Probability Theory-I : SI 427

K. Suresh Kumar
Department of Mathematics
Indian Institute of Technology Bombay

July 29, 2024

Course Plan

0. **Foundations**

1. **Introduction.** Introduces random experiment, sample space, events, probability measure, σ-fields, probability space. [Hours 1-3]

2. **Random variables.** Definition and examples, Borel σ-field of subsets of $\mathbb{R}$, σ-field generated by a random variable. [Hours 4 - 7]

3. **Conditional probability and Independence.** Conditional probability, Law of total probability, Bayes theorem, Independence of events, Independence of σ-fields, Independence of random variables, $\liminf$ and *limsup* of events, Borel-Cantelli lemma. [Hours 8 - 12]

4. **Distributions.** Distribution function and its properties, Law of a random variable, Important distributions, Classification of random variables, pmf and pdf of random variables. [Hours 13 - 17]

5. **Random vectors, Joint distributions.** Borel σ-field of subsets of $\mathbb{R}^n$, Definition and a characterization of random vector, Law of random vector, Joint distribution function and its properties, Marginal distribution functions, Conditional pmf and conditional pdf. [Hours 18 - 23]

6. **Expectation and Conditional expectation.** Definition of expectation of discrete random variable and its properties, Simple random variables, Definition of expectation of nonnegative random variables, Definition of expectation of general random variables and properties, Expectation of continuous random variable with pdf, Monotone and Dominated convergence theorems (statements), Conditional expectation of discrete and continuous random variables. [Hours 24 - 30]

7. **Moment generating functions and Characteristic functions.** Definition and properties, Inversion theorem, Uniqueness theorem, Continuity theorem (statement). [Hours 31 - 34]

8. **Limit Theorems.** Markov and Chebyshev inequalities, Weak Law of large numbers, Strong Law of Large numbers, Central Limit Theorem, Applications. [Hours 35 -40]

Lecture 0 : Foundations

In this introduction, we quickly go through some basic mathematics used in this course. Following notations will be used without saying. $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ denote the set of all natural numbers, integers and rational numbers respectively.

Real number system is a set $\mathbb{R}$ (called the set of real numbers) containing $\mathbb{Q}$, together with two binary operations addition $+$ and multiplication $\cdot$ and an order relation $<$, i.e. $(\mathbb{R}, +, \cdot, <)$ satisfying the following.
I (Algebraic properties) Let $a, b, c \in \mathbb{R}$.

1 (associative property)

$$(a+b)+c \;=\; a+(b+c),\;\; (a\cdot b)\cdot c \;=\; a\cdot(b\cdot c).$$

2 (commutative property)

$$a+b \;=\; b+a,\;\; a\cdot b \;=\; b\cdot a.$$

3 (additive and multiplicative identity) There exists $0, 1 \in \mathbb{R}$ such that

$$a+0 \;=\; a \;=0+a,\;\; \forall\, a\in\mathbb{R},\;\; a\cdot 1 = a = \; 1\cdot a.$$

4 (additive and multiplicative inverses) For $a \in \mathbb{R}$, there exists $-a \in \mathbb{R}$ and $0 \neq a \in \mathbb{R}$, there exists $a^{-1} \in \mathbb{R}$ such that

$$a+(-a) \;=\; 0,\;\; a\cdot a^{-1} \;=1.$$

5 (distributive property)

$$a\cdot(b+c) \;=\; a\cdot b + a\cdot c.$$

II (Order properties) $\mathbb{R}$ contains a subset $\mathbb{R}^+$ (called the set of positive reals) such that

1 For each $a \in \mathbb{R}$, exactly one of the statements hold true.

$$a\in\mathbb{R}^+,\;\; a=0,\;\; -a\in\mathbb{R}^+.$$

2 If $a, b \in \mathbb{R}^+$, then $a+b, a\cdot b \in \mathbb{R}^+$.

Order properties defines an order $<$ on $\mathbb{R}$ given by, $a < b$ if $b - a \in \mathbb{R}^+$ and $a \leq b$ if $a < b$ or $a = b$.

III (Least upperbound property) If $A \subseteq \mathbb{R}$ has an upper bound[1], then it has a least upper bound.

From least upper bound property of $\mathbb{R}$, it follows that the set of natural numbers $\mathbb{N}$ is unbounded and hence we can deduce the Archimedean property.
(**Archimedean property**) Given any $x \in \mathbb{R}$, there exists a natural number n such that $n > x$.

