Preliminaries for Empirical Processes

Yilun Li

June 9, 2021

Table of Contents

Introduction

- Reference: Chapter 6, Introduction to Empirical Processes and Semiparametric Inference (Kosorok)
- This chapter presents mathematical and statistical concepts and basic ideas of empirical process, and provides a foundation for technical development in later chapters.
- Topics covered: metric space, outer expectation, linear operator and differentiation.

Introduction

- Metric space provides the descriptive language by which the most important results in stochastic processes are derived and expressed.
- Outer expectation helps to define and utilize outer modes of convergence for non-measurable quantities.
- Linear operators and differentiation are important in empirical process methods for functional delta method and Z-estimator theory.

Metric Space

Definition

A collection $\mathcal{O}$ of subsets of a set $\mathbf{X}$ is a topology in $\mathbf{X}$ if

1. $\emptyset \in \mathcal{O}$ and $\mathbf{X} \in \mathcal{O}$;
2. If $U_j \in \mathcal{O}$ for j=1, ..., m, then $\bigcap_{j=1}^{m} U_j \in \mathcal{O}$;
3. For an arbitrary collection $\{U_\alpha\} \subseteq \mathcal{O}$, we have $\bigcup_\alpha U_\alpha \in \mathcal{O}$.

$(\mathbf{X}, \mathcal{O})$ is called a topological space, and members of $\mathcal{O}$ are called the open sets in $\mathbf{X}$.

Several relevant concepts:

- (continuous map) A map $\mathbf{f}$: $\mathbf{X} \to \mathbf{Y}$ between topological spaces is continuous if $\mathbf{f}^{-1}(\mathbf{U})$ is open in $\mathbf{X}$ whenever $\mathbf{U}$ is open in $\mathbf{Y}$.
- (closed set) A set $\mathbf{B}$ in $\mathbf{X}$ is closed if and only if its complement in $\mathbf{X}$ is open.
- (closure) The closure of an arbitrary set $\mathbf{E} \subseteq \mathbf{X}$ is the smallest closed set containing $\mathbf{E}$, denoted by $\bar{\mathbf{E}}$.
- (interior) The interior of an arbitrary set $\mathbf{E} \subseteq \mathbf{X}$ is the largest open set contained in $\mathbf{E}$, denoted by $\mathbf{E}^{\circ}$.
- (dense set) A subset $\mathbf{A}$ of a topological space $\mathbf{X}$ is dense if $\bar{\mathbf{A}} = \mathbf{X}$.
- (separable space) A topological space $\mathbf{X}$ is separable if it has a countable dense subset.

Several relevant concepts (continued):

- (neighborhood) A neighborhood of a point $\mathbf{x} \in \mathbf{X}$ is any open set that contains $\mathbf{x}$.
- (Hausdorff space) A topological space $\mathbf{X}$ is Hausdorff if distinct points have disjoint neighborhoods.
- (convergence) Say a sequence of points $\{\mathbf{x}_n\}$ in a topological space $\mathbf{X}$ converges to $\mathbf{x} \in \mathbf{X}$, if every neighborhood of $\mathbf{x}$ contains all but finitely many of the $\mathbf{x}_n$'s, denoted by $\mathbf{x}_n \to \mathbf{x}$.
- (compactness) A subset $\mathbf{K}$ of a topological space is compact if for every covering $\bigcup_{\alpha \in \mathcal{I}} \mathbf{U}_\alpha \supseteq \mathbf{K}$ (where $\mathcal{I}$ is the index set and $\mathbf{U}_\alpha$ are open sets), there exists a finite subset $\mathcal{I}_0 \subseteq \mathcal{I}$ such that $\bigcup_{\alpha \in \mathcal{I}_0} \mathbf{U}_\alpha \supseteq \mathbf{K}$.
- (σ-compactness) A σ-compact set is a countable union of compact sets.

Propositions:

1. For a Hausdorff topological space $\mathbf{X}$ and a sequence $\{\mathbf{x}_n\} \subseteq \mathbf{X}$, if $\mathbf{x}_n \to \mathbf{x} \in \mathbf{X}$ and $\mathbf{x}_n \to \mathbf{y} \in \mathbf{X}$, then $\mathbf{x} = \mathbf{y}$.
2. If $\mathbf{f}\colon\ \mathbf{X} \to \mathbf{Y}$ is a continuous map between topological spaces and $\mathbf{x}_n \to \mathbf{x}$ in $\mathbf{X}$, then $\mathbf{f}(\mathbf{x}_n) \to \mathbf{f}(\mathbf{x})$ in $\mathbf{Y}$.
3. For a Hausdorff topological space $\mathbf{X}$, a subset $\mathbf{K} \subseteq \mathbf{X}$ is compact if and only if every sequence in $\mathbf{K}$ has a subsequence converging to a point in $\mathbf{K}$.
4. A compact subset of a Hausdorff topological space is closed.

Definition

A collection $\mathcal{A}$ of subsets of a set $\mathbf{X}$ is a σ-field in $\mathbf{X}$ if:

1. $\mathbf{X} \in \mathcal{A}$;
2. If $\mathbf{U} \in \mathcal{A}$, then $\mathbf{U}^C = \mathbf{X} - \mathbf{U} \in \mathcal{A}$;
3. Any countable union $\bigcup_{j=1}^{\infty} \mathbf{U}_j \in \mathcal{A}$ whenever $\mathbf{U}_j \in \mathcal{A}$ for all j.

