

# Introduction to Empirical Processes and Semiparametric Inference Lecture 06: Metric Spaces

Michael R. Kosorok, Ph.D.

Professor and Chair of Biostatistics

Professor of Statistics and Operations Research

University of North Carolina-Chapel Hill

## Introduction to Part II

The goal of Part II is to provide an in depth coverage of the basics of empirical process techniques which are useful in statistics:

- Chapter 6: mathematical background, metric spaces, outer expectation, linear operators and functional differentiation.
- Chapter 7: stochastic convergence, weak convergence, other modes of convergence.
- Chapter 8: empirical process techniques, maximal inequalities, symmetrization, Glivenk-Canteli results, Donsker results.
- Chapter 9: entropy calculations, VC classes, Glivenk-Canteli and Donsker preservation.
- Chapter 10: empirical process bootstrap.

- Chapter 11: additional empirical process results.
- Chapter 12: the functional delta method.
- Chapter 13: Z-estimators.
- Chapter 14: M-estimators.
- Chapter 15: Case-studies II.

# Topological Spaces

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

- (i)  $\emptyset \in \mathcal{O}$  and  $X \in \mathcal{O}$ , where  $\emptyset$  is the empty set;
- (ii) If  $U_j \in \mathcal{O}$  for  $j = 1, \dots, m$ , then  $\bigcap_{j=1, \dots, m} U_j \in \mathcal{O}$ ;
- (iii) If  $\{U_\alpha\}$  is an arbitrary collection of members of  $\mathcal{O}$  (finite, countable or uncountable), then  $\bigcup_\alpha U_\alpha \in \mathcal{O}$ .

When  $\mathcal{O}$  is a topology in  $X$ , then  $X$  (or the pair  $(X, \mathcal{O})$ ) is a *topological space*, and the members of  $\mathcal{O}$  are called the *open sets* in  $X$ .

For a subset  $A \subset X$ , the *relative topology* on  $A$  consists of the sets  $\{A \cap B : B \in \mathcal{O}\}$ : check that this is a topology.

A map  $f : X \mapsto Y$  between topological spaces is *continuous* if  $f^{-1}(U)$  is open in  $X$  whenever  $U$  is open in  $Y$ .

A set  $B$  in  $X$  is *closed* if and only if its complement in  $X$ , denoted  $X - B$ , is open.

The *closure* of an arbitrary set  $E \in X$ , denoted  $\overline{E}$ , is the smallest closed set containing  $E$ .

The *interior* of an arbitrary set  $E \in X$ , denoted  $E^\circ$ , is the largest open set contained in  $E$ .

A subset  $A$  of a topological space  $X$  is *dense* if  $\overline{A} = X$ .

A topological space  $X$  is *separable* if it has a countable dense subset.

A *neighborhood* of a point  $x \in X$  is any open set that contains  $x$ .

A topological space is *Hausdorff* if distinct points in  $X$  have disjoint neighborhoods.

A sequence of points  $\{x_n\}$  in a topological space  $X$  *converges* to a point  $x \in X$ , denoted  $x_n \rightarrow x$ , if every neighborhood of  $x$  contains all but finitely many of the  $x_n$ .

Suppose  $x_n \rightarrow x$  and  $x_n \rightarrow y$ .

Then  $x$  and  $y$  share all neighborhoods, and  $x = y$  when  $X$  is Hausdorff.

If a map  $f : X \mapsto Y$  between topological spaces is continuous, then  $f(x_n) \rightarrow f(x)$  whenever  $x_n \rightarrow x$  in  $X$ ; to see this,

- Let  $\{x_n\} \subset X$  be a sequence with  $x_n \rightarrow x \in X$ .
- Then for any open  $U \subset Y$  containing  $f(x)$ , all but finitely many  $\{x_n\}$  are in  $f^{-1}(U)$ , and thus all but finitely many  $\{f(x_n)\}$  are in  $U$ .
- Since  $U$  was arbitrary, we have  $f(x_n) \rightarrow f(x)$ .



