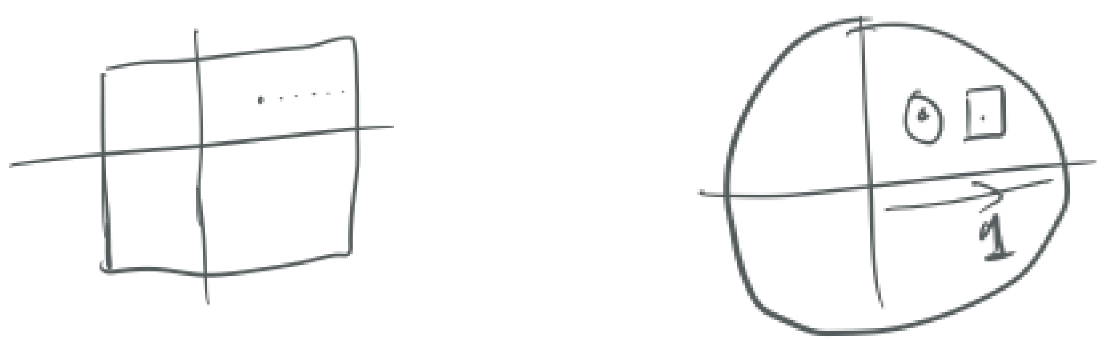
1.1 Notation

We have the following:

1. Natural numbers - $\mathbb{N} = \{1, 2, \dots\}$.
2. Integers - $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$.
3. Whole numbers - $\mathbb{Z}_+ = \mathbb{N} \cup \{0\} = \{0, 1, 2, \dots\}$. These are also the non-negative integers.
4. Real numbers - $\mathbb{R} = (-\infty, \infty)$.
5. Non-negative real numbers - $\mathbb{R}_+ = [0, +\infty)$.

We will need to talk about sets on the real line so we will start with intervals on the real-line:

- Closed interval - $[a, b] : a \leq x \leq b$.
- Open interval - $(a, b) : a < x < b$.
- Half open and half closed intervals - $(a, b] : a < x \leq b$ and $[a, b) : a \leq x < b$.



(a) Example closed set in $\mathbb{R}^2$.

(b) Example open set in $\mathbb{R}^2$.

Figure 1: Example of closed and open sets.

Open and closed intervals generalize to open and closed sets; see Figure 1 for examples. We have the following:

1. Closed set C - here given a sequence $\{x_i\}_{i=1}^{\infty}$ in C, i.e., $x_i \in C \forall i = 1, 2, \ldots$, such that $\lim_{n\to\infty} x_n$ exists, then $\lim_{n\to\infty} x_n \in C$. In other words, taking limits one cannot escape C. Figure 1a gives an example of a square in $\mathbb{R}^2$ where the edges are included in C.

2. Open set O - here $\forall$ (for all) $x \in O$ we can find a small enough ball (open interval $(x-\epsilon, x+\epsilon)$) - called neighborhood of x - that fully contained in O. Figure 1b illustrates this in $\mathbb{R}^2$ with an open circle - $\{\mathbf{x} : \|\mathbf{x}\|^2 < 1\} = \{(x_1, x_2) : x_1^2 + x_2^2 < 1\}$.

Next we discuss a probability space $(\Omega, \mathcal{F})$ where we have

- Ω - sample space or space of outcomes.
- $\mathcal{F}$ - σ-algebra (or σ-field) which is a collection of subsets of Ω.
- ω - sample point or realization where $\omega \in \Omega$. For a random process this will also stand for sample-path.

A probability distribution puts values on elements of $\mathcal{F}$.

What sort of collection of subsets is a σ-algebra? We will start with the definition of an algebra. A collection of subsets $\tilde{\mathcal{F}}$ of Ω is called an algebra if the following are true:

1. (Non-empty) Either $\Omega \in \tilde{\mathcal{F}}$ or $\emptyset \in \tilde{\mathcal{F}}$.

2. (Closure under complements) If $F \in \tilde{\mathcal{F}}$, then its complement $F^C \in \tilde{\mathcal{F}}$. Note that this implies that both Ω and $\emptyset$ are always elements of an algebra.