$(\mathbf{X}, \mathcal{A})$ is called a measurable space, and members in $\mathcal{A}$ are called measurable sets.

Definition

Suppose $(\mathbf{X}, \mathcal{A})$ is a measurable space, $\mu : \mathcal{A} \to \mathbb{R}$ is called a measure if:

1. $\mu(\mathbf{A}) \geq 0$ for any $\mathbf{A} \in \mathcal{A}$;
2. $\mu(\emptyset) = 0$;
3. For any disjoint countable collection $\{\mathbf{A}_j\}_{j=1}^{\infty} \subseteq \mathcal{A}$, $\mu(\bigcup_{j=1}^{\infty} \mathbf{A}_j) = \sum_{j=1}^{\infty} \mu(\mathbf{A}_j)$.

$(\mathbf{X}, \mathcal{A}, \mu)$ is called a measure space.

Relevant concepts:

- Suppose $\mathbf{X}$ is a measurable space and $\mathbf{Y}$ is a topological space, then a map $\mathbf{f} : \mathbf{X} \to \mathbf{Y}$ is measurable if $\mathbf{f}^{-1}(\mathbf{U})$ is measurable in $\mathbf{X}$ whenever $\mathbf{U}$ is open in $\mathbf{Y}$.
- Suppose $\mathcal{O}$ is a collection of subsets of $\mathbf{X}$, then the σ-field generated by $\mathcal{O}$ is defined as the smallest σ-field containing $\mathcal{O}$, which is equal to the intersection of all σ-field that contains $\mathcal{O}$.
- A σ-field is separable if it is generated by a countable collection of subsets.
- Suppose $\mathbf{X}$ is a topological space, then the σ-field generated by the collection of all open sets in $\mathbf{X}$ is called Borel σ-field of $\mathbf{X}$, denoted by $\mathcal{B}$. Members of $\mathcal{B}$ are called Borel sets.
- A map $\mathbf{f} : \mathbf{X} \to \mathbf{Y}$ between topological spaces is Borel-measurable if it is measurable w.r.t. the Borel σ-field of $\mathbf{X}$, (i.e. $\mathbf{f}^{-1}(\mathbf{U})$ is Borel-measurable in $\mathbf{X}$ for any open set $\mathbf{U}$ in $\mathbf{Y}$). (Thus, any continuous map between topological spaces is Borel-measurable.)

Relevant concepts (continued): For a measure space $(\mathbf{X}, \mathcal{A}, \mu)$

- μ is σ-finite if there exists a sequence $\{\mathbf{A}_j\}_{j=1}^{\infty} \subseteq \mathcal{A}$ such that $\mathbf{X} = \bigcup_{j=1}^{\infty} \mathbf{A}_j$ and $\mu(\mathbf{A}_j) < \infty$ for any j.
- If the range of μ is extended to $(-\infty, \infty]$ or $[-\infty, \infty)$, then μ is called a signed measure.
- When $\mu(\mathbf{X}) = 1$ so $(\mathbf{X}, \mathcal{A}, \mu)$ is a probability space, let

$$\bar{\mathcal{A}} = \{\mathbf{A} \cup \mathbf{N} : \mathbf{A} \in \mathcal{A}, \mathbf{N} \subseteq \mathbf{B}, \mathbf{B} \in \mathcal{A}, \mu(\mathbf{B}) = 0\} \tag{1}$$

$$\bar{\mu}(\mathbf{A} \cup \mathbf{N}) = \mu(\mathbf{A}) \tag{2}$$

Then $(\mathbf{X}, \bar{\mathcal{A}})$ is a measurable space and $\bar{\mu}$ is a well-defined probability measure on it. $\bar{\mathcal{A}}$ is called the μ-completion of $\mathcal{A}$.

Definition

A metric space $(\mathbb{D}, \mathbf{d})$ is a set $\mathbb{D}$ along with a metric $\mathbf{d} : \mathbb{D} \times \mathbb{D} \to \mathbb{R}$ that satisfies

1. $\mathbf{d}(\mathbf{x}, \mathbf{y}) \geq 0$, and $\mathbf{d}(\mathbf{x}, \mathbf{y}) = 0$ if and only if $\mathbf{x} = \mathbf{y}$;
2. $\mathbf{d}(\mathbf{x}, \mathbf{y}) = \mathbf{d}(\mathbf{y}, \mathbf{x})$;
3. $\mathbf{d}(\mathbf{x}, \mathbf{y}) \geq \mathbf{d}(\mathbf{x}, \mathbf{z}) + \mathbf{d}(\mathbf{z}, \mathbf{y})$.

for any $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{D}$.

Note: $\mathbf{d}$ is called a semimetric on $\mathbb{D}$ if it only satisfies [2][3] and

1' $\mathbf{d}(\mathbf{x}, \mathbf{y}) \geq 0$ for any $\mathbf{x}, \mathbf{y} \in \mathbb{D}$.

