Measure Theory, Limit Theorems, and Random Variables

Lee De Zhang

December 2, 2020

1 Introduction

This chapter outlines the tools necessary for a more rigorous study of probability.

2 Measure Theory

An algebra $\mathcal{F}$ is said to be a σ-algebra if

1. It is closed under complement, i.e., $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$.
2. Closed under countable union, i.e., $A_i \in \mathcal{F} \Rightarrow \cup_i A_i \in \mathcal{F}$ for countable number of A_i.

Remark 1. *Countable refers to finite or countably infinite. Some textbooks also state closed under countable intersection, but it is easy to derive that from the above conditions.*

.

We will now state the definition of a probability space

Definition 2.1 (*Probability Space*)**.** *A probability space is the triple $(\Omega, \mathcal{F}, P)$, such that Ω is the support or set of all possible outcomes (sample space), $\mathcal{F}$ is a set of events, and $P : \mathcal{F} \to [0, 1]$ which assigns probabilities to items in $\mathcal{F}$. $\mathcal{F}$ is a σ-algebra, which is a non-empty collection of subsets of Ω and satisfies the properties above.*

Without P, the double $(\Omega, \mathcal{F})$ is a measurable space, i.e., something which we can assign a measure to. We will define a measure next.

Definition 2.2 (*Measure*)**.** *A measure is a non-negative countably additive set function, such that given a σ-algebra $\mathcal{F}$ and a measure μ, then $\mu : \mathcal{F} \to \mathbb{R}$ such that*

1. $\mu(A) \geq \mu(\emptyset) = 0 \quad \forall A in \mathcal{F}$.
2. *For $i = 1, 2, \cdots$, if $A_i \in \mathcal{F}$ is a countable sequence of disjoint sets, then $\mu(\cup A_i) = \sum \mu(A_i)$.*

Remark 2. *If $\mu(\Omega) = 1$, then μ is a probability measure.*

.

From the above definition, we can deduce that a measure has the following properties.

Theorem 2.3 (*Properties of a measure*). *Suppose μ is some measure on a measurable space $(\Omega, \mathcal{F})$, then μ satisfies,*

1. *Monotonicity.* $A \subset B \Rightarrow P(B) - P(A) - P(B - A) \geq 0$ *for* $A, B \in \mathcal{F}$.
2. *Subadditivity. For* $m = 1, \cdots$, *if* $A_m \in \mathcal{F}, A \subset \cup_{m=1}^{\infty} A_m$, *then* $P(A) \leq \sum_{m=1}^{\infty} P(A_m)$.
3. *Continuity from below. If* $A_i \uparrow A$, *i.e,* $A_1 \subset A_2 \subset \cdots$ *and* $\cup A_i = A$, *then* $P(A_i) \uparrow P(A)$.
4. *Continuity from above. If* $A_i \downarrow A$, *then* $P(A_i) \downarrow P(A)$.

Proof. **Monotonicity**. Since $B - A = B \cap A^c$ and A and $B - A$ are disjoint, then $\mu(A) = \mu(B) + \mu(B - A)$ and the result follows from the the non-negativity of μ.

Subadditivity. Let $B_1 = A_1$ and $B_k = A_k - \cup_{i=1}^{k-1} A_i$, then $\mu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(B_i)$, the result follows from monotonicity.

Continuity from below. Let $B_k = A_k - A_{k-1}$. Then $\mu(A_k) = \sum_{i=1}^{k} \mu(B_i)$. Then the result follows.

Continuity from above. Observe that $A_i^c \uparrow A$, then $P(A_i^c) \uparrow P(A^c)$ and the result follows. □

To make things a bit more concrete, let us apply this to an example on the discrete probability space.

Example 2.1 (*Discrete Probability Spaces*). *Let Ω be a countable set, $\mathcal{F}$ be the set of all subsets of Ω, for $A \in \mathcal{F}$, $P(A) = \sum(x \in A)p(x) \geq 0$ and $\sum_{\omega \in \Omega} p(\omega) = 1$. If Ω is finite, then $p(\omega) = 1/|\Omega|$. For example, when rolling a fair dice, $\Omega = \{1, 2, 3, 4, 5, 6\}$.*

Example 2.2 (*Measure on continuous real line*). *Let $\mathbb{R}$ be the real line, $\mathcal{R}$ be the Boreal sets (smallest σ-algebra containing the open sets), λ be the Lebesgue measure (for $a < b$, $\lambda((a, b]) = b - a$, and $\lambda(\mathbb{R}) = \infty$. To get a probability space, let $\Omega = (0, 1)$, $\mathcal{F} = \{A \cap (0, 1) : A \in \mathcal{R}\}$, and $P = \lambda$. Then P is a Lebesgue measure restricted to the Borel subsets in interval (0,1).*

Theorem 2.4 (*σ-algebra closed under arbitrary intersection*). *If $\mathcal{F}_i, i \in I$ for an arbitrary (possibly uncountable) I, then $\cap_i \mathcal{F}_i$ is a σ-algebra.*

Proof. Since $\emptyset, \Omega \in \mathcal{F}_i$, then $\emptyset, \Omega \in \cap \mathcal{F}_\rangle$. Furthermore, $A \in \cap \mathcal{F}_i \Rightarrow A \in \mathcal{F}_i$ hence $\cap \mathcal{F}_i$ closed under intersection and complement. □

Definition 2.5 (*Smallest σ-field*). *Let $\mathcal{A}$ be a collection of subsets of Ω. Then there is a smallest σ-field containing $\mathcal{A}$, written as $\sigma(\mathcal{A})$.*

Remark 3. *The existence of $\sigma(\mathcal{A})$ is guaranteed using Theorem 2.4. Namely, let F_i be the different σ-algebra that contain $\mathcal{A}$. Then an intersection yields the smallest one.*

Theorem 2.6 (*Measure on Product Spaces*)**.** *If* $(\Omega_i, \mathcal{F}_i, P_i), i = 1, \cdots, n$ *are probability spaces, let* $\Omega = \Omega_1 \times \cdots \times \Omega_n = \{(\omega_1, \cdots, \omega_n) : \omega_i \in \Omega_i\}$, $\mathcal{F} = \prod_{i=1}^n \mathcal{F}_i$ *is the* σ*-field generated by* $\{A_1 \times \cdots \times A_n : A_i \in \mathcal{F}_i\}$. *Then* P *is a measure on* $\mathcal{F}$, *given by,*

$$P(A_1 \times \cdots \times A_n) = \prod_{i=1}^n P_i(A_i).$$

We can define a random variable on probability spaces. A real valued function X defined on Ω is a random variable if for every borel set $B \subset \mathbb{R}$,

$$X^{-1}(B) = \{\omega : X(\omega) \in B\} \in \mathcal{F}.$$

If X is a random variable, then X induces a probability measure on $\mathbb{R}$ called its distribution, by setting $\mu(A) = P(X \in A) = P(X^{-1}(A)) = P(\{\omega : X(\omega) \in B\})$.