#### An Analytical Introduction to Probability Theory

## 2. Basics of Real Analysis - II

DASN, NTU

https://personal.ntu.edu.sg/wptay/

Let  $(\mathcal{X}, d)$  be a metric space.

# 2.1 Open and Closed Sets

**Definition 2.1.** The open ball of radius  $\epsilon > 0$  is defined by

$$B(x,\epsilon) = \{ y \in \mathcal{X} : d(x,y) < \epsilon \}.$$

**Definition 2.2.** A set U is open if  $\forall x \subset U$ ,  $\exists \epsilon > 0$  such that  $B(x, \epsilon) \subset U$ . A set F is closed if  $F^c = \mathcal{X} \setminus F$  is open.

A set can be both open and closed, e.g.,  $\mathcal{X}, \emptyset$ .

**Example 2.1.** Let  $\mathcal{X} = \{x_1, x_2, \ldots\}$  be a discrete space. Consider the discrete metric

$$d(x,y) = \begin{cases} 0, & \text{if } x = y, \\ 1, & \text{if } x \neq y. \end{cases}$$

For any  $A \subset \mathcal{X}$  and  $\epsilon \in (0,1)$ ,  $B(a,\epsilon) = \{a\} \subset A$ . Therefore A is open. This also proves that A is closed as  $A^c$  is open.

If  $U_i$ ,  $i \ge 1$  are open, then  $\bigcup_{i=1}^{\infty} U_i$  is open,  $\bigcap_{i=1}^{n} U_i$  is open but  $\bigcap_{i=1}^{\infty} U_i$  may not be open.

**Example 2.2.** Suppose that  $U_i = (\frac{1}{2} - \frac{1}{i}, \frac{1}{2} + \frac{1}{i})$ , then  $U_i$  is open while  $\bigcap_{i=1}^{\infty} U_i = \{1/2\}$  is closed.

If  $F_i$  is closed, then  $\bigcup_{i=1}^{\infty} F_i$  is closed and  $\bigcap_{i=1}^{\infty} F_i$  is closed.

**Definition 2.3.** x is a limit point of A if  $\forall \epsilon > 0$ ,  $\exists y \in B(x, \epsilon) \cap A$ , and  $y \neq x$ .

Therefore, x is a limit point of A if  $\exists y_1, y_2, \ldots \in A, y_i \neq x$  for all  $i \geq 1$ , s.t.  $y_i \rightarrow x$ .

Lemma 2.1. A is closed if and only if all limit points of A are in A.

*Proof.* Suppose that A is closed, so that  $A^c$  is open. Suppose there exists a limit point x of A s.t.  $x \in A^c$ . Then  $\exists \epsilon > 0$ , s.t.  $B(x, \epsilon) \subset A^c$ , which is a contradiction to x being a limit point of A.

Suppose that all limit points of A belongs to A. Consider  $x \in A^c$ . There exists  $\epsilon > 0$  s.t.  $B(x,\epsilon) \subset A^c$  because otherwise there exists  $(y_i) \subset A$  s.t.  $y_i \to x$ , which means that x is a limit point of A leading to a contradiction. This shows that  $A^c$  is open and hence A is closed.

## 2.2 Compact Spaces

**Definition 2.4.**  $\mathcal{X}$  is sequentially compact if every sequence in  $\mathcal{X}$  has a convergent subsequence in  $\mathcal{X}$ .

**Definition 2.5.**  $\mathcal{X}$  is totally bounded if  $\forall \epsilon > 0$ , there exists a finite collection  $\{B(x_i, \epsilon) : i = 1, 2, ..., N_{\epsilon}\}$  such that

$$\mathcal{X} \subset \bigcup_{i=1}^{N_{\epsilon}} B(x_i, \epsilon).$$

Note that if a set  $\mathcal{X}$  is totally bounded, then it is bounded. The converse is not true: consider the discrete space in Example 2.1, it is bounded but not totally bounded if it is infinite.

**Theorem 2.1.** For a metric space  $(\mathcal{X}, d)$ , the following are equivalent:

- (i)  $\mathcal{X}$  is sequentially compact.
- (ii)  $\mathcal{X}$  is complete and totally bounded.
- (iii) Every open cover of  $\mathcal{X}$  has a finite subcover. We say that  $\mathcal{X}$  is compact.

