Bedlam

Contents

1. Topology

Definition 1.1 (Topology):
Let X bet a set. A **topology over X** is a subset Σ of 2^X such that:
1. $A \subseteq \Sigma \Longrightarrow \bigcup_{E \in A} E$. Infinite or finite unions of sets.
2. $A, B \in \Sigma \Longrightarrow A \cap B \in \Sigma$. Finite intersections of sets.
3. $X \in \Sigma$

Definition 1.2 (Topological Space):
(X, Σ) is a **topological space** iff. Σ is a topology of X.

Definition 1.3 (Everywhere dense):
Let (X, Σ) topological space, and $H \subseteq X$. H is said **everywhere dense in Σ** iff. $\forall E \in \Sigma, E \neq \emptyset : H \cap E = \emptyset$. We can find a bit of H in every corner of the topology Σ.

Definition 1.4 (Separable):
Let (X, Σ) be a topological space. (X, Σ) is said **separable** iff $\exists H \subseteq X, H$ is countable : H is everywhere dense $\in \Sigma$. There is a sequence of elements $\{x_n \in X\}_{n=1}^{\infty}$ such that every set in the topology contains at least one element x_i.

Definition 1.5 (Metric Space):
(X, d) is a **metric space** iff.
1. $X \neq \emptyset$
2. $d : X \times X \longrightarrow \mathbb{R}_{\geq 0}$ such that (d is a distance):
   1. $\forall x, y \in X : d(x, y) = 0 \Longrightarrow x = y$. there are no different elements at zero-distance.
   3. $\forall x, y \in X : d(x, y) = d(y, x)$. symmetry.
   2. $\forall x, y, z \in X : d(x, z) \leq d(x, y) + d(y, z)$. triangular inequality.

Definition 1.6 (open ε-ball):
Let (X, d) be a metric space, $x \in X$, and $\varepsilon \in \mathbb{R}_{>0}$. We call $B_\varepsilon(x) = \{y \in X \mid d(x, y) < \varepsilon\}$ an **open ε-ball**. A ball of ε radius centered at some point.

Definition 1.7 (Neighborhood):
Let (X, d) be a metric space, $S \subseteq X$, $x \in S$, and $\varepsilon \in \mathbb{R}_{>0}$ such that the open ε-ball $B_\varepsilon(x) \subseteq S$. Then S is said a **neighborhood of x**. A neighborhood of an element is simply a set that contains an open ball containing the element.

Definition 1.8 (Open Set):
Let (X, d) be a metric space and $U \subseteq X$. U is an **open set** iff. $\forall u \in U : \exists \varepsilon \in \mathbb{R}_{>0} : B_\varepsilon(u) \subseteq U$. An open set is simply a set which is also neighborhood for all its points.

Definition 1.9 (Induced Topology):
Let (X, d) be a metric space. Σ is said an **induced topology** iff. $\Sigma = \{U \subseteq X \mid U \text{ is an open set in } (X, d)\}$

Definition 1.10 (Metrizable):
Let (X, Σ) be a topological space. (X, Σ) is said **metrizable** iff. $\exists (X, d)$ metric space : Σ is a topology induced by (X, d).

Definition 1.11 (Cauchy Sequence):
Let (X, d) be a metric space, $[x_n \in X]$ a sequence. $[x_n]$ is said a **cauchy sequence** iff. $\forall \varepsilon \in \mathbb{R}_{>0} : \exists N \in \mathbb{N} : \forall m, n \in \mathbb{N} : d(x_n, x_m) \leq \varepsilon$. There is a point after which all pairs of elements are close to each other.

Definition 1.12 (Convergent Sequence):
Let (X, d) be a metric space, $l \in X$, $[x_n \in X]$ a sequence. $[x_n]$ is said a **convergent sequence to the limit** l iff. $\forall \varepsilon \in \mathbb{R}_{>0} : \exists N \in \mathbb{R}_{>0} : \forall n > N : d(x_n, l) < \varepsilon$. If such a limit exists the sequence is simply said **convergent**.

Definition 1.13 (Complete Metric Space):
Let (X, d) be a metric space. (X, d) is said a **complete metric space** iff. every cauchy sequence is convergent.

Definition 1.14 (Polish Space):
Let (X, Σ) be a topological space. (X, Σ) is said a **Polish Space** iff. (X, Σ) is separable, metrizable, and a complete metric space for some metric.