3. (Closure under finite unions) If $A, B \in \tilde{\mathcal{F}}$, then we have $A \cup B \in \tilde{\mathcal{F}}$ too. Using induction, it then follows that for every $n \in \mathbb{N}$ and collection of sets $A_1, \ldots, A_n \in \tilde{\mathcal{F}}$, we have $\cup_{i=1}^{n} A_i \in \tilde{\mathcal{F}}$.

Examples of algebras:

1. Trivial σ-algebra, i.e., $\tilde{\mathcal{F}} = \{\emptyset, \Omega\}$.

2. Largest σ-algebra, i.e., $\tilde{\mathcal{F}} = 2^{\Omega} = \mathcal{P}(\Omega)$, i.e., collection of all subsets of Ω. This is typically used for discrete settings:

   (a) Single coin toss - $\Omega = \{H, T\}$, $\tilde{\mathcal{F}} = 2^{\Omega} = \{\emptyset, \{H\}, \{T\}, \{H, T\}\}$.

   (b) Single six-sided dice - $\Omega = \{1, 2, \ldots, 6\}$, $\tilde{\mathcal{F}} = 2^{\Omega} = \{\emptyset, \{1\}, \ldots, \{1, 2\}, \ldots, \{1, 2, 3\}, \ldots, \{1, 2, 3, 4\}, \ldots,$ $\{1, 2, 3, 4, 5\}, \ldots, \{1, 2, 3, 4, 5, 6\}\}$. For $k = 0, 1, \ldots, 6$, there are $\binom{6}{k}$ subsets of size k.

3. $\Omega = [0, 1]$. Then $\tilde{\mathcal{F}} = \{$all **finite** unions of intervals$\}$.

Next we specify a σ-algebra. A collection of subsets $\mathcal{F}$ of Ω is a σ-algebra if

1. It is an algebra,

2. **(***)** (Closure under **countable** unions) - if we have a sequence of elements $\{A_i\}_{i=1}^{\infty}$ s.t. (such that) $A_i \in \mathcal{F}$ for all $i = 1, 2, \ldots$, then we have $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Examples:

1. **(***)** $\Omega = \mathbb{R}$, $\mathcal{F} = \mathcal{B}(\mathbb{R})$ which is the Borel σ-algebra - smallest σ-algebra that contains all intervals (alternatively all open, closed or open and closed sets).

2. **(*)** $\Omega = \mathbb{R}$, $\mathcal{F} = \{F \subseteq \Omega :$ either F or F^C is countable$\}$.

In the definition of the Borel σ-algebra we included the term of a smallest σ-algebra. This holds by containment and arises as follows. Let $\tilde{\mathcal{G}}$ be a collection of subsets of Ω, then denote by $\sigma(\tilde{\mathcal{G}})$ the smallest σ-algebra that contains $\tilde{\mathcal{G}}$. How to obtain it? Find all σ-algebra $\mathcal{F}$ that contain $\tilde{\mathcal{G}}$ and take their intersection, i.e., find all elements common to them, then this intersection is still a σ-algebra and is $\sigma(\tilde{G})$. Since $\tilde{\mathcal{G}} \subseteq 2^{\Omega}$, there is always a σ-algebra that contains $\tilde{\mathcal{G}}$. Further, it can be shown that any intersection of σ-algebras is always a σ-algebra. Then we have the following:
$\tilde{\mathcal{G}} = \{[a,b] : a,b \in \mathbb{R}\}$, then $\sigma(\tilde{\mathcal{G}}) = \mathcal{B}(\mathbb{R})$.
$\tilde{\mathcal{G}} = \{O : O \text{ open in } \mathbb{R}\}$, then too $\sigma(\tilde{\mathcal{G}}) = \mathcal{B}(\mathbb{R})$.