On $\mathbb{R}$, we introduce a metric (i.e., a distance) using the 'modulus' function

$$|x| = \begin{cases} x & \text{if} \quad x > 0 \\ 0 & \text{if} \quad x = 0 \\ -x & \text{if} \quad x < 0, \end{cases}$$

i.e., distance between $x, y \in \mathbb{R}$ is given by $|x - y|$.

Using the metric given above, we give the definition of open and closed sets. Let B be a subset of $\mathbb{R}$, we say that $x \in \mathbb{R}$ is an **interior point** of B, if there exists an $\varepsilon > 0$ such that the interval $(x - \varepsilon, x + \varepsilon)$ is a subset of B. $x \in \mathbb{R}$ is said to be **boundary point** of B if for each $\varepsilon > 0$, the interval $(x - \varepsilon, x + \varepsilon)$ intersects (i.e. intersection is non empty) both B and its compliment.

A subset O of $\mathbb{R}$ is said to be **open** if all its elements are interior points. A subset F of $\mathbb{R}$ is said to be **closed** if its compliment is open. We end the discuss on the topology on $\mathbb{R}$ with following result.

Lemma 0.1 *Any open set O in $\mathbb{R}$ can be written as a union of open intervals with rational end points. In particular, any open set can be written as a countable union of open intervals.*

Basics from Set theory: We briefly describe naive theory of sets. Let U be a universal set, i.e. a set which contains all 'well defined" objects we may come across. For us it is going to the 'Sample space'. A is a subset of U, if every element of A is an element of U and $\mathcal{P}(U)$ denote the collection of all subsets of U.

[1] $M \in \mathbb{R}$ is said to be an upper bound for the set A if $a \leq M$ for all $a \in A$

Given a family of subsets $\{A_i | i \in I\}$ (indexed by I), $\cup_{i\in I} A_i$ and $\cap_{i\in I} A_i$ are defined as

$$\{x \in U | x \in A_i \text{ for some } i\},\ \{x \in U | x \in A_i \text{ for all } i\}$$

respectively. For a subset A of U, compliment of A (in U) denoted by A^c is defined as

$$A^c \ = \ \{x \in U | x \notin A\}.$$

We also define $A \backslash B = A \cap B^c$. We will now list, some of the useful properties of sets below.

1. For A, B, C subsets of U,

$$\text{(commutative property) } A \cup B \ = \ B \cup A,\ A \cap B \ = \ B \cap A,$$

$$\text{(associative property) } (A \cup B) \cup C \ = \ A \cup (B \cup C),\ (A \cap B) \cap C \ = \ A \cap (B \cap C).$$

$$(A^c)^c \ = \ A.$$

2. For subset of U, given by $A, B_i; i \in I$, I an index set,

$$A \cap (\cup_{i\in I} B_i) \ = \ \cup_{i\in I}(A \cap B_i),\ A \cup (\cap_{i\in I} B_i) \ = \ \cap_{i\in I}(A \cup B_i)\,.$$

3. For subsets of U given by $B_i, i \in I$,

$$\Big(\cup_{i\in I} B_i\Big)^c \ = \ \cap_{i\in I} B_i^c,\ \Big(\cap_{i\in I} B_i\Big)^c \ = \ \cup_{i\in I} B_i^c.$$

Exercise: Let $S \subseteq \mathbb{N}$ be such that $1 \in S$ and satisfies 'whenever $n \in S$ so is $n + 1$'. Then $S \ = \ \mathbb{N}$. *This exercise leads to principle of mathematical induction.*

Functions and its properties: Let A and B be two non empty sets, by $f : A \to B$, we mean a function from A to B. For $C \subseteq A$, $f(C)$ defined by

$$f(C) \ = \ \{f(a) | a \in C\}$$

is called the image set of C under f and for $D \subseteq B$, $f^{-1}(D)$ defined by

$$f^{-1}(D) \ = \ \{a \in A | f(a) \in D\}$$

is called the pre image of D under f.

A function $f : A \to B$ is said to be $1{-}1$ (injective) if $a_1, a_2 \in A$, with $f(a_1) = f(a_2) \Rightarrow a_1 = a_2$. A function $f : A \to B$ is said to be onto (surjective) if for each $b \in B$, there exists a $a \in A$ such that $f(a) = b$. A function $f : A \to B$ is said to be a bijection if it is both $1-1$ and onto.