2. Measure Theory

Definition 2.1 (σ-algebra):
Let X be a set. $\Sigma \subseteq 2^X$ is said a **sigma algebra of X** iff.:
1. $X \in \Sigma$
2. $E \in \Sigma \Longrightarrow X \setminus E \in \Sigma$. close under complement.
3. $\{A_n \in \Sigma\}_{n=1}^{\infty} \Longrightarrow \bigcup_{i=1}^{\infty} A_i \in \Sigma$. close under infinite unions.

Definition 2.2 (generate σ-algebra):
Let X be a set and $G \subseteq 2^X$. The **σ-algebra generated by G**, denoted $\sigma_X(G)$, is the smallest σ-algebra such that:
1. $G \subseteq \sigma_X(G)$.
2. $\forall \Sigma$ σ-algebra : $G \subseteq \Sigma \Longrightarrow \sigma_X(G) \subseteq \Sigma$. Every other σ-algebra that contains G contains also the generated one, $\sigma_X(G)$.

Definition 2.3 (σ-algebra product):
Let Σ_1 and Σ_2 be σ-algebras on X_1 and X_2 respectively. The **product σ-algebra** denoted $\Sigma_1 \otimes \Sigma_2$ is defined as $\sigma_{X_1 \times X_2}(\{S_1 \times S_2 \mid S_1 \in \Sigma_1, S_2 \in \Sigma_2\})$

Definition 2.4 (measurable space):
(X, Σ) is said **measurable** iff. Σ is a sigma-algebra of X.

Definition 2.5 (measure):
Given (X, Σ) measurable space. $\mu : \Sigma \longrightarrow \mathbb{R} \cup \{+\infty, -\infty\}$ is said a **measure** iff.
1. $E \in \Sigma \Longrightarrow \mu(E) \geq 0$. positive.
2. $\{E_n \in \Sigma\}_{n=1}^{\infty}$ such that $E_i \cap E_j$ for $i \neq j \Longrightarrow \mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu(E_i)$. The measure of disjoint sets is is the sum of the measures of each set.
3. $\exists E \in \Sigma : \mu(E) \in \mathbb{R}_{\geq 0}$. For at least an element μ is finite.

Definition 2.6 (measure space):
(X, Σ, μ) is said a **measure space** iff. (X, Σ) is a sigma algebra and μ is a measure of (X, Σ).

Definition 2.7 (measurable function):
Let (X_1, Σ_1) and (X_2, Σ_2) be a measurable spaces. $f X_1 \longrightarrow X_2$ is said a **measurable function** iff. $\forall E \in \Sigma_2 : f^{-1}(E) \in \Sigma_1$. The pre-image of each measurable set is again measurable.

Definition 2.8 (pushforward):
Let (X_1, Σ_1, μ) be a measure space. Let (X_2, Σ_2) be a measurable space. Let $f : X_1 \longrightarrow X_2$ be a measurable function. The **pushforwad of μ under f** is the mapping $f_\# \mu : \Sigma_2 \longrightarrow \mathbb{R}_{\geq 0}$ defined as:

$$\forall E \in \Sigma_2 : f_\# \mu(E) = \mu(f^{-1}(E))$$

Proposition 2.1 (pushforward of a measure is a measure):
Let (X_1, Σ_1, μ) be a measure space. Let (X_2, Σ_2) be a measurable space. Let $f : X_1 \longrightarrow X_2$ be a measurable function. Then $\left(X_2, \Sigma_2, f_\# \mu\right)$ is a measure space.

Proof 2.1 (of Proposition 2.1):
To prove that statement, we need to prove only the axioms of a measure.

1. Let $E \in \Sigma_2$, we need to show that $f_\# \mu(E) \geq 0$. This is trivial by definition of pushforward and measure.
2. Let $[E_n \in \Sigma_2]_{n=1}^{\infty}$ be a sequence of pairwise disjoint sets. We need to show that: $f_\# \mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} f_\# \mu(E_n)$.

$$\begin{aligned} f_\# \mu\left(\bigcup_{n=1}^{\infty} E_n\right) &= \mu\left(f^{-1}\left(\bigcup_{n=1}^{\infty} E_n\right)\right) \text{ definition of pushforward} \\ &= \mu\left(\bigcup_{n=1}^{\infty} f^{-1}(E_n)\right) \\ &= \sum_{n=1}^{\infty} \mu(f^{-1}(E_n)) \text{ definition of measure} \\ &= \sum_{n=1}^{\infty} f_\# \mu(E_n) \text{ definition of pushforward} \end{aligned}$$