We now review the important concept of *compactness*:

- A subset  $K$  of a topological space is *compact* if for every set  $A \supset K$ , where  $A$  is the union of a collection of open sets  $\mathcal{S}$ ,  $K$  is also contained in some finite union of sets in  $\mathcal{S}$ .
- When the topological space involved is also Hausdorff, then compactness of  $K$  is equivalent to the assertion that every sequence in  $K$  has a convergent subsequence (converging to a point in  $K$ ).
- This result implies that compact subsets of Hausdorff topological spaces are necessarily closed.
- Note that a compact set is sometimes called a *compact* for short.
- A  $\sigma$ -*compact* set is a countable union of compacts.

A collection  $\mathcal{A}$  of subsets of a set  $X$  is a  $\sigma$ -field in  $X$  (sometimes called a  $\sigma$ -algebra) if:

- (i)  $X \in \mathcal{A}$ ;
- (ii) If  $U \in \mathcal{A}$ , then  $X - U \in \mathcal{A}$ ;
- (iii) The countable union  $\bigcup_{j=1}^{\infty} U_j \in \mathcal{A}$  whenever  $U_j \in \mathcal{A}$  for all  $j \geq 1$ .

Note that (iii) clearly includes finite unions.

When (iii) is only required to hold for finite unions, then  $\mathcal{A}$  is called a *field*.

When  $\mathcal{A}$  is a  $\sigma$ -field in  $X$ , then  $X$  (or the pair  $(X, \mathcal{A})$ ) is a *measurable space*, and the members of  $\mathcal{A}$  are called the *measurable sets* in  $X$ .

If  $X$  is a measurable space and  $Y$  is a topological space, then a map  $f : X \mapsto Y$  is *measurable* if  $f^{-1}(U)$  is measurable in  $X$  whenever  $U$  is open in  $Y$ .

If  $\mathcal{O}$  is a collection of subsets of  $X$  (not necessary open), then there exists a smallest  $\sigma$ -field  $\mathcal{A}^*$  in  $X$  so that  $\mathcal{O} \subset \mathcal{A}^*$ .

This  $\mathcal{A}^*$  is called the  $\sigma$ -field *generated* by  $\mathcal{O}$ .

To see that such an  $\mathcal{A}^*$  exists:

- Let  $\mathcal{S}$  be the collection of all  $\sigma$ -fields in  $X$  which contain  $\mathcal{O}$ .
- Since the collection of all subsets of  $X$  is one such  $\sigma$ -field,  $\mathcal{S}$  is not empty.
- Define  $\mathcal{A}^*$  to be the intersection of all  $\mathcal{A} \in \mathcal{S}$ .
- Clearly,  $\mathcal{O} \in \mathcal{A}^*$  and  $\mathcal{A}^*$  is in every  $\sigma$ -field containing  $\mathcal{O}$ .
- All that remains is to show that  $\mathcal{A}^*$  is itself a  $\sigma$ -field.

To show this,

- Assume that  $A_j \in \mathcal{A}^*$  for all integers  $j \geq 1$ .
- If  $\mathcal{A} \in \mathcal{S}$ , then  $\bigcup_{j \geq 1} A_j \in \mathcal{A}$ .
- Since  $\bigcup_{j \geq 1} A_j \in \mathcal{A}$  for every  $\mathcal{A} \in \mathcal{S}$ , we have  $\bigcup_{j \geq 1} A_j \in \mathcal{A}^*$ .
- Also  $X \in \mathcal{A}^*$  since  $X \in \mathcal{A}$  for all  $\mathcal{A} \in \mathcal{S}$ ; and for any  $A \in \mathcal{A}^*$ , both  $A$  and  $X - A$  are in every  $\mathcal{A} \in \mathcal{S}$ .

