EECS126 Course Notes

Anmol Parande

Spring 2021 - Professor Thomas Courtade

Disclaimer: These notes reflect 126 when I took the course (Spring 2021). They may not accurately reflect current course content, so use at your own risk. If you find any typos, errors, etc, please raise an issue on the GitHub repository.

Contents

1 Introduction to Probability

Definition 1 *A probability space is a triple $(\Omega, \mathcal{F}, P)$ where Ω is a set of objects called the sample space, $\mathcal{F}$ is a family of subsets of Ω called events, and the probability measure $P : \mathcal{F} \to [0, 1]$.*

One key assumption we make is that $\mathcal{F}$ is a σ-algebra containing Ω, meaning that countably many complements, unions, and intersections of events in $\mathcal{F}$ are also events in $\mathcal{F}$. The probability measure P must obey **Kolmogorov's Axioms**.

1. $\forall A \in \mathcal{F},\ P(A) \geq 0$
2. $P(\Omega) = 1$
3. If $A_1, A_2, \dots \in \mathcal{F}$ and $\forall i \neq j,\ A_i \bigcap A_j = \emptyset$, then $P\left(\bigcup_{i\geq 1} A_i\right) = \sum_{i\geq 1} P(A_i)$

We choose Ω and $\mathcal{F}$ to model problems in a way that makes our calculations easy.

Theorem 1

$$P(A^c) = 1 - P(A)$$

Theorem 2 (Inclusion-Exclusion Principle)

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{k=1}^{n} (-1)^{k+1} \left(\sum_{1\leq i_1 < \cdots < i_k \leq n} P(A_{i_1} \cap \cdots \cap A_{i_k})\right)$$

Theorem 3 (Law of Total Probability) *If $A_1, A_2, \cdots$ partition Ω (i.e A_i are disjoint and $\cup A_i = \Omega$), then for event B,*

$$P(B) = \sum_i P(B \cap A_i)$$

1.1 Conditional Probability

Definition 2 *If B is an event with $P(B) > 0$, then the conditional probability of A given B is*

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Intuitively, conditional probabilty is the probability of event A given that event B has occurred. In terms of probability spaces, it is as if we have taken $(\Omega, \mathcal{F}, P)$ and now have a probabilty measure $P(\cdot|C)$ belonging to the space $(\Omega, \mathcal{F}, P(\cdot|C))$.

Theorem 4 (Bayes Theorem)

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

1.2 Independence

Definition 3 *Events A and B are independent if $P(A \cap B) = P(A)P(B)$*

If $P(B) > 0$, then A, B are independent if and only if $P(A|B) = P(A)$. In other words, knowing B occurred gave no extra information about A.

Definition 4 *If A, B, C with $P(C) > 0$ satisfy $P(A \cap B|C) = P(A|C)P(B|C)$, then A and B are conditionally independent given C.*

Conditional independence is a special case of independence where A and B are not necessarily independent in the original probability space which has the measure P, but are independent in the new probability space conditioned on C with the measure $P(\cdot|C)$.

2 Discrete Probability

Definition 5 *A random variable is a function $X : \Omega \to \mathbb{R}$ with the property $\forall \alpha \in \mathbb{R},\ \{\omega \in \Omega : \ X(\omega) \leq \alpha\} \in \mathcal{F}$.*

The condition in definition 5 is necessary to compute $P(X \leq \alpha),\ \forall \alpha \in \mathbb{R}$. In a sense, it binds the probability space to the random variable.