

# **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

<b>1</b>	<b>Probability Spaces and Random Variables</b>	<b>1</b>
1.1	Discrete Probability Spaces . . . . .	4
1.2	Uncountable probability spaces . . . . .	5
<b>2</b>	<b>Independence, Martingales, and Conditioning</b>	<b>7</b>
<b>3</b>	<b>Characteristic Functions</b>	<b>7</b>
<b>4</b>	<b>Markov Chains</b>	<b>7</b>
<b>5</b>	<b>Generating Functions and Branching Processes</b>	<b>7</b>
<b>6</b>	<b>Convergence of Random Variables</b>	<b>7</b>

# 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  $\Omega$  is the *sample space* or the space of possible outcomes,  $\mathcal{F}$  is the set of events which are subsets of  $\Omega$  and  $\mathbb{P}: \mathcal{F} \rightarrow [0, 1]$  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

$$\Omega = \{1, 2, 3, 4, 5, 6\}. \quad (1.1)$$

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

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{|A|}{6} \text{ for all } A \in \mathcal{F}. \quad (1.2)$$

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

$$\mathbb{P}(\{2, 4, 6\}) = \frac{3}{6} = 0.5. \quad (1.3)$$

In this example, it is simple enough to use all possible subsets as 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  $\sigma$ -algebra.

**Definition 1.3** ( $\sigma$ -algebra). A non-empty collection of subsets of  $\Omega$  is a  $\sigma$ -algebra of  $\Omega$  if it satisfies 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  $\sigma$ -algebra is closed under complements and countable unions.

**Proposition 1.4.** *The definition of the  $\sigma$ -algebra implies that  $\sigma$ -algebras are also closed under countable intersection because*

$$\bigcap_i A_i = \left( \bigcup_i A_i^c \right)^c. \quad (1.4)$$

Minimally, we know that a  $\sigma$ -algebra must contain the empty set and the space on which it lives.

**Exercise 1.5.** Show that any  $\sigma$ -algebra of  $\Omega$  contains both the entire space  $\Omega$  and the empty set.

This formalizes the notion of events and combinations of events expressed earlier. Now, we'll develop a formal notion of probability using measure theory.

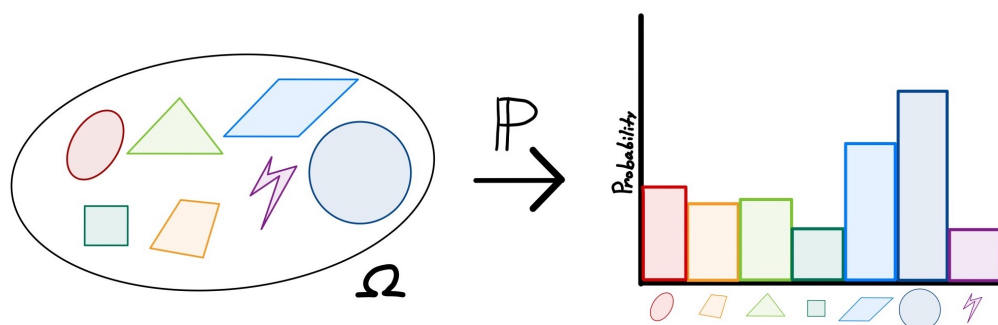
When we have a probability space  $\Omega$  and a  $\sigma$ -algebra  $\mathcal{F}$  on  $\Omega$ , we can define a measure on the space  $(\Omega, \mathcal{F})$ . *Describe what a measure is intuitively.*

**Definition 1.6** (Measure). A *measure* is a function  $\mu : \mathcal{F} \rightarrow [0, \infty)$  which satisfies:

- (i)  $\mu(A) \geq \mu(\emptyset) = 0$  for all  $A \in \mathcal{F}$
- (ii) if  $A_i \in \mathcal{F}$  is a countable sequence of disjoint sets, then

$$\mu\left(\bigcup_i A_i\right) = \sum_i \mu(A_i). \quad (1.5)$$

In the case that  $\mu(\Omega) = 1$ , we will call  $\mu$  a *probability measure* and denote it as  $\mathbb{P}$ .



Marlin Figgins (2020)  
Twitter: @marlinfiggins

Figure 1: A probability measure  $\mathbb{P}$  takes measurable subsets  $E \in \mathcal{F}$  to a numerical value in  $[0, 1]$ .

