EE 520: Random Processes **Fall 2021**

Lecture 1
The Probability Model

Instructor Name: John Lipor

Recommended Reading: Pishro-Nik: 1.1 - 1.3; Gubner: 1.1 - 1.4

1 The Probability Model

To think formally about probability, we require a model, which we refer to as a *probability space*. This space will consist of possible outcomes of an experiment, a set of outcomes we care about, and a way to measure the probability of outcomes. We now define these formally and give some examples.

Definition 1. The set of all possible outcomes of an experiment is the **sample space**, denoted by Ω.

Example 1. Toss a U.S. coin. The possible outcomes are $\Omega = \{\text{heads, tails}\}$. Since tossing coins is one of the most important things a probabilist can do, we will often abbreviate these, writing $\Omega = \{H, T\}$.

Example 2. Rolling a six-sided die gives $\Omega = \{1, 2, 3, 4, 5, 6\}$.

We're typically interested in subsets of outcomes (e.g., rolling an even number). We call these collections of outcomes *events*, and form them by taking a subset of all the possible outcomes. Note that this means any event is a set itself.

Definition 2. An **event** is a collection (or set) of outcomes.

Example 3. Roll a die. One possible event is $E = \{x \in \Omega : x \text{ is even}\}$, yielding $E = \{2, 4, 6\}$.

Example 4. Roll a die and consider $E = \{x \in \Omega : x \text{ is even and } x \leq 3\}$, which yields

$$E = \{2, 4, 6\} \cap \{1, 2, 3\} = \{2\} .$$

We now have a set of possible outcomes (sample space), as well as a way to define sets of outcomes that may be of interest (events). It turns out that in the case of uncountably infinite sample spaces (more on this later), the probability world would break if we tried to measure the probability of every possible event. Instead, we focus on a subset of Ω consisting of events whose probability we do want to measure; we call these events *measurable* and put them all in a set called a *sigma algebra*.

Definition 3. A **σ-algebra** or **σ-field** is a collection $\mathcal{F}$ of subsets of Ω such that

- $\emptyset \in \mathcal{F}$ ($\emptyset$ is called the empty set)
- $A \in \mathcal{F}$ implies $A^c \in \mathcal{F}$ for all events A (closed under complements)
- $A_1, A_2, \dots \in \mathcal{F}$ implies $\cup_{n=1}^{\infty} A_n \in \mathcal{F}$ (closed under countable unions)

Example 5. The smallest possible σ-algebra is $\mathcal{F} = \{\emptyset, \Omega\}$.

Example 6. Consider some event A. Then $\mathcal{F} = \{\emptyset, A, A^c, \Omega\}$ is a valid σ-algebra.

On Homework 1, you will think through the smallest σ-algebra containing two sets. In general, we don't think consciously about the underlying σ-algebra, but it is an important technical detail that comes into play when studying machine learning problems like multi-armed bandits and reinforcement learning. In the case of a countable sample space, we can safely measure all possible events, so we take take $\mathcal{F} = 2^{\Omega}$ (known as the *power set*).

Now that we have a space of outcomes and events we might want to think about, the last ingredient is how to measure the probability of a given event. This is done through a *probability measure*.

Definition 4. A **probability measure** P on a set $(\Omega, \mathcal{F})$ is a function $P : \mathcal{F} \to [0, 1]$ such that

- $P(\emptyset) = 0$ and $P(\Omega) = 1$ (called the impossible and almost sure events, respectively)
- $P(A) \geq 0$, $\forall A \in \mathcal{F}$ (probabilities are nonnegative)
- Let $A_1, A_2, \dots \in \mathcal{F}$ be disjoint, so that

  $$A_i \cap A_j = \emptyset,\ \forall i \neq j.$$

  Then

  $$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i),$$

  which is referred to as *countable additivity.*

Example 7. Roll a die. The probability of not rolling anything is zero ($P(\emptyset) = 0$), and the probability of getting something is 1 ($P(\Omega) = 1$). Let $A_1 = \{1, 3, 4\}$ and $A_2 = 5$. Then

$$P(A_1 \cup A_2) = P(\{1, 3, 4\} \cup \{5\}) = P(\{1, 3, 4\}) + P(\{5\}).$$

If the die is fair, we have

$$P(A_1 \cup A_2) = 1/2 + 1/6 = 4/6.$$

More generally, for an event A, we have

$$P(A) = \sum_{i \in A} p_i,$$

where p_i is the probability of rolling $i \in 1, 2, \dots, 6$. If the die is "fair," then $p_i = 1/6$ for all i, and we have

$$P(A) = \frac{1}{6} |A|,$$

where $|A|$ denotes the *cardinality* of A (number of elements in A).

In light of the fact that events are sets, a probability measure is simply a way to measure how "big" a given set is. As stated above, P takes in a measurable set and returns its "size," which in our case is a probability between 0 and 1. The three ingredients above are all required to define a probability model, which we can now do.

Definition 5. A **probability space** is a tuple $\{\Omega, \mathcal{F}, P\}$, where Ω is the sample space, $\mathcal{F}$ is a σ-algebra, and P is a probability measure.

1.1 Properties of Probability Measures

Our primary concern is with probability measures. Some important properties of these are below.

- $P(A) = 1 - P(A^c)$ (probability of complements)
- $A \subset B$ implies $P(A) \leq P(B)$ (monotonicity)
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (inclusion-exclusion principle)
- $P\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} P(A_i)$ (union bound / countable subadditivity)