- A semimetric space is also a topological space with open sets generated by applying arbitrary unions to the open r-balls

$$Br(\mathbf{x}) = \{\mathbf{y} : \mathbf{d}(\mathbf{x}, \mathbf{y}) < r\}. \tag{3}$$

  where $r > 0$ and $\mathbf{x} \in \mathbb{D}$.

- A metric space is also a Hausdorff space, and in this case, a sequence $\{\mathbf{x}_n\} \subseteq \mathbb{D}$ converges to $\mathbf{x} \in \mathbb{D}$ if $\mathbf{d}(\mathbf{x}_n, \mathbf{x}) \to 0$.
- For a semimetric space, $\mathbf{d}(\mathbf{x}_n, \mathbf{x}) \to 0$ only ensures that $\mathbf{x}_n$ converges to elements in the equivalence class of $\mathbf{x}$, where the equivalence class of $\mathbf{x}$ consists of all $\{\mathbf{y} \in \mathbb{D} : \mathbf{d}(\mathbf{x}, \mathbf{y}) = 0\}$.

Relevant concepts:

- Two metrics $\mathbf{d}_1$ and $\mathbf{d}_2$ on $\mathbb{D}$ are strongly equivalent if there exists $\alpha, \beta > 0$ such that

$$\alpha \mathbf{d}_1(\mathbf{x}, \mathbf{y}) \leq \mathbf{d}_2(\mathbf{x}, \mathbf{y}) \leq \beta \mathbf{d}_1(\mathbf{x}, \mathbf{y}) \quad (\forall \mathbf{x}, \mathbf{y} \in \mathbb{D}) \tag{4}$$

- Suppose $(\mathbb{D}, \mathbf{d})$ is a semimetric space, $\{\mathbf{x}_n\} \subseteq \mathbb{D}$ is called a Cauchy sequence if $\mathbf{d}(\mathbf{x}_m, \mathbf{x}_n) \to 0$ as $m, n \to \infty$. $(\mathbb{D}, \mathbf{d})$ is complete if any Cauchy sequence converges to a point in $\mathbb{D}$.
- Two metric spaces are isometric if there is a distance-perseving bijection between them.
- A map $\mathbf{f} : \mathbf{X} \to \mathbf{Y}$ between topological spaces is a homeomorphism if $\mathbf{f}$ is a continuous bijection and $\mathbf{f}^{-1}$ is continuous.
- A Polish space is a space which is homeomorphic to a separable and complete metric space.
- A Suslin set is the image of a Polish space under continuous mapping. If a Suslin set is also a Hausdorff topological space, then it is called a Suslin space.
- A subset $\mathbf{K}$ is totally bounded (or precompact) if for any $r > 0$, $\mathbf{K}$ can be covered by finite many open r-balls.

Definition: Normed Space

A normed space $(\mathbb{D}, \| \cdot \|)$ is a vector space $\mathbb{D}$ equipped with a norm $\| \cdot \|$ which is a map $\mathbb{D} \to \mathbb{R}$ such that, for all $\mathbf{x}, \mathbf{y} \in \mathbb{D}$ and $\alpha \in \mathbb{R}$,

1. $\| \mathbf{x} \| \geq 0$, and $\| \mathbf{x} \| = 0$ if and only if $\mathbf{x} = 0$;
2. $\| \alpha \mathbf{x} \| = |\alpha| \| \mathbf{x} \|$;
3. $\| \mathbf{x} + \mathbf{y} \| \leq \| \mathbf{x} \| + \| \mathbf{y} \|$.

- A seminorm is a map that only satisfies [2][3] and

  1' $\| \mathbf{x} \| \geq 0$ $(\forall \mathbf{x} \in \mathbb{D})$.

- A normed space is a metric space with $\mathbf{d}(\mathbf{x}, \mathbf{y}) = \| \mathbf{x} - \mathbf{y} \|$.
- A complete normed space is called a Banach space.

Useful Conclusions:

- A map $\mathbf{f} : \mathbb{D} \to \mathbb{E}$ between two semimetric space is continuous at $\mathbf{x} \in \mathbb{D}$ if and only if for all $\{\mathbf{x}_n\} \subseteq \mathbb{D}$ and $\mathbf{x}_n \to \mathbf{x}$, we have $\mathbf{f}(\mathbf{x}_n) \to \mathbf{f}(\mathbf{x})$.
- Suppose $\mathbb{D}$ is a metric space and $\mathbf{f} : \mathbb{D} \to \mathbb{R}$, then the following are equivalent:
  1. For any $c \in \mathbb{R}$, $\{\mathbf{y} : \mathbf{f}(y) \geq c\}$ is a closed set.
  2. For any $\mathbf{y}_0 \in \mathbb{D}$, $\limsup_{\mathbf{y} \to \mathbf{y}_0} \mathbf{f}(\mathbf{y}) \leq \mathbf{f}(\mathbf{y}_0)$.
- Every metric space $\mathbb{D}$ has a completion $\bar{\mathbb{D}}$ which has a dense subset isometric with $\mathbb{D}$.
- If a metric space $\mathbb{D}$ is separable, then the Borel σ-field of $\mathbb{D}$ is also separable.
- Any open subset of a Polish space is also Polish.
- Suppose $(\mathbb{D}, \mathbf{d})$ is a complete semimetric space, then
  - A subset $\mathbf{K} \subseteq \mathbb{D}$ is compact if and only if $\mathbf{K}$ is closed and totally bounded.
  - $\mathbf{K} \subseteq \mathbb{D}$ is totally bounded if and only if every sequence in $\mathbf{K}$ has a Cauchy subsequence.

