Math 154: Probability Theory, Lecture Notes

Kevin Yang

CONTENTS

1. WEEK 1, STARTING TUE. JAN. 23, 2024

1.1. Probability spaces and events.

Definition 1.1. Take a set Ω. A σ-algebra $\mathscr{F}$ is a collection of subsets of Ω such that

- $\Omega, \emptyset \in \mathscr{F}$.
- If $\{A_n\}_{n=1}^{\infty}$ is a collection of sets in $\mathscr{F}$, then $\cup_{n=1}^{\infty} A_n \in \mathscr{F}$ and $\cap_{n=1}^{\infty} A_n \in \mathscr{F}$.

Sets in $\mathscr{F}$ are called *events*. A probability measure $\mathbb{P}$ on $(\Omega, \mathscr{F})$ is a function $\mathbb{P} : \mathscr{F} \to [0, 1]$ such that

- $\mathbb{P}(\emptyset) = 0$ and $\mathbb{P}(\Omega) = 1$
- If $\{A_n\}_{n=1}^{\infty}$ is a pairwise disjoint collection of sets in $\mathscr{F}$, then $\mathbb{P}(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$.
- If $\{E_n\}_{n=1}^{\infty}$ are in $\mathscr{F}$ and $E_1 \subseteq E_2 \subseteq \ldots$, then $\mathbb{P}(E_n) \to \mathbb{P}(\cup_{k=1}^{\infty} E_k)$.
- If $\{B_n\}_{n=1}^{\infty}$ are in $\mathscr{F}$ and $B_1 \supseteq B_1 \supseteq \ldots$, then $\mathbb{P}(B_n) \to \mathbb{P}(\cap_{n=1}^{\infty} B_n)$.
- The previous two bullet points are necessary parts of the definition. They must follow

The data $(\Omega, \mathscr{F}, \mathbb{P})$ is called a *probability space*.

Example 1.2. A coin is tossed. In this case, $\Omega = \{H, T\}$ (heads or tails). We can take $\mathscr{F} = 2^{\Omega}$. It contains $\{H, T\}$ (the coin lands heads or tails), $\{H\}$ (the coin lands heads), $\{T\}$ (the coin lands tails), and $\emptyset$ (the coin lands neither heads or tails). We have $\mathbb{P}(H) = 1 - \mathbb{P}(T)$, and $\mathbb{P}(\{H, T\}) = 1$ and $\mathbb{P}(\emptyset) = 0$. If it is a fair coin, then $\mathbb{P}(H), \mathbb{P}(T) = \frac{1}{2}$.

Example 1.3. A six-sided dice is thrown. $\Omega = \{1, 2, 3, 4, 5, 6\}$. We can take $\mathscr{F} = 2^{\Omega}$. In general, if Ω is finite, one should always take $\mathscr{F} = 2^{\Omega}$. If $X \in \mathscr{F}$ has size 1, then $\mathbb{P}(X) = \frac{1}{6}$. Then, use the additivity property to extend all of $\mathbb{P}$. (For example, $\mathbb{P}(\{1, 2\}) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$.

Lemma 1.4. *(1) $\mathbb{P}(A^C) = 1 - \mathbb{P}(A)$, where $A^C = \Omega \setminus A$.*
(2) If $B \supseteq A$, then $\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A) \geqslant \mathbb{P}(A)$.
(3) If $A_1, \ldots, A_n \in \mathscr{F}$, then

$$\mathbb{P}\left(\cup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i) - \sum_{i<j} \mathbb{P}(A_i \cap A_j) + \sum_{i<j<k} \mathbb{P}(A_i \cap A_j \cap A_k) - \ldots \tag{1.1}$$

$$+ (-1)^{n+1} \mathbb{P}(A_1 \cap \ldots \cap A_n). \tag{1.2}$$

For $n = 2$, this reduces to $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.
(4) If $A_1, \ldots, A_n, \ldots \in \mathscr{F}$, then $\mathbb{P}(\cup_{n=1}^{\infty} A_n) \leqslant \sum_{n=1}^{\infty} \mathbb{P}(A_n)$. This is the union bound

