Math 502AB - Lecture 1

Dr. Jamshidian

August 21, 2017

1 Chapter 1: Probability Theory

- The foundation of statistics is probability theory, and the foundation of probability theory is set theory.
- **Definition:**

  The set of all possible outcomes of an experiment is referred to as the **sample space**, and it is denoted by $\mathcal{S}$
- **Definition:**

  An **event** is defined as any subset of the sample space
- Our task: Define a probability function $P(\cdot)$ on subsets of $\mathcal{S}$ which give us probabilities
- **Example:**

  Say you flip a coin two times. The sample space is then:

$$\mathcal{S} = \{HH, HT, TH, TT\}$$

1.1 Basic Definitions

1. **Probability Space:** A *probability space* is a triplet

$$(\mathcal{S}, \mathcal{B}, P)$$

   where $\mathcal{S}$ is a set of outcomes, $\mathcal{B}$ is a set of events (A σ-algebra, or Borel field, which is defined as sets over your sample space), and P which is a function that maps $P : \mathcal{B} \to [0, 1]$, i.e. assigning probabilities to elements of $\mathcal{B}$ (events).

2. **σ-algebra:** If $\mathcal{B}$ is a σ-algebra, then it consists of subsets of $\mathcal{S}$ which satisfy the following properties:

   (a) $\emptyset \in \mathcal{B}$

   (b) $A \in \mathcal{B} \Rightarrow A^c \in \mathcal{B}$

(c) If $A_1, A_2, ... \in \mathcal{B}$, then $\bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{B}$

Examples of σ-algebra, or Borel Fields:

(a) Trivial σ-algebra (Borel field)

$$\mathcal{B} = \{\emptyset, \mathcal{S}\}$$

(b) Consider the set $\mathcal{S} = \{\text{H,T}\}$, the set of outcomes of a coin flip. Then a σ-algebra would be:

$$\mathcal{B} = \{\emptyset, \{H\}, \{T\}, \{H, T\}\}$$

(c) Given the sample space $\mathcal{S} = (-\infty, \infty)$, then the σ-algebra is the set of all intervals of the form $[a, b]$, $(a, b]$, $[a, b)$, (a, b)

3. Without P, the couple $(\mathcal{S}, \mathcal{B})$ is called a **measurable space**. This means that, if we have these two things, then we can put a measure on $\mathcal{B}$

4. **Measure:** A non-negative countably additive set function, that is, a function:

$$\mu : \mathcal{B} \to \mathbb{R}$$

with the following parameters:

(a) $\mu(A) \geq \mu(\emptyset) = 0$, $\forall A \in \mathcal{B}$

(b) If $A_1, A_2, ... \in \mathcal{B}$ is a countable or finite sequence of disjoint sets in $\mathcal{B}$, then:

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

Essentially, since the sets are *disjoint* we can add together the measures.

5. **Probability Measure:** Following up from the previous definition of *measure space*, if $\mu(\mathcal{S}) = 1$, we call μ a **probability measure**, and we denote it by $P(\cdot)$.

1.2 Axioms of Probability

To sum, a probability measure, $P(\cdot)$, is defined on $(\mathcal{S}, \mathcal{B})$ with the following properties (axioms):

1. $P(A) \geq 0$, $\forall A \in \mathcal{B}$
2. $P(\mathcal{S}) = 1$
3. If $A_1, A_2, ...$ are disjoint, then $P(\bigcup\limits_{i=1}^{\infty} A_i) = \sum\limits_{i=1}^{\infty} P(A_i)$. This is also referred to as **Countable Additivity**.

1.2.1 A Few Results:

1. $P(\emptyset) = 0$

   Proof: Let $A_1 = A_2 = ...$ be a set of $\emptyset$ sets. Obviously, if they are empty sets, then they are disjoint ($A_i \cap A_j = \emptyset$, $\forall i \neq j$). By set theory, we have:

   $$P(\emptyset) = P(\bigcup_{i=1}^{\infty} A_i)$$

   Using the third axiom, we can write this as:

   $$\begin{aligned} P(\bigcup_{i=1}^{\infty} A_i) &= \sum_{i=1}^{\infty} P(A_i) \\ \Rightarrow P(\emptyset) &= \sum_{i=1}^{\infty} P(\emptyset) \\ &\Rightarrow P(\emptyset) = 0, \quad \text{by axiom 1} \end{aligned}$$

2. $P(A^c) = 1 - P(A)$
3. If A and B are two events $P(B \cap A^c) = P(B \backslash A) = P(B) - P(A \cap B)$
4. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
5. $A \subset B \Rightarrow P(A) \leq P(B)$

1.2.2 A Few Notes:

- Countable Additivity $\Rightarrow$ Finite Additivity
- **Finite Additivity:** If $A_1, A_2, ..., A_n$ are n disjoint sets, then:

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

  Proof: Let $A_{n+1} = A_{n+2} = ...$ all be $\emptyset$. Then we have:

  $$\begin{aligned} P\left(\bigcup_{i=1}^{n} A_i\right) &= P\left[\left(\bigcup_{i=1}^{n} A_i\right) \bigcup \left(\bigcup_{i=n+1}^{\infty} A_i\right)\right] \\ &= P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i), \quad \text{by countable additivity} \\ &= \sum_{i=1}^{n} P(A_i) + \sum_{i=n+1}^{\infty} P(A_i) \\ &= \sum_{i=1}^{n} P(A_i), \quad \text{as wanted.} \end{aligned}$$

1.2.3 $P(\cdot)$ is a Continuous set Function

In calculus, we said $f(x)$ is continuous at $x = a$ if:

$$f(a) = f\left(\lim_{x \to a} x\right) = \lim_{x \to a} f(x)$$

Let $A_1 \supset A_2 \supset ...$ be a decreasing sequence of events. Then:

$$\lim_{n \to \infty} A_n = \bigcap_{i=1}^{\infty} A_i$$

Similarly, let $A_1 \subset A_2 \subset ...$ be a sequence of increasing events. Then:

$$\lim_{n \to \infty} A_n = \bigcup_{i=1}^{\infty} A_i$$

We will prove next class a theorem which will tell us:

$$P\left[\lim_{n \to \infty} A_i\right] = \lim_{n \to \infty} P(A_i)$$