- For an arbitrary set T, define $\ell^\infty(T) = \{\mathbf{f} : T \to \mathbb{R} : \mathbf{f} \text{ is bounded}\}$
- For $\forall \mathbf{f}_1, \mathbf{f}_2 \in \ell^\infty(T)$ and $\forall \alpha_1, \alpha_2 \in \mathbb{R}$, define

$$(\alpha_1 \mathbf{f}_1 + \alpha_2 \mathbf{f}_2)(t) = \alpha_1 \mathbf{f}_1(t) + \alpha_2 \mathbf{f}_2(t) \tag{5}$$

  then $\ell^\infty(T)$ is a linear space.
- For $\mathbf{f} \in \ell^\infty(T)$, define $\| \mathbf{f} \|_T = \sup_{\mathbf{t} \in T} |\mathbf{f}(\mathbf{t})|$, then $(\ell^\infty(T), \| \cdot \|_T)$ is a normed space. And $\| \mathbf{f} \|_T$ is called the uniform norm of $\mathbf{f}$.
- It can be proved that $(\ell^\infty(T), \| \cdot \|_T)$ is a Banach space, and is separable if and only if T is countable.

Metric Space

Examples (continued)

- For a semimetric ρ on T, define

  $UC(T,\rho) = \{\mathbf{f} : T \to \mathbb{R} : \mathbf{f} \text{ is bounded, uniformly } \rho\text{-continuous}\}$

  where uniformly ρ-continuous is defined as

$$\lim_{\delta \downarrow 0} \sup_{\rho(\mathbf{s},\mathbf{t})<\delta} |\mathbf{f}(\mathbf{s}) - \mathbf{f}(\mathbf{t})| = 0 \tag{6}$$

  then $UC(T,\rho)$ is a subspace of $\ell^\infty(T)$.

Theorem (Arzela-Ascoli)

1. Suppose ρ is a semimetric on T and (T,ρ) is totally bounded, $\mathbf{K} \subseteq UC(T,\rho)$, then $\bar{\mathbf{K}}$ is compact if and only if
   - (1) $\exists \mathbf{t}_0 \in T$ such that $\sup_{\mathbf{x}\in\mathbf{K}} |\mathbf{x}(\mathbf{t}_0)| < \infty$;
   - (2) $\lim_{\delta\downarrow 0} \sup_{\mathbf{x}\in\mathbf{K}} \left(\sup_{\mathbf{s},\mathbf{t}\in T, \rho(\mathbf{s},\mathbf{t})<\delta} |\mathbf{x}(\mathbf{s}) - \mathbf{x}(\mathbf{t})| \right) = 0$
2. Suppose $\mathbf{K} \subseteq \ell^\infty(T)$, then $\bar{\mathbf{K}}$ is σ-compact if and only if there exists a semimetric ρ such that (T,ρ) is totally bounded and $\mathbf{K} \subseteq UC(T,\rho)$.

- Suppose $(\mathbb{D}, \mathbf{d})$ and $(\mathbb{E}, \mathbf{e})$ are two metric spaces.
- For $\forall \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{D}$, $\forall \mathbf{y}_1, \mathbf{y}_2 \in \mathbb{E}$, define

$$\rho\left((\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2)\right) = \mathbf{d}(\mathbf{x}_1, \mathbf{x}_2) \vee \mathbf{e}(\mathbf{y}_1, \mathbf{y}_2). \tag{7}$$

  Then $(\mathbb{D} \times \mathbb{E}, \rho)$ forms a metric space (Cartesian product space).
- $(\mathbf{x}_n, \mathbf{y}_n) \to (\mathbf{x}_0, \mathbf{y}_0)$ in $\mathbb{D} \times \mathbb{E}$ if and only if $\mathbf{x}_n \to \mathbf{x}_0$ in $\mathbb{D}$ and $\mathbf{y}_n \to \mathbf{y}_0$ in $\mathbb{E}$.

Outer Expectation

- Suppose $(\Omega, \mathcal{A}, P)$ is a probability space. In some statistical problems, the map of interest $T : \Omega \to \bar{\mathbb{R}} = [-\infty, \infty]$ may not be measurable. Hence, we need to introduce the concept of outer expectation.

Definition (outer expectation and inner expectation)

Suppose $T : \Omega \to \bar{\mathbb{R}}$ is an arbitrary map.

- Define the outer expectation of T w.r.t. the probability measure P as

$$E^*(T) = \inf\{E(U) |\ U : \Omega \to \bar{\mathbb{R}} \text{ measurable},\ U \geq T,\ E(U) \text{ exists}\} \tag{8}$$

- Define the inner expectation of T w.r.t. the probability measure P as

$$E_*(T) = -E^*(-T) = \sup\{E(U) |\ U : \Omega \to \bar{\mathbb{R}} \text{ measurable},\ U \leq T,\ E(U) \text{ exists}\} \tag{9}$$

Here, $E(U)$ exists means that at least one of $E(U^+)$ and $E(U^-)$ is finite.

