EE325 Short Notes

Aditya Byju

Course Professors: Prof. Bikash Kumar Dey & Prof. D. Manjunath
Ref: Prof's video lectures
Couldn't complete ☹

Probability and Random Processes

September 2021

Contents

Set Theory

- **set** - a collection of well-defined objects
- **Russell's paradox** - the paradox defines the set S of all sets that are not members of themselves, but note that:
  - if S contains itself, then S must be a set that is not a member of itself by the definition of S, which is contradictory
  - if S does not contain itself, then S is one of the sets that is not a member of itself, and is thus contained in S by definition - also a contradiction

  this contradiction is called Russel's paradox
- **De-morgan's laws:**
  - $$\overline{A \cup B} = \overline{A} \cap \overline{B}$$
  - $$\overline{A \cap B} = \overline{A} \cup \overline{B}$$

  these laws apply to any number of sets
- $$A \subseteq B \leftrightarrow \overline{B} \subseteq \overline{A}$$
- **symmetric difference:**
$$\begin{aligned} A \bigtriangleup B &= (A - B) \cup (B - A) \\ &= (A \cup B) - (A \cap B) \end{aligned}$$
- **cartesian product:**
$$A \times B = \{ \ (a, b) \mid a \in A, \ \ b \in B \ \}$$
- A^∞ is the set of sequences of elements from A
- In general, $A^B =$ set of maps of B into A
- $|A|$ denotes the cardinality of a set
- Power set, $P(A) =$ set of subsets of A
$$|P(A)| = 2^{|A|}, \text{where } |A| \text{ is finite}$$
- A relation of A into B is a subset $R \subseteq A \times B$. If $(a, b) \in R$, then we write aRb.
- **equivalence relation** - a relation R of A into A is called an equivalence relation if:
  - It is reflexive i.e. $(a, a) \in R \ \forall \ a \in A$
  - It is symmetric i.e. $\forall a, b \in A$, if $(a, b) \in R$ then $(a, b) \in R$
  - It is transitive i.e. $\forall a, b, c \in A$, if $(a, b), \ (b, c) \in R$ then $(a, c) \in R$
- **equivalence class** - any equivalence relation partitions the set into a disjoint union of subsets, which are called equivalence classes, such that two elements are related iff they are in the same equivalence class
- Given any partition of A, i.e $P_i \subseteq A$; $i \in \mathbb{I}$ such that $\cup_{i \ \in \ \mathbb{I}} P_i = A \ \ \& \ \ P_i \cap P_j = \emptyset \ \ \forall i, \ j \in \mathbb{I}, \ i \neq j$, one can define an equivalence relation R $((a, b) \in R)$ iff $a, \ b \in P_i$ for some i
- A function or mapping or map f of A into B is a relation such that $\forall a \in A, \ \exists$ a unique $b \in B$ such that $(a, b) \in f$. Here b is called the image of a, and a is called the pre-image of b.

Cardinality

- **one-to-one (injective):** $f : A \to B$ is said to be injective if every element in the range R has a unique pre-image
- **onto (surjective):** $f : A \to B$ is said to be surjective if Range, $(R) = B$, i.e., every element in B has a pre-image in A
- **bijective:** $f : A \to B$ is said to be bijective if it is both injective and surjective
- **cardinality of a set:** is the number of elements in the set
- Comparing cardinality of two sets:
  - two sets A and B are said to be equicardinal if there exists a bijective function from A to B $\to$ $|A| = |B|$
  - set B has cardinality greater than or equal to set A if there exists a one-to-one function from A to B $\to$ $|B| \geq |A|$
  - set B has cardinality strictly greater than set A if there exists a one-to-one function from A to B, but no bijective function $\to$ $|B| > |A|$
- A set is said to be countably infinite if it is equicardinal with $\mathbb{N}$
- A set is said to be countable if it is finite or countably infinite
- Countable union of countable sets is countable
- A set is said to be uncountable if its cardinality is strictly greater than that of $\mathbb{N}$
- **Lemma:** the set of all infinite length binary strings $\{0,1\}^\infty$ is uncountable. It's proof is given by Cantor's diagonalization argument.
- The sets $[0,1]$, $\mathbb{R}$, $\mathbb{R}\backslash\mathbb{Q}$ are uncountable
- Dyadic rational is a rational number of the form $\frac{a}{2^b}$