Proof.

1) (i)  $\Leftrightarrow$  (ii):

We first show that (i)  $\Rightarrow$  (ii). Since  $\mathcal{X}$  is sequentially compact, every Cauchy sequence in  $\mathcal{X}$  has a convergent subsequence. From Lemma 1.11, the Cauchy sequence also converges in  $\mathcal{X}$ , so  $\mathcal{X}$  is complete. Suppose that  $\mathcal{X}$  is not totally bounded. Then  $\exists \epsilon > 0$  so that  $\mathcal{X}$  cannot be covered by a finite collection of open balls. Choose any  $x_1 \in \mathcal{X}$ . Then  $\exists x_2 \notin B(x_1, \epsilon)$ . Similarly,  $\exists x_3 \notin B(x_1, \epsilon) \cup B(x_2, \epsilon)$ , and so on. The sequence  $(x_n)$  does not contain any convergent subsequence since  $d(x_i, x_j) \geq \epsilon$  for any  $i \neq j$ . This is a contraction to (i).

We next show that (ii)  $\Rightarrow$  (i). Since  $\mathcal{X}$  is totally bounded, for each  $m \geq 1$ , there exists a finite cover  $\{B(x_{m,k}, 1/m) : k = 1, ..., M_m\}$  of  $\mathcal{X}$ . Consider any infinite sequence  $(y_n)$  in  $\mathcal{X}$ . We assume that  $y_n$  are distinct because if there are infinitely many  $y_n$  that are the same, then there is a trivial convergent subsequence. Then there is a  $B(x_{1,k_1}, 1)$  that contains a subsequence  $(y_{1,n})$  of  $(y_n)$ . Similarly, there is a  $B(x_{2,k_2}, 1/2)$  that contains a further subsequence  $(y_{2,n})$  of  $(y_{1,n})$ , and so on. Consider the "diagonal" subsequence  $(y_{m,m})_{m=1}^{\infty}$ . This sequence is Cauchy and since  $\mathcal{X}$  is complete, it converges.

2) (i)  $\Leftrightarrow$  (iii): We show (iii)  $\Rightarrow$  (i). To do that, we first prove the following facts:

- (a) Any compact  $A \subset \mathcal{X}$  is closed. In particular, if  $\mathcal{X}$  is compact, then it is closed. Let  $x \in A^c$  and  $U_n = \{y : d(y,x) > 1/n\}$  for  $n \geq 1$ . Every  $y \in \mathcal{X}$  with  $y \neq x$  has d(y,x) > 0 so y belongs to some  $U_n$ . Therefore,  $\{U_n : n \geq 1\}$  covers A and there must be a finite subcover. Let N to be the largest index in the subcover, i.e., every  $y \in A$  lies in some  $U_n$  where  $n \leq N$ . Then  $B(x,1/N) \subset A^c$  and A is closed.
- (b) If  $\mathcal{X}$  is compact and  $A \subset \mathcal{X}$  is closed, then A is compact. Let  $\{U_n\}$  be an open cover of A. Then  $\{U_n\} \cup \{A^c\}$  is an open cover of  $\mathcal{X}$ . There is a finite subcover, say,  $\{U_1, \ldots, U_N, A^c\}$  of  $\mathcal{X}$ . Then  $\{U_1, \ldots, U_N\}$  is a finite open cover of A.

Suppose  $\mathcal{X}$  is compact. Assume there is a sequence  $(x_n)$  that has no convergent subsequences. In particular, this sequence has infinitely distinct points  $y_1, y_2, \ldots$  Since there is no convergent subsequence, there is some open ball  $B_k$  containing each  $y_k$  and no other  $y_i$ . The set  $A = \{y_1, y_2, \ldots\}$  is closed as it has no limit points, so it is compact. But  $\{B_k\}$  is an open cover of A and has no finite subcover, a contradiction. Therefore  $(x_n)$  has a convergent subsequence whose limit lies in  $\mathcal{X}$  as  $\mathcal{X}$  is closed.