Outer Expectation

Lemma

For any $T : \Omega \to \bar{\mathbb{R}}$, there exists a minimal measurable majorant $T^* : \Omega \to \bar{\mathbb{R}}$ with

(1) T^* is measurable and $T^* \geq T$ (a.s.);

(2) For every measurable $U : \Omega \to \bar{\mathbb{R}}$ with $U \geq T$ (a.s.), $U \geq T^*$ (a.s.)

(3) For any T^* satisfying (1)(2), $E^*(T) = E(T^*)$ as long as $E(T^*)$ exists. The last statement is true if $E^*(T) < \infty$.

Thus if both T^* and T^{**} satisfy (1) and (2), then $T^* = T^{**}$ (a.s.)

Similarly, define $T_* = -(-T)^*$ as the maximal measurable majorant of T. Then

(1') T_* is measurable and $T_* \leq T$ (a.s.);

(2') For every measurable $U : \Omega \to \bar{\mathbb{R}}$ with $U \leq T$ (a.s.), $U \leq T_*$ (a.s.)

(3') For any T_* satisfying (1')(2'), $E_*(T) = E(T_*)$ as long as $E(T_*)$ exists. The last statement is true if $E_*(T) > -\infty$.

Outer Expectation

Definition (outer probability and inner probability)

For any $\mathbf{B} \subseteq \Omega$, define

- the outer probability of $\mathbf{B}$ w.r.t. the probability measure P as

  $P^*(\mathbf{B}) = \inf\{P(\mathbf{A}) : \mathbf{A} \in \mathcal{A}, \mathbf{A} \supseteq \mathbf{B}\}$;

- the inner probability of $\mathbf{B}$ w.r.t. the probability measure P as

  $P_*(\mathbf{B}) = 1 - P^*(\mathbf{B}^c) = \sup\{P(\mathbf{A}) : \mathbf{A} \in \mathcal{A}, \mathbf{A} \subseteq \mathbf{B}\}$

Then we can prove that for any $\mathbf{B} \subseteq \Omega$

- $P^*(\mathbf{B}) = E^*(I_{\mathbf{B}})$, $P_*(\mathbf{B}) = E_*(I_{\mathbf{B}})$;
- $\mathbf{B}^* = \{\omega : (I_{\mathbf{B}})^*(\omega) \geq 1\}$ is measurable with $\mathbf{B}^* \supseteq \mathbf{B}$, $P^*(\mathbf{B}) = P(\mathbf{B}^*)$ and $(I_{\mathbf{B}})^* = I_{\mathbf{B}^*}$;
- $\mathbf{B}_* = [(\mathbf{B}^c)^*]^c$ with $P_*(\mathbf{B}) = P(\mathbf{B}_*)$;
- $(I_{\mathbf{B}})^* + (I_{\mathbf{B}^c})_* = 1$.

Outer Expectation

Properties of outer expectation: For arbitrary map $S, T : \Omega \to \mathbb{R}$, the following statements are true almost surely

- $S_* + T^* \leq (S+T)^* \leq S^* + T^*$ with all equalities if S is measurable.
- $S_* + T_* \leq (S+T)_* \leq S_* + T^*$ with all equalities if T is measurable.
- $(S-T)^* \geq S^* - T^*$
- $|S^* - T^*| \leq |S-T|^*$
- For any $c \in \mathbb{R}$, $[I_{(T>c)}]^* = I_{(T^*>c)}$ and $[I_{(T\geq c)}]_* = I_{(T_* \geq c)}$
- $(S \vee T)^* = S^* \vee T^*$
- $(S \wedge T)^* \leq S^* \wedge T^*$ with equality if S or T is measurable.

Outer Expectation

Properties of outer probability: For any $\mathbf{A}, \mathbf{B} \subseteq \Omega$,

- $(\mathbf{A} \cup \mathbf{B})^* = \mathbf{A}^* \cup \mathbf{B}^*$, $(\mathbf{A} \cap \mathbf{B})_* = \mathbf{A}_* \cap \mathbf{B}_*$.
- $(\mathbf{A} \cap \mathbf{B})^* \subseteq \mathbf{A}^* \cap \mathbf{B}^*$, $(\mathbf{A} \cup \mathbf{B})_* \supseteq \mathbf{A}_* \cup \mathbf{B}_*$, with equality if either $\mathbf{A}$ or $\mathbf{B}$ is measurable.
- If $\mathbf{A} \cap \mathbf{B} = \emptyset$, then

$$P_*(\mathbf{A}) + P_*(\mathbf{B}) \leq P_*(\mathbf{A} \cup \mathbf{B}) \leq P^*(\mathbf{A} \cup \mathbf{B}) \leq P^*(\mathbf{A}) + P^*(\mathbf{B}) \tag{10}$$

Lemma

Let $T : \Omega \to \mathbb{R}$ be an arbitrary map and suppose $\phi : \mathbb{R} \to \mathbb{R}$ is monotone with an extension to $\bar{\mathbb{R}}$.