Thus  $\mathcal{A}^*$  is indeed a  $\sigma$ -field.

A  $\sigma$ -field is *separable* if it is generated by a countable collection of subsets.

Note: we have already defined “separable” as a characteristic of certain topological spaces.

There is a connection between the two definitions which we will point out shortly in our discussion on metric spaces.

When  $X$  is a topological space, the smallest  $\sigma$ -field  $\mathcal{B}$  generated by the open sets is called the *Borel  $\sigma$ -field* of  $X$ .

Elements of  $\mathcal{B}$  are called *Borel sets*.

A function  $f : X \mapsto Y$  between topological spaces is *Borel-measurable* if it is measurable with respect to the Borel  $\sigma$ -field of  $X$ .

Clearly, a continuous function between topological spaces is also Borel-measurable.

For a  $\sigma$ -field  $\mathcal{A}$  in a set  $X$ , a map  $\mu : \mathcal{A} \mapsto \bar{\mathbb{R}}$  is a *measure* if:

- (i)  $\mu(A) \in [0, \infty]$  for all  $A \in \mathcal{A}$ ;
- (ii)  $\mu(\emptyset) = 0$ ;
- (iii) For any disjoint sequence  $\{A_j\} \in \mathcal{A}$ ,  
$$\mu\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mu(A_j) \text{ (countable additivity).}$$

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



If  $X = A_1 \cup A_2 \cup \dots$  for some finite or countable sequence of sets in  $\mathcal{A}$  with  $\mu(A_j) < \infty$  for all indices  $j$ , then  $\mu$  is  $\sigma$ -finite.

If  $\mu(X) = 1$ , then  $\mu$  is a *probability measure*.

For a probability measure  $P$  on a set  $\Omega$  with  $\sigma$ -field  $\mathcal{A}$ , the triple  $(\Omega, \mathcal{A}, P)$  is called a *probability space*.

If the set  $[0, \infty]$  in Part (i) is extended to  $(-\infty, \infty]$  or replaced by  $[-\infty, \infty)$  (but not both), then  $\mu$  is a *signed measure*.

For a measure space  $(X, \mathcal{A}, \mu)$ , let  $\mathcal{A}^*$  be the collection of all  $E \subset X$  for which there exists  $A, B \in \mathcal{A}$  with  $A \subset E \subset B$  and  $\mu(B - A) = 0$ , and define  $\mu(E) = \mu(A)$  in this setting.

Then  $\mathcal{A}^*$  is a  $\sigma$ -field,  $\mu$  is still a measure, and  $\mathcal{A}^*$  is called the  *$\mu$ -completion of  $\mathcal{A}$* .

# Metric Spaces

A *metric space* is a set  $\mathbb{D}$  together with a *metric*.

A metric or *distance function* is a map  $d : \mathbb{D} \times \mathbb{D} \mapsto [0, \infty)$  where:

- (i)  $d(x, y) = d(y, x)$ ;
- (ii)  $d(x, z) \leq d(x, y) + d(y, z)$  (the *triangle inequality*);
- (iii)  $d(x, y) = 0$  if and only if  $x = y$ .

A *semimetric* or *pseudometric* satisfies (i) and (ii) but not necessarily (iii).

Technically, a metric space consists of the pair  $(\mathbb{D}, d)$ , but usually only  $\mathbb{D}$  is given and the underlying metric  $d$  is implied by the context.

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

$B_r(x) \equiv \{y : d(x, y) < r\}$  for  $r \geq 0$  and  $x \in \mathbb{D}$  (where  $B_0(x) \equiv \emptyset$ ).

A metric space is also Hausdorff, and, in this case, a sequence  $\{x_n\} \in \mathbb{D}$  converges to  $x \in \mathbb{D}$  if  $d(x_n, x) \rightarrow 0$ .