We now show (i)  $\Rightarrow$  (iii). Suppose that  $\mathcal{X}$  is sequentially compact. Let  $\{G_{\alpha} \subset \mathcal{X} : \alpha \in I\}$  be an open cover of  $\mathcal{X}$ . We claim that there exists  $\epsilon > 0$  such that every ball  $B(x, \epsilon)$  is contained in some  $G_{\alpha}$ . Suppose not. Then for each positive integer  $n, \exists y_n \in \mathcal{X}$  such that  $B(y_n, 1/n)$  is not contained in any  $G_{\alpha}$ . By hypothesis, there exists a subsequence  $y_{n_i} \to y \in \mathcal{X}$ . Since  $\{G_{\alpha} \subset \mathcal{X} : \alpha \in I\}$  is an open cover of  $\mathcal{X}$ , there exists  $G_{\alpha_0} \ni y$  and  $\epsilon > 0$  such that  $B(y, \epsilon) \subset G_{\alpha_0}$ . Choose  $n_i$  sufficiently large so that  $d(y_{n_i}, y) < \epsilon/2$  and  $1/n_i < \epsilon/2$ . Then  $B(y_{n_i}, 1/n_i) \subset G_{\alpha_0}$ , a contradiction.

Since  $\mathcal{X}$  is sequentially compact, it is totally bounded, so there exists a finite collection of balls of radius  $\epsilon$  s.t.  $\{B(x_i, \epsilon) : i = 1, 2, ..., N\}$  covers  $\mathcal{X}$ . Choose  $\alpha_i \in I$  s.t.  $B(x_i, \epsilon) \subset G_{\alpha_i}$ . Then  $\{G_{\alpha_i} : i = 1, 2, ..., N\}$  is a finite subcover of  $\mathcal{X}$ , which means  $\mathcal{X}$  is compact.

**Theorem 2.2** (Heine-Borel).  $A \subset \mathbb{R}^k$  is compact iff A is closed and bounded.

*Proof.* Since the following proof can be repeated in every dimension, it suffices to prove only for  $A \subset \mathbb{R}$ .

We first prove sufficiency. Suppose A is closed and bounded and  $(x_n) \subset A$ . Then  $x_n$  is bounded since A is bounded. By Lemma 1.6, there exists convergent  $(x_{n_i})$  such that  $x_{n_i} \to x \in \mathbb{R}$ . Since A is closed,  $x \in A$ . Therefore, A is compact.

Next, suppose that A is compact. Then from Theorem 2.1, A is complete and totally bounded, hence closed and bounded in  $\mathbb{R}$ .

## 2.3 Continuity

**Definition 2.6.** A function  $f:(\mathcal{X}, d_{\mathcal{X}}) \mapsto (\mathcal{Y}, d_{\mathcal{Y}})$  is (pointwise) continuous at  $x \in \mathcal{X}$  if  $\forall \epsilon > 0, \exists \delta_x > 0$  s.t.  $d_{\mathcal{Y}}(f(x), f(z)) < \epsilon$ , whenever  $d_{\mathcal{X}}(x, z) < \delta_x$ .

Note that the definition is equivalent to saying that  $\forall \epsilon > 0$ ,  $\exists \delta_x > 0$  s.t.  $B(x, \delta_x) \subset f^{-1}(B(f(x), \epsilon))$ .

**Lemma 2.2.**  $f:(\mathcal{X},d_X)\mapsto(\mathcal{Y},d_Y)$  is continuous iff  $f^{-1}(U)$  is open in  $\mathcal{X}$  for every open  $U\subset\mathcal{Y}$ .

*Proof.* Suppose f is continuous and U is open in  $\mathcal{Y}$ . For each  $y \in U$ , there exists an open ball  $B(y,r) \subset U$ . Since f is continuous, for each  $x \in f^{-1}(\{y\})$ ,  $\exists \delta_x > 0$  s.t.  $B(x,\delta_x) \subset f^{-1}(B(y,r)) \subset f^{-1}(U)$ . Therefore,  $f^{-1}(U) = \bigcup_{x \in f^{-1}(U)} B(x,\delta_x)$  is open.

To prove the other direction, we have for every  $x \in \mathcal{X}$  and  $\epsilon > 0$ ,  $f^{-1}(B(f(x), \epsilon))$  is open and contains x, so it contains an open ball around x. Thus, f is continuous at x.