Basics of probability

- **sample space (S or $\boldsymbol{\Omega}$)** - a set of outcomes of a random experiment
- If the sample space is finite or countably infinite, then we say that it is discrete
- **event** - a subset of a sample space
- For a sample space Ω, the set Ω is called sure event and the null set $(\emptyset)$ is called impossible event
- For discrete sample spaces we usually take $P(\Omega)$, the power set of Ω as the set of events
- If $A \cap B = \emptyset$ then A and B are called disjoint events
- Properties of events/subsets:
  - commutativity - $A \cup B = B \cup A$ & $AB = BA$
  - associativity - $(A \cup B) \cup C = A \cup (B \cup C)$ & $(A \cap B) \cap C = A \cap (B \cap C)$
  - distributivity - $(A \cup B)C = AC \cup BC$ & $AB \cup C = (A \cup C)(B \cup C)$
- Let the event space denoted by $\mathcal{F}$. Then, a probability measure on $(\Omega, \mathcal{F})$ is a function $P : \mathcal{F} \to [0,1]$ satisfying:

- $P(\emptyset) = 0,\ P(\Omega) = 1$
- If $A_1,\ A_2, \ldots$ is a collection of disjoint events, i.e., $A_i \cup A_j = \emptyset \quad \forall\ i \neq j$, then $P(\cup_{i\ =\ 1}^{\infty} A_i) = \sum_{i\ =\ 1}^{\infty} P(A_i)$

- The triplet $(\Omega,\ \mathcal{F},\ P)$ is called probability space
- Properties of probability measure P:
  - $P(\overline{A}) = 1 - P(A)$
  - If $B \supseteq A$, then $P(A) \geq P(B)$
  - $P(A \cup B) = P(A) + P(B) - P(A \cup B)$
  - $P(\cup_{i\ =\ 1}^{n} A_i) = \sum_{j\ =\ 1}^{n} (-1)^{j-1} (\sum_{s \subseteq [1:n],\ |s| = j} P(\cup_{i \in s} A_i))$
- **Lemma:**
  - let $A_1 \subseteq A_2 \subseteq A_3 \subseteq \ldots$ be an increasing sequence of events and let $A = \lim_{i \to \infty} A_i = \cup_{i=1}^{\infty} A_i$. Then $P(A) = \lim_{i \to \infty} P(A_i)$.
  - let $B_1 \supseteq B_2 \supseteq B_3 \supseteq \ldots$ be a decreasing sequence of events and let $B = \lim_{i \to \infty} B_i = \cap_{i=1}^{\infty} B_i$. Then $P(B) = \lim_{i \to \infty} P(B_i)$.
- **inclusion-exclusion principle (inclusion-exclusion bounds)** - If $E_1,\ E_2,\ E_3, \ldots,\ E_n$ are n events. Then:

$$P_r(\bigcup_{i\ =\ 1}^{n} E_i) = \sum_{i\ =\ 1}^{n} P_r(E_i) - \sum_{\{i,\ j\}\ \subseteq\ [1:n]} P_r(E_i E_j) + \sum_{\{i,\ j,\ k\}\ \subseteq\ [1:n]} P_r(E_i E_j E_k) - \ldots (-1)^{n-1} P_r(E_1 \ldots E_n)$$