For a semimetric space,  $d(x_n, x) \rightarrow 0$  ensures only that  $x_n$  converges to elements in the *equivalence class* of  $x$ , where the equivalence class of  $x$  consists of all  $\{y \in \mathbb{D} : d(x, y) = 0\}$ .

Accordingly, the closure  $\overline{A}$  of a set  $A \in \mathbb{D}$  is not only the smallest closed set containing  $A$ , as stated earlier, but  $\overline{A}$  also equals the set of all points that are limits of sequences  $\{x_n\} \in A$ .

In addition, two semimetrics  $d_1$  and  $d_2$  on a set  $\mathbb{D}$  are considered equivalent (in a topological sense) if they both generate the same open sets, and equivalent metrics yield the same convergent subsequences.

A map  $f : \mathbb{D} \mapsto \mathbb{E}$  between two semimetric spaces is *continuous at a point*  $x$  if and only if  $f(x_n) \rightarrow f(x)$  for every sequence  $x_n \rightarrow x$ .

The map  $f$  is continuous (in the topological sense) if and only if it is continuous at all points  $x \in \mathbb{D}$ .

The following lemma helps to define *semicontinuity* for real valued maps:

LEMMA 1. *Let  $f : \mathbb{D} \mapsto \mathbb{R}$  be a function on the metric space  $\mathbb{D}$ ; Then*

*TFAE:*

(i) *For all  $c \in \mathbb{R}$ , the set  $\{y : f(y) \geq c\}$  is closed.*

(ii) *For all  $y_0 \in \mathbb{D}$ ,  $\limsup_{y \rightarrow y_0} f(y) \leq f(y_0)$ .*

A function  $f : \mathbb{D} \mapsto \mathbb{R}$  satisfying either (i) or (ii) (and hence both) of the conditions in Lemma 1 is said to be *upper semicontinuous*.

A function  $f : \mathbb{D} \mapsto \mathbb{R}$  is *lower semicontinuous* if  $-f$  is upper semicontinuous.

Using Condition (ii), it is easy to see that a function which is both upper and lower semicontinuous is also continuous.

The set of all continuous and bounded functions  $f : \mathbb{D} \mapsto \mathbb{R}$ , which we denote  $C_b(\mathbb{D})$ , plays an important role in weak convergence on the metric space  $\mathbb{D}$ , which we will explore in Chapter 7.

It is not hard to see that the Borel  $\sigma$ -field on a metric space  $\mathbb{D}$  is the smallest  $\sigma$ -field generated by the open balls.

It turns out that the Borel  $\sigma$ -field  $\mathcal{B}$  of  $\mathbb{D}$  is also the smallest  $\sigma$ -field  $\mathcal{A}$  making all of  $C_b(\mathbb{D})$  measurable; To see this,

- Note that any closed  $A \subset \mathbb{D}$  is the preimage of the closed set  $\{0\}$  for the continuous bounded function  $x \mapsto d(x, A) \wedge 1$ , where for any set  $B \subset \mathbb{D}$ ,  $d(x, B) \equiv \inf\{d(x, y) : y \in B\}$ .
- Thus  $\mathcal{B} \subset \mathcal{A}$ .
- Since it is obvious that  $\mathcal{A} \subset \mathcal{B}$ , we now have  $\mathcal{A} = \mathcal{B}$ .  $\square$



A Borel-measurable map  $X : \Omega \mapsto \mathbb{D}$  defined on a probability space  $(\Omega, \mathcal{A}, P)$  is called a *random element* or *random map* (or *random variable*) with values in  $\mathbb{D}$ .

Borel measurability is, in many ways, the natural concept to use on metric spaces since it connects nicely with the topological structure.

A *Cauchy sequence* is a sequence  $\{x_n\}$  in a semimetric space  $(\mathbb{D}, d)$  such that  $d(x_n, x_m) \rightarrow 0$  as  $n, m \rightarrow \infty$ .

A semimetric space  $\mathbb{D}$  is *complete* if every Cauchy sequence has a limit  $x \in \mathbb{D}$ .