**Lemma 2.3.** If f is continuous and  $B \in \mathcal{X}$  is compact, then  $f(B) \triangleq \{f(x) : x \in B\}$  is compact.

*Proof.* Consider  $y_n = f(x_n)$ .  $\exists (x_{n_i}) \text{ s.t. } x_{n_i} \to x \in B$ . From the continuity of  $f, y_{n_i} = f(x_{n_i}) \to f(x) \in f(B)$ .

From Theorem 2.2, f(B) is closed and bounded. Therefore,  $\sup_{x \in B} f(x)$  and  $\inf_{x \in B} f(x)$  are both achieved on B.

**Definition 2.7.** f is uniformly continuous if  $\forall \epsilon > 0$ ,  $\exists \delta > 0$  s.t.  $d_{\mathcal{Y}}(f(x), f(z)) < \epsilon$ ,  $\forall d_{\mathcal{X}}(x, z) < \delta$ .

#### Example 2.3.

- (a) If |f(x) f(y)| < Ld(x, y) for some positive constant L and all  $x, y \in \mathcal{X}$ , we say that f is Lipschitz continuous. It is clear that f is uniformly continuous on  $\mathcal{X}$ .
- (b) If  $f : \mathbb{R} \to \mathbb{R}$  is differentiable with  $\sup |f'| < \infty$ , where f' is the derivative of f, then f is uniformly continuous from the mean value theorem.

**Lemma 2.4.** Suppose that  $f: K \mapsto \mathcal{Y}$ , where  $K \subset \mathcal{X}$  is compact. If f is continuous, then f is uniformly continuous.

Proof. Fix  $\epsilon > 0$ . For each  $x \in \mathcal{X}$ ,  $\exists \delta_x > 0$ , s.t.  $d_{\mathcal{Y}}(f(x), f(y)) < \epsilon/2$  whenever  $d_{\mathcal{X}}(x, y) < \delta_x$ . We have  $\{B(x, \delta_x/2) : x \in K\}$  is an open cover of K. Since K is compact, exists a finite subcover  $\{B(x_i, \delta_{x_i}/2) : 1 \le i \le n\}$ . Let  $\delta = \min\{\delta_{x_i}, \dots, \delta_{x_n}\} > 0$ .

For each  $x \in K$ ,  $\exists x_i$ , s.t.  $d_{\mathcal{X}}(x, x_i) < \delta_{x_i}/2$ . Then for all y s.t.  $d_{\mathcal{X}}(x, y) < \delta/2$ , we have

$$d_{\mathcal{X}}(y, x_i) \le d_{\mathcal{X}}(x, y) + d_{\mathcal{X}}(x, x_i) < \delta_{x_i}.$$

Therefore, we obtain

$$d_{\mathcal{Y}}(f(x), f(y)) \le d_{\mathcal{Y}}(f(x), f(x_i)) + d_{\mathcal{Y}}(f(x_i), f(y)) \le \epsilon,$$

which shows that f is uniformly continuous.

Note that in the above proof, y need not be in K. We obtain a slightly stronger result here.

**Corollary 2.1.** Suppose  $f: \mathcal{X} \mapsto \mathcal{Y}$  is continuous, and  $K \subset \mathcal{X}$  is compact. Then  $\forall \epsilon > 0$ ,  $\exists \delta > 0$  such that  $d_{\mathcal{Y}}(f(x), f(y)) < \epsilon$  whenever  $x \in K$ ,  $y \in \mathcal{X}$  and  $d_{\mathcal{X}}(x, y) < \delta$ .