- **union bound:**

$$P_r(\bigcup_{i\ =\ 1}^{n} E_i) \leq \sum_{i\ =\ 1}^{n} P_r(E_i)$$

- For random experiments that have equiprobable outcomes the probability of any event E is:

$$P_r(E) = \frac{|E|}{|\Omega|}$$

- Uniform probability measure means all outcomes are equally likely

Conditional probability

- Conditional probability is defined as:

$$P(A/B) = \frac{P(A \cap B)}{P(B)}, \quad if\ P(B) > 0$$

- For any event B, such that $P(B) > 0$, $P_B(A) = P(A/B) \quad \forall A \subseteq \Omega$. $P_B(A)$ is a valid probability measure, i.e., it satisfies all the properties of a probability measure.
- **Multiplication rule:**
  - $P(A \cap B) = P(B)\, P(A/B)$
  - $P(A \cap B \cap C) = P(B)\, P(A/B)\, P(C/A \cap B)$
  - in general $P(\cap_{i\ =\ 1}^{n} A_i) = P(A_1) \prod_{i\ =\ 2}^{n} P(A_i/A_1 \cap A_2 \cap \ldots \cap A_{i-1})$

- **Law of total probability:** let A be an event and $\{B_\text{i},\ \text{i} \in \mathbb{I}\}$ be countable collection of events that partition Ω $(P(B_i) > 0,\ \forall \text{i})$. Then:

$$P(A) = \sum_{\text{i} \in \mathbb{I}} P(B_\text{i}) P(A/B_\text{i})$$

- **Bayes' theorem:** let A be an event and $\{B_\text{i},\ \text{i} \in \mathbb{I}\}$ be countable collection of events that partition Ω $(P(B_i) > 0,\ \forall \text{i})$. Then:

$$P(B_\text{i}/A) = \frac{P(B_\text{i}) P(A/B_\text{i})}{P(A)}$$

- **independence** - two events A and B are said to be independent (under probability measure P) if $P(A \cap B) = P(A)P(B)$
- If A, B are independent events, and $P(B) > 0$ then $P(A/B) = P(A)$
- **Lemma:** if A & B are independent events then A & $\overline{B}$ are also independent events
- The events $A_1,\ A_2, \ldots,\ A_n$ are said to be independent if for all non-empty subsets $I \in \{1, 2, 3, \ldots, n\}$ we have:

$$P(\bigcap_I A_\text{i}) = \prod_I A_\text{i}$$

- Pairwise independence does not imply independence
- **conditional independence** - A & B are conditionally independent given C $(P(C) > 0)$ if $P((A \cap B)/C) = P(A/C)P(B/C)$
- Conditional independence does not imply independence and vice versa

Borel-Cantelli Lemma

- Let $\{A_n\}$ be a sequence of events over a sample space Ω, and a probability measure P. Then the event $A(i.o.) = \{A_n$ occurs for infinitely many $n\}$ is given by:

$$A(i.o.) = \cap_{n = 1}^{\infty} \cup_{m = n}^{\infty} A_m$$

- **First Borel-Cantelli lemma:** let $\{A_n\}$ be a sequence of events over a sample space Ω, and a probability measure P. If:

$$\sum_{n = 1}^{\infty} P(A_n) < \infty$$

  then $P(A(i.o.)) = 0$
- **Second Borel-Cantelli lemma:** let $\{A_n\}$ be a sequence of events over a sample space Ω, and a probability measure P. If:

$$\sum_{n = 1}^{\infty} P(A_n) = \infty$$

  then $P(A(i.o.)) = 1$
- **Stirling's formula:** $n! \sim \sqrt{2\pi n}(\frac{n}{e})^n$, when $n \to \infty$

Probability measures

- $\nexists$ a uniform probability measure $\mu : P(\Omega) \to [0,1]$ that satisfies the following two conditions:
  - for $0 \le a < b \le 1$, $\mu([a,b]) = b - a$
  - translational invariance: for $A \subseteq [0,1]$, and $\forall x \in [0,1]$, $\mu(A) = \mu(A \oplus x)$, where $A \oplus x = \{a + x \mid a \in A,\ a + x \le 1\} \cup \{a + x - 1 \mid a \in A,\ a + x > 1\}$