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

**Theorem 1.7** (Properties of measure).

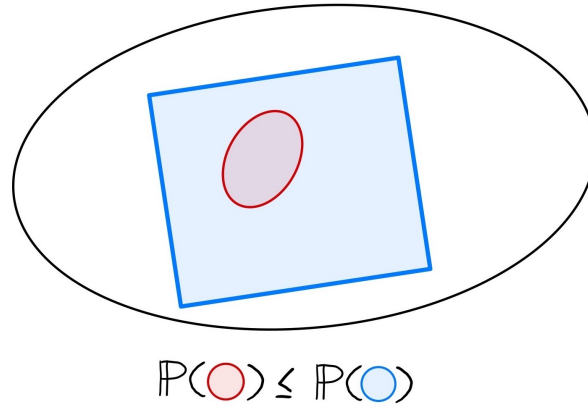
(i) Monotonicity. If  $A \subset B$ , then  $\mu(A) \leq \mu(B)$ .

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

$$\mu(A) \leq \sum_{j \in \mathbb{N}} \mu(A_j). \quad (1.6)$$

*Proof of monotonicity.* Finish proof. □

*Proof of subadditivity.* Finish proof. □



Marlin Figgins (2020)  
Twitter: @marlinfiggins

Figure 2: If a probability measure is consistent with our formulation, it ought to be monotonic. Appending additional outcomes to an event  $E$  should never decrease its probability of occurring.

**Theorem 1.8** (Continuity of measure).

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

$$\mu(A_i) \uparrow \mu(A) \text{ as } i \rightarrow \infty. \quad (1.7)$$

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

$$\mu(A_i) \downarrow \mu(A) \text{ as } i \rightarrow \infty. \quad (1.8)$$

**Exercise 1.9.** Prove that in a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$

$$\mathbb{P}(A) = 1 - \mathbb{P}(A^c) \text{ for any event } A \in \mathcal{F}. \quad (1.9)$$

## 1.1 Discrete Probability Spaces

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

**Theorem 1.10.** *Suppose we have a countable set  $\Omega$  with a  $\sigma$ -algebra  $\mathcal{F} = \mathcal{P}(\Omega)$ . Then any function  $p: \Omega \rightarrow [0, 1]$  such that*

$$\sum_{\omega \in \Omega} p(\omega) = 1 \quad (1.10)$$

*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*

$$\mathbb{P}(A) = \sum_{\omega \in A} p(\omega). \quad (1.11)$$

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

**Example 1.11** (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

$$\Omega = \{H, TH, TTH, TTH, \dots\}. \quad (1.12)$$

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  $(\frac{1}{2})^n$  for having  $n$  consecutive flips. This tells us that

$$p(\underbrace{T \cdots T}_n H) = \left(\frac{1}{2}\right)^n. \quad (1.13)$$

We can check that this indeed sums to one:

$$\sum_{\omega \in \Omega} p(\omega) = \sum_{i=1}^{\infty} \left(\frac{1}{2}\right)^i. \quad (1.14)$$

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

$$A = \{\text{Heads appears before third toss.}\} = \{H, TH\}, \quad (1.15)$$

$$B = \{\text{There are an even number of tails.}\} = \{H, TTH, TTTTH, \dots\}. \quad (1.16)$$

We can then compute these probabilities as:

$$\mathbb{P}(A) = \frac{1}{2} + \frac{1}{4}, \quad (1.17)$$

$$\mathbb{P}(B) = \frac{1}{2} + \frac{1}{2^3} + \frac{1}{2^5} + \cdots = \frac{2}{3}. \quad (1.18)$$

## 1.2 Uncountable probability spaces

We'll begin by further describing  $\sigma$ -algebra, so that we can understand what an uncountable  $\sigma$ -algebra must satisfy.