# 2.4 Riemann Integral

We consider a function  $f:[a,b] \to \mathbb{R}$  in this section. A partition  $P=(x_0,\ldots,x_N)$  is defined by  $x_0=a < x_1 < x_2 < \cdots < x_N=b$ . We say that Q is a refinement of P if  $P \subset Q$ . Let  $\mathcal{P}$  be the collection of all partitions. For  $P \in \mathcal{P}$ , define

$$U(f, P) = \sum_{i=1}^{N} \sup_{[x_{i-1}, x_i]} f \cdot (x_i - x_{i-1}),$$

$$L(f, P) = \sum_{i=1}^{N} \inf_{[x_{i-1}, x_i]} f \cdot (x_i - x_{i-1}).$$

Lemma 2.5.  $L(f, P) \leq U(f, Q), \ \forall P, Q \in \mathcal{P}$ 

Proof.

$$\begin{split} L(f,P) &\leq L(f,P \cup Q) \quad \text{since } \inf_{I_1} f \leq \inf_{I_2} f \text{ if } I_1 \supset I_2 \\ &\leq U(f,P \cup Q) \\ &\leq U(f,Q) \quad \text{since } \sup_{I_1} f \geq \sup_{I_2} f \text{ if } I_1 \supset I_2. \end{split}$$

From the above lemma, we have

$$\sup_{P \in \mathcal{P}} L(f, P) \le \inf_{P \in \mathcal{P}} U(f, P) \tag{1}$$

**Definition 2.8.**  $f:[a,b] \mapsto \mathbb{R}$  is Riemann integrable if equality in (1) holds, i.e.,

$$\forall \epsilon > 0, \ \exists P \in \mathcal{P} \text{ s.t. } U(f, P) - L(f, P) < \epsilon.$$

**Example 2.4.** The following function is not Riemann integrable as L(f, P) = 0 and U(f, P) = 1 for all  $P \in \mathcal{P}$ :

$$f(x) = \begin{cases} 1, & \text{if } x \in \mathbb{Q} \cap [0, 1], \\ 0, & \text{otherwise.} \end{cases}$$

**Definition 2.9.** A set  $A \subset \mathbb{R}$  has Lebesgue measure zero if  $\forall \epsilon > 0$ , there exists open intervals  $(\alpha_1, \beta_1), (\alpha_2, \beta_2), \ldots s.t.$ 

$$A \subset \bigcup_{i=1}^{\infty} (\alpha_i, \beta_i) \text{ and } \sum_{i=1}^{\infty} (\alpha_i - \beta_i) < \epsilon.$$

If a countable sequence of sets  $A_1, A_2, \ldots$  each of which has Lebesgue measure zero, then the union  $\bigcup_{i=1}^{\infty} A_i$  has Lebesgue measure zero. To see this, let  $\epsilon > 0$  and  $A_j$  be covered by  $\bigcup_i (\alpha_{ij}, \beta_{ij})$  with  $\sum_i (\alpha_{ij} - \beta_{ij}) < \epsilon/2^j$ .

**Theorem 2.3** (Henri Lebesgue). Suppose  $f : [a,b] \mapsto \mathbb{R}$  is bounded. Then f is Riemann integrable iff  $\exists A \subset [a,b]$  of Lebesgue measure zero s.t. f is continuous on  $[a,b] \setminus A$ .

*Proof.* We first show that if f is Riemann integrable, then its set of discontinuities has Lebesgue measure zero. Observe that  $y \in (a, b)$  is a point of discontinuity of f iff  $\exists j \in \mathbb{Z}_+$  s.t.  $\sup_I f - \inf_I f \geq 1/j$  for all open intervals  $I \subset (a, b)$  containing y. Let

$$S_j = \left\{ y \in (a, b) : \sup_I f - \inf_I f \ge \frac{1}{j} \ \forall \text{ open intervals } I \subset (a, b) \text{ with } y \in I \right\}.$$

Then, the set of discontinuities of f in (a,b) is  $\bigcup_{j=1}^{\infty} S_j$ . For  $\epsilon > 0$ , since f is Riemann integrable, there exists some partition  $P = (x_0, \ldots, x_N)$  s.t.

$$U(f,P) - L(f,P) = \sum_{i=1}^{N} (\sup_{[x_{i-1},x_i]} f - \inf_{[x_{i-1},x_i]} f) \cdot (x_i - x_{i-1}) < \frac{\epsilon}{j}.$$
 (2)