Next we will discuss random variables but we will start by considering functions. Let $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ be two probability spaces. Then we will consider functions $f : \Omega_1 \mapsto \Omega_2$. E.g., we can take $\Omega_1 = \Omega_2 = \mathbb{R}$ and $\mathcal{F}_1 = \mathcal{F}_2 = \mathcal{B}(\mathbb{R})$ with example functions being $f(x) = x$, $f(x) = x^2$ and $f(x) = \text{sign}(x)$. An important point for random variables will be understanding how the inverse map of f is with respect to $\mathcal{F}_1$ and $\mathcal{F}_2$. Then we have the following:

- For $B \subseteq \Omega_2$, define the inverse image of B, $f^{-1}(B) \subseteq \Omega_1$, by $f^{-1}(B) := \{\omega_1 \in \Omega_1 : f(\omega_1) \in B\}$.
- Then set $\tilde{\mathcal{G}} = \{f^{-1}(B) : B \in \mathcal{F}_2$ and obtain $\sigma(\tilde{\mathcal{G}})$.
- Then the relevant question will be: *How does $\sigma(\tilde{\mathcal{G}})$ compare to $\mathcal{F}_1$?*
  This will be important for random variables.

Given a probability space $(\Omega, \mathcal{F})$ we next define a probability distribution or measure $\mathbb{P}$ as follows. The probability distribution **assigns values to elements of $\mathcal{F}$ that are between** 0 **and** 1, i.e., $\mathbb{P} : \mathcal{F} \mapsto [0,1]$, such that:

1. (Normalized) $\mathbb{P}(\Omega) = 1$.
2. (Countably additive) If $\{A_i\}_{i=1}^{\infty}$ (also written as $\{A_i\}_{i\in\mathbb{N}}$) is a *countable* and *disjoint* collection of subsets of Ω *from* $\mathcal{F}$ (i.e., $A_i \in \mathcal{F}$ for all $i \in \mathbb{N}$), then we have
$$\mathbb{P}\left(\cup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i),$$
where the infinite sum is understood as a limit, i.e., $\sum_{i=1}^{\infty} \mathbb{P}(A_i) = \lim_{n\to\infty} \sum_{i=1}^{n} \mathbb{P}(A_i)$. Note that disjoint (same as mutually exclusive) means that for all $i, j \in \mathbb{N}$ with $i \neq j$, we have $A_i \cap A_j = \emptyset$.

Examples:

1. Bernoulli distribution: $\Omega = \{0,1\}$, $\mathcal{F} = 2^{\Omega} = \{\emptyset, \{0\}, \{1\}, \{0,1\}\}$. Then given $p \in [0,1]$ we can define $\mathbb{P}$ by assigning the following values to the elements of $\mathcal{F}$: $\{0, 1-p, p, 1\}$.
2. Uniform distribution on $[0,1]$: $\Omega = [0,1]$, $\mathcal{F} = \mathcal{B}([0,1])$ (Borel σ-algebra of $[0,1]$ which is defined similarly as for $\mathbb{R}$) and the probability measure is Lebesgue measure:
Given $0 \leq a \leq b \leq 1$, $\mathbb{P}([a,b]) = b - a$, i.e., the probability of an interval is its length, and for $B \in \mathcal{B}([0,1])$, $\mathbb{P}(B) := \int_0^1 1_B(x)dx$ where the indicator function of a set B is given by
$$1_B(x) = \begin{cases} 1 & \text{if } x \in B; \\ 0 & \text{otherwise,} \end{cases}$$
where otherwise corresponds to $x \notin B$.
3. Normal distribution with mean μ and variance $\sigma^2 > 0$: $\Omega = \mathbb{R}$ and $\mathcal{F} = \mathcal{B}(\mathbb{R})$ (Borel σ-algebra). Then we have for $B \in \mathcal{F}$,
$$\mathbb{P}(B) = \int_{-\infty}^{\infty} 1_B(x) \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx.$$

References

[1] P. G. Hoel, S. C. Port and C. J. Stone, "Introduction to stochastic processes," 1986, Waveland Press.

[2] P. Brémaud, "Markov chains: Gibbs fields, Monte Carlo simulation, and queues," 2013, Vol. 31, Springer Science & Business Media.