Now we list some properties which are left as exercises.

Let $f : A \to B$ and $\{A_i, i \in I\}$ be a family of subsets of A and $\{B_i, i \in I\}$ be a family of subsets of B, then

1. $$f(\cup_{i\in I} A_i) = \cup_{i\in I} f(A_i).$$

2. $$f(\cap_{i\in I} A_i) \subseteq \cap_{i\in I} f(A_i),$$

   the inclusion is in general strict.

3. $$f^{-1}(\cup_{i\in I} B_i) = \cup_{i\in I} f^{-1}(B_i).$$

4. $$f^{-1}(\cap_{i\in I} B_i) = \cap_{i\in I} f^{-1}(B_i).$$

5. $$f^{-1}(B_1^c) = (f^{-1}(B_1))^c.$$

The following exercise is the key for the principle of induction.

Exercise: Let $S \subseteq \mathbb{N}$ be such that $1 \in S$ and when ever $n \in S$, then $n+1 \in S$. Show that $S = \mathbb{N}$.

Sequences and series: We consider real valued sequences and series. A sequence $\{a_n\}$ is said to converge to a if for each $\varepsilon > 0$, there exists a natural number N such that $|a_n - a| < \varepsilon$ for all $n \geq N$. In this case, we write $\lim_{n\to\infty} a_n = a$.

Now we list some useful results.

1. If $\{a_n\}$ is an increasing sequence (i.e., $a_n \leq a_{n+1}$ for all n), which is bounded from above, then $\lim_{n\to\infty} a_n = \sup_{n\geq 1} a_n$. An analogous result holds true for decreasing sequences.

2. (i) Let $\{a_n\}$ and $\{b_n\}$ be two convergent sequences such that $a_n \leq b_n$ for all n. Then
$$\lim_{n\to\infty} a_n \leq \lim_{n\to\infty} b_n.$$

(ii) Let $\{a_n\}, \{b_n\}$ and $\{c_n\}$ be three sequences such that $a_n \leq b_n \leq c_n$ for all n and $\lim_{n\to\infty} a_n = \lim_{n\to\infty} c_n = a$, then $\{b_n\}$ is convergent and $\lim_{n\to\infty} b_n = a$.

Given a real valued sequence $\{a_n\}$, we define

$$\begin{aligned}\limsup_{n\to\infty} a_n &= \lim_{n\to\infty} \sup\{a_n, a_{n+1}, \cdots\} \\ \liminf_{n\to\infty} a_n &= \lim_{n\to\infty} \inf\{a_n, a_{n+1}, \cdots\}.\end{aligned}$$

Exercise: Show that (i)

$$\liminf_{n\to\infty} a_n \leq \limsup_{n\to\infty} a_n.$$

(ii) If

$$\liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n < \infty,$$

then $\lim_{n\to\infty} a_n$ exists and is

$$\liminf_{n\to\infty} a_n.$$

A series $\sum_{n=1}^{\infty} a_n$ is said to converge to a number a, if the sequence of partial sums $s_n = \sum_{k=1}^{n} a_k$ converges to a. In this case, we write $\sum_{n=1}^{\infty} a_n = a$.

A series $\sum_{n=1}^{\infty} a_n$ is said to converge absolutely if the series $\sum_{n=1}^{\infty} |a_n|$ converges.

In general convergence doesn't imply absolute convergence. For example, the series $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$ converges but not absolutely. This is left it as an exercise.[2]

The series $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^{\alpha}}, 0 < \alpha < \infty$ is an important family of series. These are used to test the convergence/divergence of many other series as we have seen in Example **??** . So we will quickly take a look the convergence of the above series. Consider the series $\sum_{n=1}^{\infty} \frac{1}{n^2}$.

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{1}{n^2} &\leq 1 + \frac{1}{2^2} + \frac{1}{2^2} + \overbrace{\frac{1}{4^2} + \cdots + \frac{1}{4^2}}^{4 \text{ times}} + \overbrace{\frac{1}{8^2} + \cdots + \frac{1}{8^2}}^{8 \text{ times}} + \cdots \\ &= 1 + \frac{1}{2} + \frac{1}{2^2} + \cdots \text{(geometric series)}\end{aligned}$$

[2] Alternating series test (Leibnz test): If $a_n \downarrow 0$, then $\sum_{n=1}^{\infty} (-1)^n a_n$ converges.

