MATH6710 **Summer**

Probability Theory I

Notes

Authors: Shaleen Baral

Contents

1. Introduction

We will be covering the following material

a. Measure Theory,
b. Random Variables,
c. Law of Large Numbers,
d. Weak Convergence, Central Limit Theorems.

The main textbook is R. Durrett, Probability, Theory & Examples (5th edition). See the following books for an alternate perspective

- D. WIlliams, Probability with Martingales
- K.L. Chung, A Course in Probability

In *naive* probability theory, we consider a countable sample space $\Omega \subseteq \mathbb{N}$ of possible outcomes and a function $\mathbb{P} : \mathcal{P}(\Omega) \to [0, 1]$ such that

a. $\mathbb{P}(\Omega) = 1$
b. If we have disjoint *events* $A, B \in \mathcal{P}(\Omega)$ then $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ *(finite additivity)*

Proposition 1.1 ($\mathbb{P}$ is monotone): Let $A \subseteq B \subseteq \Omega$. Then

$$\mathbb{P}(A) \leq \mathbb{P}(B)$$

Proof: Consider the disjoint events A and $B - A$. Then,

$$\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B - A) \geq \mathbb{P}(A).$$

□

Proposition 1.2 ($\mathbb{P}(A^C)$):

$$\mathbb{P}(A^C) = 1 - \mathbb{P}(A)$$

Proof:

$$\begin{gathered} 1 = \mathbb{P}(\Omega) = \mathbb{P}(\Omega - A) + \mathbb{P}(A) \\ \Longrightarrow \mathbb{P}(A^C) = 1 - \mathbb{P}(A). \end{gathered}$$

□

Proposition 1.3 (inclusion-exclusion): For any events $A, B \subseteq \Omega$, we have

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

Proof: We have $A \cup B = (A - B) \cup (B - A) \cup (A \cap B)$. □

Corollary 1.3.1 (finite subadditivity): For any events $A, B \subseteq \Omega$, we have

$$\mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B)$$

Why should we be dissatisfied with finite additivity?

Consider an infinite sequence of independent, random variables

$$X_j = \begin{cases} +1 & \text{w.p. } \frac{1}{2} \\ -1 & \text{w.p. } \frac{1}{2} \end{cases}$$

We consider the random walk defined by $S_n = \sum_{j=1}^{N} X_j$. Then, define the event

$$\begin{aligned} A &= \{S \text{ visits } 0 \text{ infinitely often}\} \\ &= \bigcap_{N \geq 1} \underbrace{\{\exists k \geq N, S_k = 0\}}_{A_n} \\ &= \bigcap_{N \geq 1} A_n \end{aligned}$$

We want to be able to use 'finite observations' or 'approximations', A_n to compute A. Ideally,

$$\mathbb{P}(A) = \lim_{N \to \infty} \mathbb{P}(A_N)$$

Note that for $N' \geq N$, we have $A_{N'} \subseteq A_N$. Note also that

$$\mathbb{P}\left(\bigcap_{N \geq 1} A_N\right) = 1 - \mathbb{P}\left(\bigcup_{N \geq 1} A_N^C\right)$$

The sequence of complements is increasing $A_1^C \subseteq A_2^C \subseteq \dots$.

Before moving on note that the prior statement is equivalent to sayign

$$\mathbb{P}(B) = \lim_{N \to \infty} \mathbb{P}(B_N)$$

for $B_1 \subseteq B_2 \subseteq \dots$ and $B = \bigcup_{N \geq 1} B_N$.

We convert $\{B_N\}$ to a disjoint family by considering the family given by $B_i \setminus B_{i-1}$ (for convenience, take $B_0 = \emptyset$). Then, using finite additivity, we would like to say

$$\mathbb{P}(B) = \lim_{N \to \infty} \sum_{j=1}^{N} \mathbb{P}(B_j \setminus B_{j-1})$$

Equivalently, in our theory of probability we would want, for a *countable*, disjoint family $\{C_i\}$ the following holds

$$\mathbb{P}\left(\bigcup_{j=1}^{\infty} C_j\right) = \sum_{j=1}^{\infty} \mathbb{P}(C_j).$$

Why do we want countable and not arbitrary additivity?

Consider a probability measure on $[0, 1]$. Then, if we have arbitrary additivity,

$$\mathbb{P}\left(\bigcup_{j \in I} A_i\right) = \sum_{j \in I} \mathbb{P}(A_j),$$

it leads to contradictions of the following form

$$\mathbb{P}\left(\bigcup_{j \in [0,1]} j\right) = \sum_{j \in [0,1]} \mathbb{P}(\{j\})$$

The left hand side must be 1 whereas the equality holds only under very particular probability measures. In fact, this already rules out, for example the uniform measure on $[0, 1]$.

Now, we move on to actually developing a measure theory that incorporates these ideas.

Definition 1.1:

$$\mu^\star(A) = \inf\left\{ \sum_{j=1}^{N} |B_j| \text{ for } \{B_j\} \text{ such that } A \subseteq \bigcup B_j \right\}$$
$$\mu_\star(A) = \sup\left\{ \sum_{j=1}^{N} |B_j| \text{ for } \{B_j\} \text{ such that } A \subseteq \bigcup B_j \right\}$$

Definition 1.2 (Jordan-measurable): We call A *Jordan measurable* if $\mu^\star(A) = \mu_\star(A)$.

Proposition 1.4: The Jordan measure is finitely additive.

Proof: Consider disjoint sets A, B. Then note that any open cover of A and B give rise to an open cover of $A \cup B$. From this, we immediately have $\mu(A \cup B) \leq \mu(A) + \mu(B)$.

For the converse, consider a sufficiently fine open cover $\{U_i\}$ of $A \cup B$. Then define open covers for A and B as follows.

□