Empirical Process Theory

Yasuyuki Matsumura (Kyoto University)

Last Updated: April 16, 2025

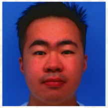
https://yasu0704xx.github.io

Yasuyuki Matsumura

Graduate School of Economics, Kyoto University

Home | Research | Japanese

About Me



I am a second-year master's student in Graduate School of Economics, Kyoto University.

Under the guidance of Yoshihiko Nishiyama and Takahide Yanagi, my research focuses on Econometric Theory and Statistics.

Contact: matsumura.yasuyuki.w85 [at] kyoto-u.jp

Address: Yoshida Honmachi, Sakyo, Kyoto 606-8501, Japan

Other Links: Linktree

Introduction

- The empirical distribution of a random sample is the uniform discrete measure on the observations.
- In this chapter, we study the convergence of this measure and the convergence of the corresponding function, which leads to the uniform law of large numbers and the functional central limit theorem.
- Here we review Chapter 18 of Hansen (2022) [2].
- Useful references include Pollard (1990) [3]; Andrews (1994) [1]; van der Vaart (1998) [4]; and van der Vaart and Wellner (2023) [5].

Preliminaries

Stochastic Convergence in Metric Spaces[1]

- Before starting with emprical process theory, we briefly review the concept of stochastic convergence in metric spaces.
- This section is associated with Chapter 18 of van der Vaart (1998) [4].

18.	Stochastic Convergence in Metric Spaces	255
18.1.	Metric and Normed Spaces	255
18.2.	Basic Properties	258
18.3.	Bounded Stochastic Processes	260
	Problems	263
19.	Empirical Processes	265
19.1.	Empirical Distribution Functions	265
19.2.	Empirical Distributions	269
19.3.	Goodness-of-Fit Statistics	277
19.4.	Random Functions	279
19.5.	Changing Classes	282
19.6.	Maximal Inequalities	284
	Problems	289

[1] To be skipped in the class.

- A metric space is a set $\mathbb{D}$ equipped with a metric.
- A metric function (distance function) is a map $d : \mathbb{D} \times \mathbb{D} \to [0, \infty)$ with the properties

$$d(x, y) = d(y, x), \tag{1}$$

$$d(x, z) \leq d(x, y) + d(y, z) \quad (\text{triangle inequality}), \tag{2}$$

$$d(x, y) = 0 \text{ iff } x = y. \tag{3}$$

- A semimetric satisfies (1) and (2), but not necessarily (3).
- An open ball is a set of the form $\{y \mid d(x, y) < r\}$.
- A subset of a metric space is open iff it is the union of open balls; it is closed iff its complement is open.

Metric Space ii

- A sequence x_n converges to x iff $d(x_n, x) \to 0$; this is denoted by $x_n \to x$.
- The closure $\bar{A}$ of a set $A \subset \mathbb{D}$ consists of all points that are the limit of a sequence in A; it is the smallest closed set containing A.
- The interior $\mathring{A}$ is the collection of all points x s.t. $x \in G \subset A$ for some open set G; it is the largest open set contained in A.
- A function $f : \mathbb{D} \to \mathbb{E}$ between two metric spaces is continuous at a point X iff $f(x_n) \to f(x)$ for every sequence $x_n \to x$; it is continuous at every x iff the inverse image $f^{-1}(G)$ of every open set $G \subset \mathbb{E}$ is open in $\mathbb{D}$.

Metric Space iii

- A subset of a metric space is dense iff its closure is the whole space.
- A metric space is separable iff it has a countable dense subset.
- A subset K of a metric space is compact iff it is closed and every sequence in K has a converging subsequence.
- A subset K is totally bounded iff for every $\epsilon > 0$ it can be covered by finitely many balls of radius ϵ.
- A semimetric space is complete if every Cauchy sequence [2] has a limit.
- A subset of a complete semimetric space is compact iff it is totally bounded and closed.

[2] Cauchy sequence: a sequence s.t. $d(x_n, x_m) \to 0$ as $n, m \to \infty$.

Normed Space

- A normed space $\mathbb{D}$ is a vector space equipped with a norm.
- A norm is a map $||\cdot|| : \mathbb{D} \to [0, \infty)$ s.t., for every $x, y \in \mathbb{D}$ and $\alpha \in \mathbb{R}$,

$$||x + y|| \leq ||x|| + ||y|| \quad \text{(triangle inequality)}, \tag{4}$$

$$||\alpha x|| = |\alpha| \, ||x||, \tag{5}$$

$$||x|| = 0 \text{ iff } x = 0. \tag{6}$$

- A semi-norm satisfies (4) and (5), but not necessarily (6).
- Given a norm, a metric can be defined by $d(x, y) = ||x - y||$.