Hence using comparison test, the series $\sum_{n=1}^{\infty}\frac{1}{n^2}$ converges.

Now let us consider

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{1}{n} &= 1+\frac{1}{2}+\overbrace{\frac{1}{3}+\frac{1}{4}}+\overbrace{\frac{1}{5}+\cdots+\frac{1}{8}}^{4\text{ terms}}+\overbrace{\frac{1}{9}+\cdots+\frac{1}{16}}^{8\text{ terms}}+\cdots\\ &\geq 1+\frac{1}{2}+\overbrace{\frac{1}{4}+\frac{1}{4}}+\overbrace{\frac{1}{8}+\cdots+\frac{1}{8}}^{4\text{ terms}}+\overbrace{\frac{1}{16}+\cdots+\frac{1}{16}}^{8\text{ terms}}+\cdots\\ &= \sum_{n=1}^{\infty}\frac{1}{2}.\end{aligned}$$

Hence by comparison test, the series $\sum_{n=1}^{\infty}\frac{1}{n}$ diverges.

Now we give a general convergence /divergence result for the above series.

Lemma 0.2 *The series* $\displaystyle\sum_{n=1}^{\infty}\frac{1}{n^\alpha}$ *converges if* $1<\alpha<\infty$ *and diverges to infinity if* $0<\alpha\leq 1$.

Proof: The proof I am going to give is based the comparison with the integrals $\int_1^\infty \frac{1}{x^\alpha}dx$. We know that the above Riemann integral diverges to infinity if $0<\alpha\leq 1$ and is $\frac{1}{\alpha-1}$ for $\alpha>1$.

Now using

$$\frac{1}{n^\alpha}\leq\int_{n-1}^{n}\frac{dx}{x^\alpha}, \alpha>1 \text{ and } \int_n^{n+1}\frac{dx}{x^\alpha}\leq\frac{1}{n^\alpha}, 0<\alpha\leq 1,$$

we get

$$\int_1^\infty\frac{dx}{x^\alpha}\leq\sum_{n=1}^{\infty}\frac{1}{n^\alpha}, 0<\alpha\leq 1$$

and

$$\sum_{n=2}^{\infty}\frac{1}{n^\alpha}\leq\int_1^\infty\frac{dx}{x^\alpha},\ \alpha>1.$$

From the above it follows that (exercise) series $\displaystyle\sum_{n=1}^{\infty}\frac{1}{n^\alpha}$ converges if $1<\alpha<\infty$ and diverges to infinity if $0<\alpha\leq 1$.

Real valued functions: Real valued functions $f: I\to\mathbb{R}$, where I is an interval ,is very often used. Recall that an interval I is said to be open if all

$x \in I$ are interior points.

We say that $f : I \to \mathbb{R}$, I an open interval, is continuous at $x_0 \in I$, if for each $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon.$$

$f : I \to \mathbb{R}$ is said to be continuous if it is continuous at all $x_0 \in \mathbb{R}$.

We will be using the following equivalent statement for a function to be continuous. $f : \mathbb{R} \to \mathbb{R}$ is continuous iff $f^{-1}(O)$ is open for all O open in $\mathbb{R}$.

Suppose f is continuous. Observing that any open set can be written as a countable union of finite open intervals, it is enough to show that $f^{-1}(I)$ is open for each finite open interval I.

For $y \in I$, choose $\varepsilon > 0$ such that $(y-\varepsilon,\ y+\varepsilon) \subseteq I$. Since f is continuous at $x \in f^{-1}(y)$, there exists $\delta > 0$ such that

$$x' \in (x - \delta,\ x + \delta) \ \Rightarrow\ f(x') \in (y - \varepsilon,\ y + \varepsilon) \ \Rightarrow\ x' \in f^{-1}\Big((y - \varepsilon,\ y + \varepsilon)\Big).$$

i.e.

$$(x - \delta,\ x + \delta) \subseteq f^{-1}\Big((y - \varepsilon,\ y + \varepsilon)\Big) \subseteq f^{-1}(I).$$

i.e. each $x \in f^{1}(I)$ is an interior point. Hence $f^{-1}(I)$ is open. This concludes the prove of first part. Converse is left as an exercise.

We say that $f : I \to \mathbb{R}$ is uniformly continuous if for each $\varepsilon > 0$, there exists a $\delta > 0$ such that for $x, y \in I$,

$$|x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon.$$

Exercise: Give an example of a continuous function which is not uniformly continuous.