  Here, $P(\Omega)$ is called the Vitali set.
- Let $\mathcal{F}$ be a collection of subsets of Ω. $\mathcal{F}$ is said to be a $\sigma-$algebra of Ω if it satisfies:
  - if $A \in \mathcal{F}$ then $\overline{A} \in \mathcal{F}$
  - if $A_\text{i} \in \mathcal{F}$ ($\text{i} \ge 1$) is a countable sequence of sets, then $\cup_\text{i} A_i \in \mathcal{F}$
- A function $\mu : \mathcal{F} \to \mathbb{R} \cup \{-\infty, \infty\}$ is called a measure on $(\Omega, \mathcal{F})$ if it satisfies:
  - $\mu(A) \ge \mu(\emptyset) = 0 \quad \forall A \in \mathcal{F}$
  - if $A_\text{i} \in \mathcal{F}$ is countable sequence of disjoint sets, then $\mu(\cup_\text{i} A_\text{i}) = \sum_\text{i} \mu(A_\text{i})$

  If $\mu(\Omega) = 1$, then μ is called a probability measure
- **Theorem:** let μ be a measure on $(\Omega, \mathcal{F})$:
  - if $A \subseteq B$, then $\mu(A) \le \mu(B)$ [monotone]
  - if $A \subseteq \cup_{m\,=\,1}^{\infty} A_m$ than $\mu(A) \le \sum_{m\,=\,1}^{\infty} \mu(A_m)$ [subadditive]
  - if$A_i \uparrow A$, i.e., $A_1 \subset A_2 \subset \ldots$ and $\cup_\text{i} A_\text{i} = A$, then $\mu(A_\text{i}) \uparrow \mu(A)$ [continuity from below]
  - if$A_i \downarrow A$, i.e., $A_1 \supset A_2 \supset \ldots$ and $\cap_\text{i} A_\text{i} = A$, then $\mu(A_\text{i}) \downarrow \mu(A)$ [continuity from above]
- Probability mass function is a $p : \Omega \to [0,1]$ such that $\sum_{\omega\,\in\,\Omega} p(\omega) = 1$. Now $P(A) = \sum_{\omega\,\in\,A} p(\omega)$ is a probability measure.
- For any collection of $\sigma-$fields $\mathcal{F}_\text{i}\ :\ \text{i} \in \mathbb{I}$, $\cup_i \mathcal{F}_i$ is a $\sigma-$field
- In general, $\cap_i \mathcal{F}_i$ is not a $\sigma-$field
- For any $A \subset P(\Omega)$, $\sigma(A) = \cap_{A \subset \mathcal{F}} \mathcal{F}$ (where $\mathcal{F}$ is a $\sigma-$field) is the smallest $\sigma-$field containing A, or the $\sigma-$field generated by A
- **Borel $\sigma-$field and Borel sets:** For $\mathbb{R}$, the $\sigma-$field $\mathcal{R}$ generated by the open sets is called the Borel $\sigma-$field. The sets in it are the Borel sets. Similarly, for $\mathbb{R}^d$ the $\sigma-$field $\mathcal{R}^d$ generated by the open sets is called the Borel $\sigma-$field.
- **Stieltjes measure function (SMF):** let $F : \mathbb{R} \to \mathbb{R}$ be a function such that:
  - F is non-decreasing
  - F is right-continuous, i.e., $\lim_{y \downarrow x} F(y) = F(x)$
  - $F(\infty) = \lim_{x \to \infty} = 1$, and $F(-\infty) = 0$ (For probability measure)
- **Theorem:** for every Stieltjes measure function F there is a unique measure μ on $(\mathbb{R}, \mathcal{R})$ with $\mu(\ (a,b]\ ) = F(b) - F(a)$. If F satisfies the third property above, then μ is a probability measure.
- **Lebesgue measure:** this is the natural length measure on $\mathbb{R}$. This corresponds to $F(x) = x$.
- **Generalization of Stieltjes measures:** in $\mathbb{R}^d$, the three SMF conditions do not ensure a measure corresponding to the function. To ensure a measure we need an extra condition:
  - $\Delta_A F \ge 0$ for all rectangles A, where $\Delta_A F = \sum_{v \in V} sgn(v) F(v)$
- **Theorem:** suppose $F : \mathbb{R}^d \to [0,1]$ satisfies the four SMF properties, then $\exists$ a unique measure μ on $(\mathbb{R}^d,\ \mathcal{R}^d)$ so that $\mu(A) = \Delta_A F$ for all finite rectangles