Lemma 1.5. *Let $\{A_n\}_{n=1}^{\infty}$ be in $\mathscr{F}$. Then $(\cup_{n=1}^{\infty} A_n)^C = \cap_{n=1}^{\infty} A_n^C$ and $(\cap_{n=1}^{\infty} A_n)^C = \cup_{n=1}^{\infty} A_n^C$. One can take $A_n = \emptyset$ or $A_n = \Omega$ for all $n \geqslant N$ for some N to take finite unions and intersections.*

Example 1.6. Let $A, B \in \mathscr{F}$. Suppose $\mathbb{P}(A) = \frac{3}{4}$ and $\mathbb{P}(B) = \frac{1}{3}$. We can bound $\mathbb{P}(A \cap B)$ as follows. First,

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cup B). \tag{1.3}$$

We know $\mathbb{P}(A \cup B) \leqslant 1$, so $\mathbb{P}(A \cap B) \geqslant \frac{3}{4} + \frac{1}{3} - 1 = \frac{1}{12}$. Also, we know $\mathbb{P}(A \cup B) \geqslant \mathbb{P}(A)$, so $\mathbb{P}(A \cap B) \leqslant \frac{3}{4} + \frac{1}{3} - \frac{3}{4} = \frac{1}{3}$.

1.2. **Conditional probability.**

Definition 1.7. Take $B \in \mathscr{F}$ so that $\mathbb{P}(B) > 0$. The *conditional probability of A given B* is

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}. \tag{1.4}$$

The idea is that one takes Ω, and restricts to a smaller probability space with set B. The σ-algebra is just given by taking $\mathscr{F}$ and intersecting with B (feel free to try to show that this is a σ-algebra). $\mathbb{P}(\cdot|B)$ is the "natural" probability measure on this probability space.

Example 1.8. Two fair dice are thrown. Condition on the first showing 3. What is the probability that the sum of the two rolls is > 6? Let A be the event where the sum of the two rolls is > 6 and B is the event where the first roll is a 3. We have

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A \cap B)}{\frac{1}{6}}. \tag{1.5}$$

Note that $A \cap B$ is the event where the second roll is $4, 5, 6$, and the first roll is a 3. In particular, there are 3 outcomes out of 36 that are okay, so the probability of $\mathbb{P}(A \cap B) = \frac{3}{36}$. This shows $\mathbb{P}(A|B) = \frac{1}{2}$.

Example 1.9. A coin is flipped twice independently. What is the probability that both are heads, given that one is a heads. It is not $\frac{1}{2}$. Indeed, let A be the event of two heads, and B is the event where one is a heads. There are four total outcomes, three of which have at least one heads. So $\mathbb{P}(B) = \frac{3}{4}$. On the other hand, $A \cap B$ is just the event of two heads, so its probability is $\frac{1}{4}$. This shows $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{1}{3}$.

Lemma 1.10 (Law of total probability)**.** *We say that $B_1, \ldots, B_n \in \mathscr{F}$ form a partition of Ω if they are pairwise disjoint, positive probability, and $\cup_{i=1}^n B_i = \Omega$. For any partition $B_1, \ldots, B_n$ and any event A, we have*

$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A|B_i)\mathbb{P}(B_i). \tag{1.6}$$

In particular, for any events A, B (where $B \neq \Omega, \emptyset$), we have $\mathbb{P}(A) = \mathbb{P}(A|B)\mathbb{P}(B) + \mathbb{P}(A|B^C)\mathbb{P}(B^C)$.

Theorem 1.11 (Bayes' formula)**.** *This will be helpful for the homework For any events A, B of positive probability, we have $\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}$.*

1.3. **Independence.**

Definition 1.12. We say events A, B are *independent* if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. Independent and disjoint are *totally different* notions! This is the same as $\mathbb{P}(A|B) = \mathbb{P}(A)$.