Borel σ-Field

Def 18.1 (Borel σ-field)

- The Borel σ-field on a metric space $\mathbb{D}$ is the smallest σ-field that contains the open sets (and then also the closed sets).
- A function defined relative to (One or two) metric spaces is called Borel-measurable if it is measurable relative to the Borel σ-field(s).
- A Borel-measurable map $X : \Omega \to \mathbb{D}$ defined on a probability space $(\Omega, \mathcal{U}, \mathbb{P})$ is referred to as a random element with values in $\mathbb{D}$.

Lemma 18.2

- A continuous map between metric spaces is Borel-measurable.

Proof.

- A map $g : \mathbb{D} \to \mathbb{E}$ is continuous iff the inverse image $g^{-1}(G)$ of every open set $G \subset \mathbb{E}$ is open in $\mathbb{D}$.
- In particular, for every open set G, the set $g^{-1}(G)$ is a Borel set in $\mathbb{D}$.
- By definition, the open sets in $\mathbb{E}$ generate the Borel σ-field.
- Thus, the inverse image of a generator of the Borel sets in $\mathbb{E}$ is contained in the Borel σ-field in $\mathbb{D}$.
- Because the inverse image $g^{-1}(\mathcal{G})$ of a σ-field $\mathcal{B}$ generates the σ-field $g^{-1}(\mathcal{B})$, the inverse image of every Borel set is a Borel set.

Examples

- Here we review several examples:
  - Euclidean spaces
  - Extended real line
  - Uniform norm
  - Skorohod space
  - Uniformly continuous functions
  - Product spaces

Example 18.3 (Euclidean Spaces)

- The Euclidean space $\mathbb{R}^k$ is a normed space w.r.t. the Euclidean norm[3], but also w.r.t. many other norms (such as $||x|| = \max_i |x_i|$), all of which are equivalent.
- By the Heine-Borel theorem, a subset of $\mathbb{R}^k$ is compact iff it is closed and bounded.
- A Euclidean space is separable, with, for instance, the vectors with rational coordinates as a cuntable dense subset.
- The Borel σ-field is the usual σ-field is the usual σ-field, generated by the intervals of the type $(-\infty, x]$.

[3] Its square is $||x||^2 = \sum_{i=1}^k x_i^2$

Example 18.4 (Extended Real Line)

- The extended real line $\bar{R} = [-\infty, \infty]$ is the set consisting of all real numbers and the additional elements $-\infty$ and ∞.
- It is a metric space w.r.t.

$$d(x, y) = \left|\Phi(x) - \Phi(y)\right|,$$

  where Φ can be any fixed, bounded, strictly increasing continuous function. For instance, Φ can be the normal distribution function, which satisfies $\Phi(-\infty) = 0$ and $\Phi(\infty) = 1$.
- Convergence of a sequence $x_n \to x$ w.r.t. this metric has the ysual meaning, also if the limit x is $-\infty$ or ∞ (normally we would say that x_n "diverges").
- Consequently, every sequence has a converging subsequence and hence the extended real line is compact.

Example 18.5 (Uniform Norm)

Example 18.6 (Skorohod Space)

Example 18.7 (Uniformly Continuous Functions)

Example 18.8 (Product Sets)

Framework

Glivenko-Cantelli Theorem

Packing, Covering, and Bracketing Numbers

Uniform Law of Large Numbers

Functional Central Limit Theory

Conditions for Asymptotic Equicontinuity

Donsker's Theorem

Refernces

References i

Andrews, D. W. K. (1994). "Empirical process methods in econometrics," *Handbook of Econometrics 4*(37), Robert F. Engle and Daniel L.McFadden, eds., 2247-2294, Elsevier.

Hansen, B. E. (2022). *Probability and Statistics for Economists.* Princeton.

Pollard, D. (1990). *Empirical Processes: Theory and Applications.* NSF-CBMS Regional Conference Series in Probability and Statistics, Volume 2. IMS and ASA.

van der Vaart, A. W. (1998). *Asymptotic Statistics.* Cambridge.

References ii

van der Vaart, A. W. and J. A. Wellner (2023) *Weak Convergence and Empirical Processes.* Springer.