IEOR 151 – Lecture 1
Probability Review

1 Definitions in Probability and Their Consequences

1.1 Defining Probability

A probability space $(\Omega, \mathcal{F}, \mathbb{P})$ consists of three elements:

- A sample space Ω is the set of all possible outcomes.
- The σ-algebra $\mathcal{F}$ is a set of events, where an event is a set of outcomes.
- The measure $\mathbb{P}$ is a function that gives the probability of an event. This function $\mathbb{P}$ satisfies certain properties, including: $\mathbb{P}(A) \geq 0$ for an event A, $\mathbb{P}(\Omega) = 1$, and $\mathbb{P}(A_1 \cup A_2 \cup \ldots) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \ldots$ for any countable collection $A_1, A_2, \ldots$ of mutually exclusive events.

Some useful consequences of this definition are:

- For a sample space $\Omega = \{o_1, \ldots, o_n\}$ in which each outcome o_i is equally likely, it holds that $\mathbb{P}(o_i) = 1/n$ for all $i = 1, \ldots, n$.
- $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$, where $\overline{A}$ denotes the complement of event A.
- For any two events A and B, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.
- If $A \subseteq B$, then $\mathbb{P}(A) \leq \mathbb{P}(B)$.
- Consider a finite collection of mutually exclusive events $B_1, \ldots, B_m$ such that $B_1 \cup \ldots \cup B_m = \Omega$ and $\mathbb{P}(B_i) > 0$. For any event A, we have $\mathbb{P}(A) = \sum_{k=1}^m \mathbb{P}(A \cap B_k)$.

1.2 Conditional Probability

The conditional probability of A given B is defined as

$$\mathbb{P}[A|B] = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Some useful consequences of this definition are:

- Law of Total Probability: Consider a finite collection of mutually exclusive events $B_1, \ldots, B_m$ such that $B_1 \cup \ldots \cup B_m = \Omega$ and $\mathbb{P}(B_i) > 0$. For any event A, we have

$$\mathbb{P}(A) = \sum_{k=1}^m \mathbb{P}[A|B_k]\mathbb{P}(B_k).$$

- Bayes' Theorem: It holds that

$$\mathbb{P}[B|A] = \frac{\mathbb{P}[A|B]\mathbb{P}(B)}{\mathbb{P}(A)}.$$

1.3 Independence

Two events A_1 and A_2 are defined to be independent if and only if $\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1)\mathbb{P}(A_2)$. Multiple events $A_1, A_2, \ldots, A_m$ are mutually independent if and only if for every subset of events

$$\{A_{i_1}, \ldots, A_{i_n}\} \subseteq \{A_1, \ldots, A_m\},$$

the following holds:

$$\mathbb{P}(\cap_{k=1}^{n} A_{i_k}) = \Pi_{k=1}^{n} \mathbb{P}(A_{i_k}).$$

Multiple events $A_1, A_2, \ldots, A_m$ are pairwise independent if and only if every pair of events is independent, meaning $\mathbb{P}(A_n \cap A_k) = \mathbb{P}(A_n)\mathbb{P}(A_k)$ for all distinct pairs of indices n, k. Note that pairwise independence does not always imply mutual independence! Lastly, an important property is that if A and B are independent and $\mathbb{P}(B) > 0$, then $\mathbb{P}[A|B] = \mathbb{P}(A)$.

1.4 Random Variables

A random variable is a function $X(\omega) : \Omega \to \mathcal{B}$ that maps the sample space Ω to a subset of the real numbers $\mathcal{B} \subseteq \mathbb{R}$, with the property that the set $\{w : X(\omega) \in b\} = X^{-1}(b)$ is an event for every $b \in \mathcal{B}$. The distribution function (d.f.) of a random variable X is defined by

$$F_X(u) = \mathbb{P}(\omega : X(\omega) \leq u).$$

2 Stochastic Convergence

2.1 Convergence in Distribution

A sequence of random variables $X_1, X_2, \ldots$ converges in distribution to a random variable X if

$$\lim_{n\to\infty} F_{X_n}(u) = F_X(u),$$

for every point u at which $F_X(u)$ is continuous. This is denoted by $X_n \xrightarrow{d} X$. Note that $F_{X_n}(u)$ is the distribution function for X_n, and $F_X(u)$ is the distribution function for X.

2.2 Convergence in Probability

A sequence of random variables $X_1, X_2, \ldots$ converges in probability to a random variable X if for all $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| \geq \epsilon) = 0.$$

2.3 Relationships Between Modes of Convergence

There are several important points to note:

- Convergence in probability implies convergence in distribution.
- Convergence in distribution does not always imply convergence in probability.
- If X_n converges in distribution to a constant x_0, then X_n also converges in probability to x_0.