Introduction to Probability Measure Spaces

1 Probability Space

Definition. A *probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ is given by

1. a set Ω
2. a σ-algebra $\mathcal{F}$ of subsets of Ω
3. a function $\mathbb{P} : \mathcal{F} \to [0,1]$ where $\mathbb{P}(\Omega) = 1$ and $\mathbb{P}$ is a measure on $\mathcal{F}$

We call Ω the *sample space*, and an element $\omega \in \Omega$ is an *outcome*. Elements of $\mathcal{F}$ are referred to as *events* and $\mathbb{P}(E)$ is the *probability* of an event E. In particular, we say $(\Omega, \mathcal{F}, \mathbb{P})$ is a *discrete probability space* if Ω is finite or countably infinite.

Here is a simple example of a discrete probability space: let $\Omega = \{H, T\}$ be the result of a flip of a p-coin. Then, $\mathcal{F} = \{\emptyset, \{H\}, \{T\}, \{H, T\}\}$, with probabilities $0, p, 1-p, 1$.

We now discuss some elementary concepts in probability from a measure theoretic viewpoint.

2 Independence

The notion of independence in probability spaces is the same as what we are used to.

Definition. Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Two events A, B are *independent* if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B)$$

A collection of events $\{A_i\}$ is *independent* if for any subcollection $\{A_j\}$,

$$\mathbb{P}(\cap_j A_j) = \sum_j \mathbb{P}(A_j)$$

3 Random Variables

We start by defining a Borel set.

Definition. The *Borel σ-algebra* of $\mathbb{R}$ is the smallest σ-algebra containing all open sets of $2^{\mathbb{R}}$. A *Borel set* is any element of the Borel σ-algebra.

Definition. A *random variable* on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a function $X : \Omega \to \mathbb{R}$ such that for every Borel set B, $X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\}$. A random variable X is *discrete* if the range of X is finite or countably infinite, and *continuous* otherwise.

Note that we can define a function on the Borel sets $\mathcal{B}$ by $\mu_X : \mathcal{B} \to \mathbb{R}$ that sends B to $\mathbb{P}(X^{-1}(B))$, and this forms a probability measure space $(\mathbb{R}, \mathcal{B}, \mu_X)$.

This gives rise to the following definition:

Definition. The *distribution function* for a random variable X is given by

$$F_X(x) = \mu_X((-\infty, x])$$

We can think of this as computing the probability that X takes on value less than or equal to x. In particular, for continuous random variables, we have the following definition:

Definition. If for a continuous random variable X, there exists a function $f : \mathbb{R} \to [0, \infty)$ such that

$$\mu_X((a,b)) = \int_a^b f(x)dx$$

for all $a < b \in \mathbb{R}$, we call f the *density* of X.

Note that the density must integrate to 1 over $\mathbb{R}$. Furthermore, it is not necessarily true that a continuous random variable has a density.

4 Expectation

Definition. For a non-negative random variable X, we define its *expectation* by

$$\mathbb{E}(X) = \int X d\mathbb{P}$$

For a random variable X that takes on positive and negative values, if $\mathbb{E}(|X|) < \infty$, we define its *expectation* by

$$\mathbb{E}(X) = \int X d\mathbb{P}$$

Otherwise, we say that the expectation is not defined. Note that the integrals in this definition are all Lebesgue integrals.

For the discrete case, if the range of X consists of values $a_1, a_2, ...$, we have the standard formula

$$\mathbb{E}(X) = \sum_j a_j \mathbb{P}(X = a_j)$$

Lemma. For a random variable X with distribution μ_X,

$$\mathbb{E}(X) = \int_{\mathbb{R}} x d\mu_X$$

The idea for the proof of this lemma is to assume X is non-negative, split up $\mathbb{R}_{\geq 0}$ into countably many intervals $[\frac{k-1}{n}, \frac{k}{n})$, and define a discrete random variable X_n that is equal to $\frac{k}{n}$ whenever X is on this interval. Then, we have that $X_n - \frac{1}{n} \leq X \leq X_n$, so taking expectations and applying the discrete formula, we obtain the result as n approaches infinity. To do the general case we can split X into positive and negative parts.

Definition. A *Borel measurable function* is a function f such that $f^{-1}(U)$ is a Borel set for any open set U.

Lemma. For a random variable X with distribution μ_X and a Borel measurable function g,

$$\mathbb{E}(X) = \int_{\mathbb{R}} g(x) d\mu_X$$

The idea for the proof of this lemma is to approximate a non-negative g by an increasing sequence of simple functions, for which the result follows easily. Then, applying the monotone convergence theorem yields the result. We can prove the general case by splitting into positive and negative parts of g.