AMATH 561A: Probability and Random Processes

Marlin Figgins

Oct. 1, 2020

Introduction

This is a collection of my notes taken during AMATH 561A during fall quarter 2020 at the University of Washington.

Table of Contents

1. Probability Spaces and Random Variables

Definition 1.1 (Probability Space: Version A). A *probability space* is a triple $(\Omega, \mathcal{F}, \mathbb{P})$ where Ω is the *sample space* or the space of possible outcomes, $\mathcal{F}$ is the set of events which are subsets of Ω and $\mathbb{P}\colon \mathcal{F} \to [0,1]$ is a probability measure that assigns probabilities to events.

Example 1.2 (Rolling a fair standard die). In this case, we have six possible outcomes as our die is six-sided. Therefore, our sample space is

$$\Omega = \{1, 2, 3, 4, 5, 6\}. \tag{1.3}$$

The set of events in this case are all possible subsets of Ω i.e. the power set $\mathcal{P}(\Omega)$. Finally, since we specificed that the die is fair, each outcome is equally likely, so our probability measure $\mathbb{P}$ is uniform over Ω. Therefore, our probability of a given event is just the fraction of our possible out comes which are in the event.

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{|A|}{6} \text{ for all } A \in \mathcal{F}. \tag{1.4}$$

For example, we would say that the probability of rolling an even number would be given by

$$\mathbb{P}(\{2, 4, 6\}) = \frac{3}{6} = 0.5. \tag{1.5}$$

In this example, it is simple enough to use all possible subets are our set of events, but as we attempt to deal with more complicated probability spaces, we'll need to develop a more rigorous notion of what constitutes an event, so that our probability measures will have certain properties of interest. This idea is formalized by the σ-algebra.

Definition 1.6 (σ-algebra). A non-empty collection of subsets of Ω is a *σ-algebra* of Ω if it satistifies the following:

(i) For each event $A \in \mathcal{F}$, then its complement $A^c \in \mathcal{F}$

(ii) If we have a countable sequence of sets $A_i \in \mathcal{F}$, then their union $\bigcup_i A_i \in \mathcal{F}$.

These two conditions form the statement that a σ-algebra is closed under complements and countable unions.

Proposition 1.7. *The definition of the σ-algebra implies that σ-algebras are also closed under countable intersection because*

$$\bigcap_i A_i = \left(\bigcup_i A_i^c \right)^c. \tag{1.8}$$

Exercise 1.9. Show that any σ-algebra of Ω contains both the entire space Ω and the empty set.

When we have a probability space Ω and a σ-algebra $\mathcal{F}$ on Ω, we can define a measure on the space $(\Omega, \mathcal{F})$. *Describe what a measure is intuitively.*

Definition 1.10 (Measure)**.** A *measure* is a function $\mu : \mathcal{F} \to [0, \infty)$ which satisfies:

(i) $\mu(A) \geq \mu(\emptyset) = 0$ for all $A \in \mathcal{F}$

(ii) if $A_i \in F$ is a countable sequence of disjoint sets, then

$$\mu\left(\bigcup_i A_i\right) = \sum_i \mu(A_i). \tag{1.11}$$

In the case that $\mu(\Omega)$, we will call μ a *probability measure* and denote it as $\mathbb{P}$.

Our definition of measure allows us to ensure that our notion of probabilty satisifies some of the intuitive properties one might expect of probabilities.

Theorem 1.12 (Properties of measure)**.**

(i) Monotonicity. *If $A \subset B$, then $\mu(A) \leq \mu(B)$.*

(ii) Subadditivity. *If $A \subset \bigcup_{j\in\mathbb{N}} A_j$, then*

$$\mu(A) \leq \sum_{j\in\mathbb{N}} \mu(A_m). \tag{1.13}$$

Proof of monotonicity. Finish proof. □

Proof of subadditivity. Finish proof. □

Theorem 1.14 (Continuity of measure)**.**

(i) Continuity from above. *If we have a sequence of increasing subsets $A_1 \subset A_2 \subset \cdots$ such that $\bigcup_i A_i = A$, then*

$$\mu(A_i) \uparrow \mu(A) \text{ as } i \to \infty. \tag{1.15}$$

(ii) Continuity from below. *If we have a sequence of decreasing subsets $A_1 \supset A_2 \supset \cdots$ such that $\bigcup_i A_i = A$, then*

$$\mu(A_i) \downarrow \mu(A) \text{ as } i \to \infty. \tag{1.16}$$

Exercise 1.17. Prove that in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$

$$\mathbb{P}(A) = 1 - \mathbb{P}(A^c) \text{ for any event } A \in \mathcal{F}. \tag{1.18}$$

Discrete Probability Spaces

Constructing a discrete probability space. We'll now focus on the case of discrete probability spaces. Let Ω be a countable set. That is, either finite or countably infinite. Also, let $\mathcal{F}$ be the set of all subsets of Ω which we will call the power set $\mathcal{P}(\Omega)$. We can know endow $(\Omega, \mathcal{F})$ with a probability measure.

Theorem 1.19. *Suppose we have a countable set Ω with a σ-algebra $\mathcal{F} = \mathcal{P}(\Omega)$. Then any function $p\colon \Omega \to [0,1]$ such that*

$$\sum_{\omega \in \Omega} p(\omega) = 1 \tag{1.20}$$

induces a probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ as follows. For any event $A \in \mathcal{F}$, we define the probability of A as

$$\mathbb{P}(A) = \sum_{\omega \in A} p(\omega). \tag{1.21}$$

We call the function p the probability mass function *of $\mathbb{P}$.*

Example 1.22 (Repeated fair coins)**.** Consider the following experiment: We have a fair coin and we flip it until it lands on heads. We can write the possible outcomes as sequences of heads and tails, so that

$$\Omega = \{H, TH, TTH, TTH, \dots\}. \tag{1.23}$$

As before, we can let $\mathcal{F} = \mathcal{P}(\Omega)$. Let's motivate our choice of probability mass function p. Assuming that the probability of heads and tails at every flip is $\frac{1}{2}$, then we have a probability of $\left(\frac{1}{2}\right)^n$ for having n consequective flips. This tells us that

$$p(\underbrace{T \cdots T}_{n \text{ tails}} H) = \left(\frac{1}{2}\right)^n. \tag{1.24}$$

We can check that this indeed sums to one:

$$\sum_{\omega \in \Omega} p(\omega) = \sum_{i=1}^{\infty} \left(\frac{1}{2}\right)^n. \tag{1.25}$$

As shown above, this probability mass function induces a probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$. We can now use this measure the compute the probability of events like:

$$A = \{\text{Heads appears before third toss.}\} = \{H, TH\}, \tag{1.26}$$

$$B = \{\text{There are an even number of tails.}\} = \{H, TTH, TTTTH, \dots\}. \tag{1.27}$$

We can then compute these probabilities as:

$$\mathbb{P}(A) = \frac{1}{2} + \frac{1}{4}, \tag{1.28}$$

$$\mathbb{P}(B) = \frac{1}{2} + \frac{1}{2^3} + \frac{1}{2^5} + \cdots = \frac{2}{3}. \tag{1.29}$$

2. Independence, Martingales, and Conditioning

3. Characteristic Functions

4. Markov Chains

5. Generating Functions and Branching Processes

6. Convergence of Random Variables