3. We need to show that $\exists E \in \Sigma_1$ such that $f_\#(E) \geq 0$. Let $E' \in \Sigma_1$ such that $\mu(E') \geq 0$ (such E' exists by defintion of measure). Then, $f(E')$ is a set that meets the requirements, that is

$$f_\#(f(E')) = \mu(f^{-1}(f(E'))) = \mu(E') \geq 0$$

■

3. Probability Theory

Definition 3.1 (Probability Space):
(Ω, Σ, p) is said a **probability space** iff.
1. (Ω, Σ, p) is a measure space.
2. $p(\Omega) = 1$.

Intuitively, Ω represents the set of all possible outcomes, it is also known as **sample space**. Σ represents the set of all possible events. These are nothing more than set of outcomes. It is also known as **event space**. p is a measure on the event space, it is also known as **probability function**. It maps events to their likelihood.

Example 3.1 (Fair Die):
Consider the probability space (Ω, Σ, p), where:
1. $\Omega = \{1, 2, 3, 4, 5, 6\}$ is the sample space, representing the possible outcomes of rolling a standard six-sided die.
2. $\Sigma = 2^\Omega$ is the event space.
3. $p : \Sigma \longrightarrow [0, 1]$ is the probability measure function, defined as $P(E) = \frac{|E|}{6}$ for any event $E \in \Sigma$.

For example, consider the event $A = \{1, 2, 3\}$, which represents rolling a 1, 2, or 3. This event is an element of Σ. The probability of event A occurring is $p(A) = \frac{|A|}{6} = \frac{3}{6} = \frac{1}{2}$.

Definition 3.2 (Coupling):
Let $(\Omega_1, \Sigma_1, \mu_1)$ and $(\Omega_2, \Sigma_2, \mu_2)$ be probability spaces. A **coupling** is a probability space $(\Omega_1 \times \Omega_2, \Sigma_1 \otimes \Sigma_2, \gamma)$ such that:
1. $\forall E \in \Sigma_1 : \gamma(E \times \Omega_2) = \mu_1(E)$. The left marginal of γ is μ_1.
2. $\forall E \in \Sigma_2 : \gamma(\Omega_1 \times E) = \mu_2(E)$. The right marginal of γ is μ_2.

Example 3.2 (Coupling a Dice and a Coin):
Consider a probability space $\mathcal{F}_1 = \left(\Omega_1 = \{1, 2, 3, 4\}, \Sigma_1 = 2^{\Omega_1}, p_1 = A \mapsto \frac{|A|}{4}\right)$ (The probability space corresponding to a 4 sided die). Further, consider a probability space $\mathcal{F}_2 = \left(\Omega_2 = \{1, 2\}, \Sigma_2 = 2^{\Omega_2}, p_2 = A \mapsto \frac{|A|}{2}\right)$ (The probability space corresponding to a coin). We can define a probability space $\mathcal{F} = (\Omega_1 \times \Omega_2, \Sigma_1 \otimes \Sigma_2, p)$ by coupling $\mathcal{F}_1$ and $\mathcal{F}_2$. Here, sample space and event space are already decided, we need to provide only a proper measure p. Such a measure can be built by providing a coupling table:

$$\begin{pmatrix} p & \{1\} & \{2\} & \{3\} & \{4\} & p_1 \\ \{1\} & \frac{1}{4} & 0 & \frac{1}{4} & 0 & \frac{1}{2} \\ \{2\} & 0 & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{2} \\ p_2 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 1 \end{pmatrix}$$

On the top row, we have the possible singleton events from $\mathcal{F}_1$. On the left column, we have the possible singleton event from $\mathcal{F}_2$. The last row and column corresponds to marginal distributions. These marginals match p_2 and

p_1 as required by the definition of coupling. The central body of this matrix represents join probabilities of the die and coin. For example, $p(\{1\} \times \{3\}) = \frac{1}{4}$.

Note that we could fill this matrix in such a way that we have a probability space but not a coupling by breaking the marginal axioms.

Retrieving event probabilities from singleton events is only matter of applying traditional probability rules.

4. Wasserstein Distance