**Proposition 1.12.** The intersection of  $\sigma$ -algebras is again a  $\sigma$ -algebra. *Let  $I \neq \emptyset$  be an arbitrary set. If we have a collection of  $\sigma$ -algebras  $\mathcal{F}_i$  for  $i \in I$  defined on  $\Omega$ , then their intersection*

$$\bigcap_{i \in I} \mathcal{F}_i \text{ is a } \sigma\text{-algebra on } \Omega. \quad (1.19)$$

*Proof.* Per an earlier exercise, all  $\sigma$ -algebras in the collection  $\mathcal{F}_i$  must contain  $\emptyset$  and  $\Omega$ . Therefore, these sets are also in the intersection. For any event  $A$  that is in every  $\mathcal{F}_i$ ,  $A^c$  must be in  $\mathcal{F}_i$  since they are closed under complements. A similar argument holds for countable collections of events  $A_j$  and their unions.  $\square$

Thus, if we are given a sample space  $\Omega$  and  $\mathcal{A}$  a collection of subsets of  $\Omega$ , then there is a smallest  $\sigma$ -algebra which contains  $\mathcal{A}$ .

**Definition 1.13** ( $\sigma$ -algebra generated by  $\mathcal{A}$ ). Given a sample space  $\Omega$  and a collection of its subsets  $\mathcal{A}$ , the  $\sigma$ -algebra generated by  $\mathcal{A}$  is the smallest  $\sigma$ -algebra containing  $\mathcal{A}$ . We denote this as  $\sigma(\mathcal{A})$ .

**Example 1.14.** Given  $\Omega = \{1, 2, 3, 4\}$  and  $\mathcal{A} = \{\{1\}, \{1, 2\}\}$ , the smallest  $\sigma$ -algebra containing  $\mathcal{A}$  is

$$\sigma\mathcal{A} = \{\emptyset, \Omega, \{1\}, \{2, 3, 4\}, \{1, 2\}, \{3, 4\}, \{1, 3, 4\}, \{2\}\} \quad (1.20)$$

**Constructing uncountable probability spaces** Let's consider the experiment of drawing a random number between  $[0, 1]$ . In this case, our sample space is clearly  $\Omega = [0, 1]$ . In this case, we would like each number to be equally likely, but we can't simply set each number to occur with probability  $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$  as with the case of the fair die because our space is uncountable! Ideally, we want to find some function  $p(x) = p$  which describes the probability that we choose  $x \in [0, 1]$ . What is the probability that our chosen number is rational (in  $\mathbb{Q}$ ). By the countable additivity of our probability measure,

$$\mathbb{P}(\mathbb{Q} \cap [0, 1]) = \mathbb{P}\left(\bigcup_{x \in \mathbb{Q} \cap [0, 1]} \{x\}\right) \quad (1.21)$$

$$= \sum_{x \in \mathbb{Q} \cap [0, 1]} p(x) = \sum_{x \in \mathbb{Q} \cap [0, 1]} p. \quad (1.22)$$

For any positive real value of  $p$ , the rightmost sum is infinite which is nonsensical as the probability of any event cannot exceed one. Thus the probability of a single value must be exactly 0. By the countable additivity of measure, this means that any countable subset of  $[0, 1]$  must have probability 0.

**The Lebesgue measure on  $[0, 1]$**  We'll try an alternative approach to constructing probability spaces. We begin by defining the probabilities of certain subsets of interest and we extend this to the  $\sigma$ -algebra containing that collection of subsets. For example, we can specify that the probability of choosing a number between  $a$  and  $b$  is

$$\mathbb{P}((a, b]) = b - a \text{ for } a \text{ and } b \text{ with } 0 \leq a < b \leq 1. \quad (1.23)$$

This is a uniform probability measure on  $[0, 1]$  and it is called the *Lebesgue measure on  $[0, 1]$* .

**Borel  $\sigma$ -algebra** We can extend the Lebesgue measure described below by extending it to the  $\sigma$ -algebra generated by intervals of the form  $(a, b]$  by using the countable additivity property of measures. This is called the *Borel  $\sigma$ -algebra* on  $[0, 1]$  is denoted by  $\mathcal{B}([0, 1])$ .  $\mathcal{B}([0, 1])$  is the smallest  $\sigma$ -algebra containing all open sets in  $[0, 1]$ . Additionally, the extension of our measure to the Borel  $\sigma$ -algebra is unique by the Caratheodory Extension theorem. That being said, the Lebesgue measure cannot be extended to all subsets of  $[0, 1]$ . That being said, we will restrict our attention to study  $([0, 1], \mathcal{B}([0, 1]))$ .