Then the following statements are true almost surely, provided they are well-defined:

- If ϕ is non-decreasing, then
  - $\phi(T^*) \geq [\phi(T)]^*$, with equality if ϕ is left-continuous on $[-\infty, \infty)$;
  - $\phi(T_*) \leq [\phi(T)]_*$, with equality if ϕ is right-continuous on $(-\infty, \infty]$;
- If ϕ is non-increasing, then
  - $\phi(T^*) \leq [\phi(T)]_*$, with equality if ϕ is left-continuous on $[-\infty, \infty)$;
  - $\phi(T_*) \geq [\phi(T)]^*$, with equality if ϕ is right-continuous on $(-\infty, \infty]$;

Theorem (Jensen's Inequality)

Let $T : \Omega \to \mathbb{R}$ be an arbitrary map with $E^*|T| < \infty$, and suppose $\phi : \mathbb{R} \to \mathbb{R}$ is convex, then

1. $E^*[\phi(T)] \geq \phi[E^*(T)] \vee \phi[E_*(T)]$
2. If ϕ is also monotone, then $E_*[\phi(T)] \geq \phi[E^*(T)] \wedge \phi[E_*(T)]$

Chebyshev's Inequality

Let $T : \Omega \to \mathbb{R}$ be an arbitrary map. $\phi : [0, \infty) \to [0, \infty)$ is positive on $(0, \infty)$ and non-decreasing, then for any $u > 0$, $P^*(|T| \geq u) \leq E^*[\phi(|T|)]/\phi(u)$

Monotone Convergence

Let $T_n, T : \Omega \to \mathbb{R}$ be arbitrary maps, with $T_n \uparrow T$ pointwise on a set of inner probability 1. Then $T_n^* \uparrow T^*$ (a.s.). Additionally, if $E^*(T_n) > -\infty$ for some n, then $E^*(T_n) \uparrow E^*(T)$.

Dominated Convergence

Let $T_n, T, S : \Omega \to \mathbb{R}$ be maps with $|T_n - T|^* \to 0(a.s.)$, $|T_n| \leq S(\forall n)$, and $E^*(S) < \infty$, then $E^*(T_n) \to E^*(T)$.

Suppose $(\Omega, \bar{\mathcal{A}}, \bar{P})$ is the P-completion of probability space $(\Omega, \mathcal{A}, P)$

- $\bar{\mathcal{A}} = \{\mathbf{A} \cup \mathbf{N} : \mathbf{A} \in \mathcal{A}, \mathbf{N} \subseteq \mathbf{B}, \mathbf{B} \in \mathcal{A}, P(\mathbf{B}) = 0\}$
- $\bar{P}(\mathbf{A} \cup \mathbf{N}) = P(\mathbf{A})$

Then

- for any $\bar{\mathcal{A}}$-measurable map $\bar{S} : (\Omega, \bar{\mathcal{A}}) \to \mathbb{R}$, there exists an $\mathcal{A}$-measurable map $S : (\Omega, \mathcal{A}) \to \mathbb{R}$ such that $P^*(S \neq \bar{S}) = 0$.
- For any $T : (\Omega, \mathcal{A}, P) \to \bar{\mathbb{R}}$, define $\bar{T} : (\Omega, \bar{\mathcal{A}}, \bar{P}) \to \bar{\mathbb{R}}, \omega \mapsto T(\omega)$.
  Let T^* be the minimal measurable majorant of T w.r.t. P.
  Let $\bar{T}^*$ be the minimal measurable majorant of $\bar{T}$ w.r.t. $\bar{P}$.
  Then $P^*(T^* \neq \bar{T}^*) = 0$.

Consider a measurable map $\phi : (\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{P}) \to (\Omega, \mathcal{A}, P)$ and any map $T : (\Omega, \mathcal{A}, P) \to \mathbb{R}$ where for $\mathbf{A} \in \mathcal{A}$, $P(\mathbf{A}) \triangleq \tilde{P} \circ \phi^{-1}(\mathbf{A}) = \tilde{P}(\phi \in \mathbf{A})$.

- $T \circ \phi : (\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{P}) \to \mathbb{R}$
- P is a probability measure on $(\Omega, \mathcal{A})$.
- Let T^* be the minimal measurable majorant of T w.r.t. P.
- By definition, $T^* \circ \phi : (\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{P}) \to \mathbb{R}$ is measurable and $T^* \circ \phi \geq T \circ \phi$.
  Thus, $T^* \circ \phi \geq (T \circ \phi)^*$.
- ϕ is perfect if $T^* \circ \phi = (T \circ \phi)^*$ (a.s.) for any bounded map $T : \Omega \to \mathbb{R}$. In this case, $\tilde{P}^*(\phi \in \mathbf{A}) = (\tilde{P} \circ \phi^{-1})^*(\mathbf{A})$ for any $\mathbf{A} \subseteq \Omega$.

- Example: coordinate projection in a product probability space is a perfect map.
- Specifically, suppose $(\Omega_1, \mathcal{A}_1, P_1)$ and $(\Omega_2, \mathcal{A}_2, P_2)$ are two probability space.
- Let $T_1 : (\Omega_1, \mathcal{A}_1, P_1) \to \mathbb{R}$ be a bounded map.
  Define $T : (\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \Omega_2, P_1 \times P_2) \to \mathbb{R}, \omega = (\omega_1, \omega_2) \mapsto T_1(\omega_1)$ to be a map from the product space to the real line.
- Also let π_1 be the projection on the first coordinate. then $T = T_1 \circ \pi_1$.
- It can be proved that π_1 is a perfect map. Thus $T^* = (T_1 \circ \pi_1)^* = T_1^* \circ \pi_1$. Thus, to find the image of any $\omega = (\omega_1, \omega_2)$ under T^*, we can ignore ω_2 and find the image of ω_1 under T_1^*.