We say a family of events $\{A_i\}_{i=1}^{\infty}$ are *jointly independent* if $\mathbb{P}(\cap_{i=1}^n A_i) = \prod_{i=1}^n \mathbb{P}(A_i)$. We say it is *pairwise independent* if A_i, A_j are independent for all $i \neq j$.

Example 1.13. Let $\Omega = \{abc, acb, cab, cba, bca, bac, aaa, bbb, ccc\}$. Each element in Ω occurs with probability $\frac{1}{9}$. Let A_k be the event where the k-th letter (for $k = 1, 2, 3$) is a. We know that A_1, A_2, A_3 are pairwise independent. Indeed, $A_1 \cap A_2$ is the event where the first and second letter are both a. Thus, $A_1 \cap A_2 = \{aaa\}$, so $\mathbb{P}(A_1 \cap A_2) = \frac{1}{9}$. Note that $\mathbb{P}(A_1)\mathbb{P}(A_2) = \frac{1}{3}\frac{1}{3} = \frac{1}{9}$. Similar arguments apply to A_1, A_3 and A_2, A_3 (try it!).

But, A_1, A_2, A_3 are not jointly independent. Indeed, $A_1 \cap A_2 \cap A_3 = \{aaa\}$, so its probability is $\frac{1}{9}$. But $\mathbb{P}(A_1)\mathbb{P}(A_2)\mathbb{P}(A_3) = \frac{1}{3}\frac{1}{3}\frac{1}{3} = \frac{1}{27}$.

Example 1.14. We pick a card uniformly at random from a deck of 52. Each has probability $\frac{1}{52}$. Let A be the event where a king is picked, and B is the event where a spade is picked. Then $\mathbb{P}(A) = \frac{4}{52} = \frac{1}{13}$, and $\mathbb{P}(B) = \frac{1}{4}$. Also, $\mathbb{P}(A \cap B) = \frac{1}{52}$. So A, B are independent.

Lemma 1.15. *If A, B are independent, then A^C, B are independent and A^C, B^C are independent.*

Example 1.16. Two fair dice are rolled independently. Let A be the event where the sum of the rolls is 7. Let B be the event where the first roll is 1. Then A, B are independent. Indeed, $\mathbb{P}(A|B) = \frac{1}{6}$ (since a six is needed on the second roll). But $\mathbb{P}(A) = \frac{6}{36}$, since for any value of the first roll, there is exactly one value of the second roll to realize A. If we change 7 to 1, then A, B are no longer independent.

1.4. **Some examples.**

(1) (Symmetric random walk, "gambler's ruin") Let's play a game. We flip a coin repeatedly. If it lands heads, I get one dollar. If it lands tails, I lose a dollar. (Suppose this is a fair coin for now.) I want to save N dollars, at which point I stop the game, so that I can retire happily. But if I end up with zero dollars at any point, we stop the game, since I can't play anymore.

Suppose I start with $0 < k < N$ dollars. What is the probability that I win?

- Let $p_k = \mathbb{P}_k(A)$ be the event that I win if we start at k dollars. By the law of total probability, if B is the event that we toss a heads, then
$$\mathbb{P}_k(A) = \mathbb{P}_k(A|B)\mathbb{P}(B) + \mathbb{P}_k(A|B^C)\mathbb{P}(B^C). \tag{1.7}$$
We have $\mathbb{P}_k(A|B) = p_{k+1}$ and $\mathbb{P}_k(A|B^C) = p_{k-1}$ and $\mathbb{P}(B), \mathbb{P}(B^C) = \frac{1}{2}$. So $p_k = \frac{1}{2}(p_{k+1} + p_{k-1})$. But also $p_0 = 0$ and $p_N = 1$. We will talk later in this class about how to solve this equation efficiently, but one can check that $p_k = 1 - \frac{k}{N}$ solves this equation.

2. Week 2, starting Tue. Jan. 30, 2024

2.1. **Random variables.**