**Example 1.15** (Infinite coin tosses). Suppose that we conduct an infinite sequence of coin tosses. Let the probability of each coin showing heads be  $p$  and the probability of tails be  $q = 1 - p$ . In this case, our sample space is all possible countable sequences of heads and tails. There is a one-to-one correspondence between  $\Omega$  and  $[0, 1]$  by considering the base two expansion of numbers between  $[0, 1]$ . As  $\Omega$  is uncountable, we begin by specifying the probability on some reasonable subset of events which we can expand to some larger  $\sigma$ -algebra. We'll represent  $w \in \Omega$  as a sequence  $\omega_1\omega_2\ldots$  where each  $w_i$  is the result of the  $i$ th coin toss. Let  $A_H = \{\omega \in \Omega \mid \omega_1 = H\}$ . That is,  $A_H$  is the event that the first coin toss is heads. Naturally, the probability of  $A_H$  should be the probability of flipping heads, so that

$$\mathbb{P}(A_H) = p \text{ and similarly } \mathbb{P}(A_T) = q = 1 - p. \quad (1.24)$$

These events generate a  $\sigma$ -algebra  $\mathcal{F}_1 = \sigma(\{A_H, A_T\})$ . In fact, we can define a series of  $\sigma$ -algebras  $\mathcal{F}_n$  based on the events of the first  $n$  coin tosses. For example, consider  $\mathcal{F}_2$ . This would contain events like:

$$A_{HH} = \{\omega \in \Omega \mid \omega_1 = H, \omega_2 = H\}, \quad A_{HT} = \{\omega \in \Omega \mid \omega_1 = H, \omega_2 = T\}, \quad (1.25)$$

$$A_{TH} = \{\omega \in \Omega \mid \omega_1 = T, \omega_2 = H\}, \quad A_{TT} = \{\omega \in \Omega \mid \omega_1 = T, \omega_2 = T\}. \quad (1.26)$$

These events would have probabilities:

$$\mathbb{P}(A_{HH}) = p^2, \quad \mathbb{P}(A_{HT}) = pq, \quad (1.27)$$

$$\mathbb{P}(A_{TH}) = pq, \quad \mathbb{P}(A_{TT}) = q^2. \quad (1.28)$$

We can similarly extend this to the events in  $\mathcal{F}_n$ . Due to our construction, we have an increasing sequence of  $\sigma$ -algebras  $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}_3 \ldots$  which describe the results of finite



coin tosses. We'll then define  $\mathcal{F}_\infty = \cup_{n \in \mathbb{N}} \mathcal{F}_n$ . This is **not** a  $\sigma$ -algebra because it is not closed under countable intersection. Consider the event  $w_H$  which is the infinite sequence of outcomes with all  $H$ . We also consider the outcomes  $\omega_{nH}$  which is where we count  $n$  heads in a row. As you can see,

$$\{\omega_H\} = \bigcap_{n \in \mathbb{N}} \{\omega_{nH}\}. \quad (1.29)$$

However,  $\{\omega_H\}$  is not in  $\mathcal{F}_n$  for any  $n$  since it depends on the results of all coin tosses. This means that  $\mathcal{F}_\infty$  cannot be a  $\sigma$ -algebra as it does not contain the countable union of  $\{\omega_{nH} \mid n \in \mathbb{N}\} \in \mathcal{F}_\infty$ . To remedy this, we can just use  $\mathcal{F} = \sigma(\mathcal{F}_\infty)$ . Once again, we can extend our probabilities to events in  $\mathcal{F}$  that are in  $\mathcal{F}_\infty$  uniquely. Unlike  $\mathcal{F}_\infty$ ,  $\mathcal{F}$  contains events which cannot be specified by any finite number of coin tosses like  $\omega_H$  which we discussed before. We can assign a probability to this event using the continuity of measure from above. Since  $\{\omega_{nH}\} \supset \{\omega_{(n+1)H}\}$  and  $\{\omega_H\} = \bigcap_{n \in \mathbb{N}} \{\omega_{nH}\}$ , it follows that

$$\mathbb{P}(\{\omega_{nH}\}) = p^n \downarrow P(\{\omega_H\}) = 0. \quad (1.30)$$

2. Independence, Martingales, and Conditioning
3. Characteristic Functions
4. Markov Chains
5. Generating Functions and Branching Processes
6. Convergence of Random Variables