Fubini's Theorem

Let $T : (\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \Omega_2, P_1 \times P_2) \to \mathbb{R}$ be an arbitrary map. For any fixed $\omega_1 \in \Omega_1$, define

$$E_2^*(T)(\omega_1) = \inf\{E_2(U) = \int_{\Omega_2} U(\omega_2)dP_2(\omega_2) |\ U : \Omega_2 \to \bar{\mathbb{R}} \text{ measurable}, U(\omega_2) \geq T(\omega_1, \omega_2), E_2(U) \text{ exists}\}. \tag{11}$$

Also, define

$$E_1^*[E_2^*(T)] = \inf\{E_1(U) = \int_{\Omega_1} U(\omega_1)dP_1(\omega_1) |\ U : \Omega_1 \to \bar{\mathbb{R}} \text{ measurable}, U(\omega_1) \geq E_2^*(T)(\omega_1), E_1(U) \text{ exists}\}. \tag{12}$$

$$E^*(T) = \inf\{E(U) = \int_{\Omega_1 \times \Omega_2} U(\omega_1, \omega_2)d(P_1 \times P_2)(\omega_1, \omega_2) |\ U : \Omega_1 \times \Omega_2 \to \mathbb{R} \text{ measurable}, U \geq T, E(U) \text{ exists}\}. \tag{13}$$

We can also define $E_{2*}(T)(\omega_1)$, $E_{1*}[E_{2*}(T)]$ and $E_*(T)$ similarly. Then

$$E_*(T) \leq E_{1*}[E_{2*}(T)] \leq E_1^*[E_2^*(T)] \leq E^*(T). \tag{14}$$

Linear Operators and Differentiation

- A linear operator is a map $T : \mathbb{D} \to \mathbb{E}$ between normed spaces with the property that $T(a\mathbf{x} + b\mathbf{y}) = aT(\mathbf{x}) + bT(\mathbf{y})$ $(\forall a, b \in \mathbb{R}, \forall \mathbf{x}, \mathbf{y} \in \mathbb{D})$.
- If $\mathbb{E} = \mathbb{R}$, then T is called a linear functional.
- T is bounded if $\| T \| \triangleq \sup_{\mathbf{x} \in \mathbb{D}, \|\mathbf{x}\|=1} \| T(x) \| < \infty$.
- T is continuous at $\mathbf{x}_0 \in \mathbb{D}$ if $T(\mathbf{x}_n) \to T(\mathbf{x}_0)$ in $\mathbb{E}$ whenever $\mathbf{x}_n \to \mathbf{x}_0$ in $\mathbb{D}$.
- Lemma: $T : \mathbb{D} \to \mathbb{E}$ is a linear operator between normed spaces. Then the following are equivalent:
  - T is continuous at a point $\mathbf{x}_0 \in \mathbb{D}$;
  - T is continuous at any point in $\mathbb{D}$;
  - T is bounded

- For normed spaces $\mathbb{D}$ and $\mathbb{E}$, let $B(\mathbb{D}, \mathbb{E})$ be the space of all bounded linear operators. For $\forall T_1, T_2 \in B(\mathbb{D}, \mathbb{E})$ and $\forall \alpha_1, \alpha_2 \in \mathbb{R}$, define

$$(\alpha_1 T_1 + \alpha_2 T_2)(\mathbf{x}) = \alpha_1 T_1(\mathbf{x}) + \alpha_2 T_2(\mathbf{x}) \tag{15}$$

  then $B(\mathbb{D}, \mathbb{E})$ is a linear space.

- With $\| T \| \triangleq \sup_{\mathbf{x} \in \mathbb{D}, \|\mathbf{x}\|=1} \| T(x) \|$, $(B(\mathbb{D}, \mathbb{E}), \| \cdot \|)$ forms a normed space.
- If $\mathbb{E}$ is a Banach space, then $B(\mathbb{D}, \mathbb{E})$ is also a Banach space.
- Generally, when $\mathbb{D}$ is not a Banach space, T has a unique continuous extension $\bar{T} : \bar{\mathbb{D}} \to \bar{\mathbb{E}}$.

- For $\forall T \in B(\mathbb{D}, \mathbb{E})$,
  the null space $N(T) \triangleq \{\mathbf{x} : T(\mathbf{x}) = \mathbf{0}\}$;
  the range space $R(T) \triangleq \{T(\mathbf{x}) : \mathbf{x} \in \mathbb{D}\}$.
- Both $N(T)$ and $R(T)$ are linear space.
- T is one-to-one if and only if $N(T) = \{\mathbf{0}\}$.
- Conclusions for inverse mapping: Let $T \in B(\mathbb{D}, \mathbb{E})$, then
  1. T has a continuous inverse $T^{-1} : R(T) \to \mathbb{D}$ if and only if $\exists c > 0 s.t. \parallel T(\mathbf{x}) \parallel \geq c \parallel \mathbf{x} \parallel \quad (\forall \mathbf{x} \in \mathbb{D})$;
  2. If $\mathbb{D}$ and $\mathbb{E}$ are complete, and T is a continuous injection, then T^{-1} is continuous if and only if $R(T)$ is closed.

