1 Probability Review

Definition 1.1. A probability space \((\Omega, \mathcal{F}, P)\) consists of set of all possible outcomes denoted by \(\Omega\) and called a sample space, a collection of subsets \(\mathcal{F}\) of sample space, and a non-negative set function probability \(P : \mathcal{F} \to [0, 1]\), with the following properties.

1. The collection of subsets of \(\mathcal{F}\) is a \(\sigma\)-algebra, that is it contains an empty set and is closed under complements and countable unions.
2. Set function \(P\) satisfies \(P(\Omega) = 1\), and for every countable pair-wise disjoint collection \(\{A_n \in \mathcal{F} : n \in \mathbb{N}\}\), we have

\[
P(\bigcup_n A_n) = \sum_n P(A_n).
\]

There is a natural order of inclusion on sets through which we can define monotonicity of probability set function \(P\). To define continuity of this set function, we need to define limits of sets.

Definition 1.2. For a sequence of sets \(\{A_n : n \in \mathbb{N}\}\), we define limit superior and limit inferior of this sequence respectively as

\[
\lim\sup_n A_n = \bigcap_n \bigcup_{k \geq n} A_k,
\]

\[
\lim\inf_n A_n = \bigcup_n \bigcap_{k \geq n} A_k.
\]

We say that limit exists if the limit superior and limit inferior are equal, and is equal to the limit of the sequence of sets.

Lemma 1.3. Probability set function is monotone and continuous.

Definition 1.4. A real valued random variable \(X\) on a probability space \((\Omega, \mathcal{F}, P)\) is a function \(X : \Omega \to \mathbb{R}\) such that for every \(x \in \mathbb{R}\), we have \(\{\omega \in \Omega : X(\omega) \leq x\} = X^{-1}(-\infty, x] \in \mathcal{F}\).

Definition 1.5. For a random variable \(X\) defined on probability space \((\Omega, \mathcal{F}, P)\), the distribution function \(F : \mathbb{R} \to [0, 1]\) is defined as

\[
F(x) = (P \circ X^{-1})(-\infty, x], \forall x \in \mathbb{R}.
\]

Lemma 1.6. Distribution function \(F\) of a random variable \(X\) is non-negative, monotone increasing, continuous from the right, and has countable points of discontinuities. Further, if \(P \circ X^{-1}(\mathbb{R}) = 1\), then

\[
\lim_{x \to -\infty} F(x) = 0, \quad \lim_{x \to \infty} F(x) = 1.
\]
Definition 1.7. Let $X$ be a random variable on a probability space $(\Omega, \mathcal{F}, P)$ with a distribution function $F$. Let $g : \mathbb{R} \to \mathbb{R}$ be a function. Then, the expectation of $g(X)$ is defined as
\[ E[g(X)] = \int_{x \in \mathbb{R}} g(x) dF(x). \]

2 Stochastic Processes

Definition 2.1. A collection of random variables $\{X_t : t \in T\}$ each defined on the same probability space $(\Omega, \mathcal{F}, P)$ is called a random process for an arbitrary index set $T$.

To define a measure on collection of random variables, we need to know its joint distribution $F : \mathbb{R}^T \to [0, 1]$. To this end, for any $x \in \mathbb{R}^T$ we need to know $F(x) = P\left( \bigcap_{t \in T} \{\omega \in \Omega : X_t(\omega) \leq x_t\} \right)$.

When the index set $T$ is infinite, any function of the above form would be zero if $x_t$ is finite for all $t \in T$. Therefore, we only look at the values of $F(x)$ when $x_t \in \mathbb{R}$ for indices $t$ in a finite set $S$ and is $x_t = \infty$ for all $t \notin S$. This leads to finite-dimensional distributions as defined below.

Definition 2.2. For any finite set $S \subseteq T$ and $x_S = \{x_t \in \mathbb{R} : t \in S\}$, we can define a finite dimensional distribution $F_S(x) = P\left( \bigcap_{t \in S} \{\omega \in \Omega : X_t(\omega) \leq x_t\} \right)$.

Set of all finite dimensional distributions of the stochastic process $\{X_t : t \in T\}$ characterizes its distribution completely.

3 Examples of Stochastic Processes

Example 3.1 (Queues). Queues are complex stochastic processes and consist of two stochastic processes arrival and service, coupled through a buffer. Number of arrivals and arrival instants could be discrete or continuous random variable. For a discrete arrival case, arrival process can be characterized by the time epochs of discrete arrivals, denoted $\{A_n : n \in \mathbb{N}\}$. Similarly, service requirement of each incoming arrival can also be a discrete or continuous random variable. Service of each discrete arrival can be considered to be a random amount of time, $\{S_n : n \in \mathbb{N}\}$. Queue can have a finite or infinite waiting area, and can be served by single or multiple servers. Important performance metrics for queues are mean waiting time of arrivals and mean queue length. These metrics are affected by the service policy that determines how to serve incoming arrivals. Few important service policies are first come first out, last in first out, processor sharing etc. Queues have applications in operations research, industrial engineering, telecommunications networks, among others.

Example 3.2 (Gambler’s ruin). One can model many gambling games with random walks, where wins or losses on each bet can be thought of as a random step. If a gambler starts with certain capital, and he wants to quit gambling after he makes a certain amount of money, one is interested in probability of a gambler getting bankrupt before it can quit gambling. These questions are related to hitting times of a random walk. Random walks have deep relations to Brownian motion.
Example 3.3 (Urn Models). In these models, one is interested in ball and urns. One is interested in distribution of balls in urns, when one can randomly throw balls into urns. Balls can be of multiple colors and may or may not be replaced after putting into urns. These models have applications in influence maximization and epidemic control.

Example 3.4 (Branching Processes). This is used by biologists to model population. In this model, one assumes that every individual in a population has a probability distribution over number of children it can have. Each children can be assumed to have identical and independent distribution for their progenies in next generation. These type of models can answer questions related to survival of species.

Example 3.5 (Random Graphs). A typical graph $G$ consists of vertex set $V$ and edge set $E \subseteq V \times V$. Both of these can be random in general. In classical settings, usually $V = [n]$ and $E$ is selected randomly from set of all possible edges $[n] \times [n]$, without self-loops. These models are exploited in study of various type of networks, and can be used to answer questions regarding network properties.