Let  $B = \{i : (x_{i-1}, x_i) \cap S_j \neq \emptyset\}$ . Then from (2), we have

$$\sum_{B} \left( \sup_{[x_{i-1}, x_i]} f - \inf_{[x_{i-1}, x_i]} f \right) \cdot (x_i - x_{i-1}) + \sum_{B^c} \left( \sup_{[x_{i-1}, x_i]} f - \inf_{[x_{i-1}, x_i]} f \right) \cdot (x_i - x_{i-1}) < \frac{\epsilon}{j}$$
 (3)

Since

$$\sum_{B^c} \left( \sup_{[x_{i-1}, x_i]} f - \inf_{[x_{i-1}, x_i]} f \right) \cdot (x_i - x_{i-1}) \ge 0$$

and

$$\sum_{B} \left( \sup_{[x_{i-1}, x_i]} f - \inf_{[x_{i-1}, x_i]} f \right) \cdot (x_i - x_{i-1}) \ge \frac{1}{j} \sum_{i \in B} (x_i - x_{i-1}),$$

 $<sup>^{1}</sup>f$  is also said to be continuous almost everywhere on [a, b].

we obtain from (3),

$$\sum_{i \in B} (x_i - x_{i-1}) < \epsilon.$$

We have

$$S_j \subset \bigcup_{i \in B} (x_{i-1}, x_i) \bigcup \{x_0, x_1, \cdots, x_N\},$$

therefore  $S_i$  has Lebesgue measure zero.

We next prove the converse. Fix an  $\epsilon > 0$ . Assume that we have A with Lebesgue measure zero, which means that there is a cover  $\bigcup_{j=1}^{\infty} (\alpha_j, \beta_j) \supset A$  s.t.  $\sum_{j=1}^{\infty} (\beta_j - \alpha_j) < \epsilon$ . Let  $K = [a, b] \setminus \bigcup_{j \geq 1} (\alpha_j, \beta_j)$ , which is closed and bounded and therefore compact by Theorem 2.2. Since f is continuous, from Corollary 2.1,  $\exists \delta > 0$ , s.t.  $|f(x) - f(y)| < \epsilon$  whenever  $x \in K$ ,  $y \in [a, b]$  and  $|x - y| < \delta$ .

We choose a partition P with  $a = x_0 < x_1 < x_2 < \dots < x_N = b$  s.t.  $\max_{1 \le i \le N} (x_i - x_{i-1}) < \delta$ . If  $[x_{i-1}, x_i] \cap K = \emptyset$ , then  $[x_{i-1}, x_i] \subset \bigcup_j (\alpha_j, \beta_j)$  and

$$\sum_{i:[x_{i-1},x_i]\cap K=\emptyset} \left(\sup_{[x_{i-1},x_i]} f - \inf_{[x_{i-1},x_i]} f\right) \cdot (x_i - x_{i-1}) \le \left(\sup_{[a,b]} f - \inf_{[a,b]} f\right) \cdot \sum_{j} (\beta_j - \alpha_j) < M\epsilon,$$

where  $M = \sup_{[a,b]} f - \inf_{[a,b]} f < \infty$ . Suppose  $[x_{i-1}, x_i] \cap K \neq \emptyset$ . Then for any  $y, z \in [x_{i-1}, x_i]$  and  $y_i \in [x_{i-1}, x_i] \cap K$ , we have

$$|f(y) - f(z)| \le |f(y) - f(y_i)| + |f(y_i) - f(z)|$$

$$< \epsilon + \epsilon$$

$$= 2\epsilon.$$

where the last inequality follows because  $|y - y_i|$ ,  $|z - y_i| < \delta$ . Therefore,

$$\sum_{i:[x_{i-1},x_i]\cap K\neq\emptyset} (\sup_{[x_{i-1},x_i]} f - \inf_{[x_{i-1},x_i]} f) \cdot (x_i - x_{i-1}) \le 2\epsilon(b-a).$$

We finally obtain

$$U(f, P) - L(f, P) \le M\epsilon + 2\epsilon(b - a) = (M + 2(b - a))\epsilon,$$

and the proof is complete.

From Theorem 2.3, we see that only a very limited class of functions f is Riemann integrable. This is not sufficient to model many practical applications. Therefore, Henri Lebesgue, a French mathematician in the 17th century, embarked on a program to introduce a much more versatile integral known as the Lebesgue integral. We will introduce this in the coming week as part of the theory of probability.