- Let $U = \{\mathbf{x} \in \mathbb{D} : \| \mathbf{x} \| \leq 1\}$ be the unit ball in a normed space $\mathbb{D}$.
- A linear operator $T : \mathbb{D} \to \mathbb{E}$ between normed space is a compact operator if $\overline{T(U)}$ is compact in $\mathbb{E}$.
- In later chapters, we will encounter linear operators in a form of $(T + K)$, where T is continuous and invertible, and K is compact.

Lemma

Let $A = T + K : \mathbb{D} \to \mathbb{E}$ be a linear operator between Banach spaces, where T is continuous, invertible and K is compact. If $N(A) = \mathbf{0}$, then A is also continuously invertible.

- An operator A is a contraction operator if $\| A \| < 1$.
- Suppose $\mathbb{D}$ is a normed space and $A : \mathbb{D} \to \mathbb{D}$ is a linear contraction operator. Then $(I - A)$ is continuous and invertible with $(I - A)^{-1} = \sum_{j=0}^{\infty} A^j$. (Here, I is the identity map.)

- Let $\mathbb{D}$ and $\mathbb{E}$ be two normed spaces, and $\phi : \mathbb{D}_\phi \subseteq \mathbb{D} \to \mathbb{E}$. ϕ is Gateaux differentiable at $\theta \in \mathbb{D}_\phi$ in the direction h, if there exists a quantity $\phi_\theta'(h) \in \mathbb{E}$ s.t.

$$\frac{\phi(\theta + t_n h) - \phi(\theta)}{t_n} \to \phi_\theta'(h) \tag{16}$$

  for any scalar sequence $t_n \to 0$.

- Limitation: Gateaux-differentiability is usually not strong enough for the applications of functional derivatives in Z-estimators and delta method.

- $\phi : \mathbb{D}_\phi \subseteq \mathbb{D} \to \mathbb{E}$ is Hadamard differentiable at $\theta \in \mathbb{D}_\phi$ if there exists a continuous linear operator $\phi'_\theta : \mathbb{D} \to \mathbb{E}$ s.t. for any scalar sequence $t_n \to 0$ and $h_n \to h \in \mathbb{D}$

$$\frac{\phi(\theta + t_n h_n) - \phi(\theta)}{t_n} \to \phi'_\theta(h). \tag{17}$$

- If (17) is required only for $h_n \to h \in \mathbb{D}_0 \subseteq \mathbb{D}$, then say ϕ is Hadamard differentiable tangentially to $\mathbb{D}_0$.
- $\phi : \mathbb{D}_\phi \subseteq \mathbb{D} \to \mathbb{E}$ is compact differentiable at $\theta \in \mathbb{D}_\phi$ if there exists a continuous linear operator $\phi'_\theta : \mathbb{D} \to \mathbb{E}$ s.t. for any compact set $\mathbf{K} \subseteq \mathbb{D}$

$$\sup_{h \in \mathbf{K}, (\theta + th) \in \mathbb{D}_\phi} \left\| \frac{\phi(\theta + th) - \phi(\theta)}{t} - \phi'_\theta(h) \right\| \to 0 \ (t \to 0) \tag{18}$$

- Hadamard differentiability is equivalent to compact differentiability.

Lemma: Chain Rule

Suppose

$\phi : \mathbb{D}_\phi \subseteq \mathbb{D} \to \mathbb{E}_\psi \subseteq \mathbb{E}$ is Hadamard differentiable at $\theta \in \mathbb{D}_\phi$ tangentially to $\mathbb{D}_0 \subseteq \mathbb{D}$,

$\psi : \mathbb{E}_\psi \subseteq \mathbb{E} \to \mathbb{F}$ is Hadamard differentiable at $\phi(\theta)$ tangentially to $\phi_\theta^{'}(\mathbb{D}_0)$.

Then $\psi \circ \phi : \mathbb{D}_\phi \to \mathbb{F}$ is Hadamard differentiable at θ tangentially to $\mathbb{D}_0$ with derivative $\psi_{\phi(\theta)}^{'} \circ \phi_\theta^{'}$

- $\phi : \mathbb{D}_\phi \subseteq \mathbb{D} \to \mathbb{E}$ is Frechet differentiable if there exists a continuous linear operator $\phi_\theta^{'} : \mathbb{D} \to \mathbb{E}$ s.t. for any bounded set $\mathbf{K} \subseteq \mathbb{D}$

$$\sup_{h \in \mathbf{K}, (\theta + th) \in \mathbb{D}_\phi} \left|\left| \frac{\phi(\theta + th) - \phi(\theta)}{t} - \phi_\theta^{'}(h) \right|\right| \to 0 \ (t \to 0), \tag{19}$$

or equivalently,

$$\| \phi(\theta + h) - \phi(\theta) - \phi_\theta^{'}(h) \| = o(\| h \|) \ (\| h \| \to 0), \tag{20}$$

- Relationships:
  - Frechet differentiability implies Hadamard differentiability
  - Hadamard differentiability implies Gateaux differentiability
- Frechet differentiability: important in Z-estimator theory.

  Hadamard differentiability: important in functional delta method.