Every metric space  $\mathbb{D}$  has a completion  $\overline{\mathbb{D}}$  which has a dense subset *isometric* with  $\mathbb{D}$ .

Two metric spaces are *isometric* if there exists a *bijection* (a one-to-one and onto map) between them which preserves distances.

When a metric space  $\mathbb{D}$  is separable, and therefore has a countable dense subset, the Borel  $\sigma$ -field for  $\mathbb{D}$  is itself a separable  $\sigma$ -field.

To see this, let  $A \in \mathbb{D}$  be a countable dense subset and consider the collection of open balls with centers at points in  $A$  and with rational radii.

Clearly, the set of such balls is countable and generates all open sets in  $\mathbb{D}$ :

- Let  $x_0 \in \mathbb{D}$ ,  $\{x_n\} \in A: x_n \rightarrow x_0$ , and  $\epsilon_n = |x_n - x_0|$ .

- It is not hard to see that

$$\bigcup_{n \geq 1} \{x : |x - x_n| < (\eta - \epsilon_n) \vee 0\} = B_\eta(x_0) \text{ for any } \eta > 0.$$

A topological space  $X$  is *Polish* if it is separable and if there exists a metric making  $X$  into a complete metric space.

Hence any complete and separable metric space is Polish.

Furthermore, any open subset of a Polish space is also Polish.

Examples of Polish spaces include Euclidean spaces and many other interesting spaces that we will explore shortly.

A subset  $K$  is *totally bounded* if and only if for every  $r > 0$ ,  $K$  can be covered by finitely many open  $r$ -balls.

Furthermore, a subset  $K$  of a complete semimetric space is compact if and only if it is totally bounded and closed.

A totally bounded subset  $K$  is also called *precompact* because every sequence in  $K$  has a Cauchy subsequence.

This relationship between compactness and total boundedness implies that a  $\sigma$ -compact set in a metric space is separable.

These definitions of compactness agree with the previously given compactness properties for Hausdorff spaces.

This happens because a semimetric space  $\mathbb{D}$  can be made into a metric—and hence Hausdorff—space  $\mathbb{D}_H$  by equating points in  $\mathbb{D}_H$  with equivalence classes in  $\mathbb{D}$ .

# Banach Spaces

A very important example of a metric space is a *normed space*.

A normed space  $\mathbb{D}$  is a vector space (also called a linear space) equipped with a *norm*, and a norm is a map  $\|\cdot\| : \mathbb{D} \mapsto [0, \infty)$  such that, for all  $x, y \in \mathbb{D}$  and  $\alpha \in \mathbb{R}$ ,

- (i)  $\|x + y\| \leq \|x\| + \|y\|$  (another triangle inequality);
- (ii)  $\|\alpha x\| = |\alpha| \times \|x\|$ ;
- (iii)  $\|x\| = 0$  if and only if  $x = 0$ .

A *seminorm* satisfies (i) and (ii) but not necessarily (iii).

A normed space is a metric space (and a seminormed space is a semimetric space) with  $d(x, y) = \|x - y\|$ , for all  $x, y \in \mathbb{D}$ .

A complete normed space is called a *Banach space*.

Two seminorms  $\|\cdot\|_1$  and  $\|\cdot\|_2$  on a set  $\mathbb{D}$  are equivalent if the following is true for all  $x, \{x_n\} \in \mathbb{D}$ :

$$\|x_n - x\|_1 \rightarrow 0 \text{ if and only if } \|x_n - x\|_2 \rightarrow 0.$$



In our definition of a normed space  $\mathbb{D}$ , we require the space to also be a vector space (and therefore it contains all linear combinations of elements in  $\mathbb{D}$ ).

However, it is sometimes of interest to apply norms to subsets  $K \subset \mathbb{D}$  which may not be linear subspaces.

In this setting, let  $\text{lin } K$  denote the *linear span of  $K$*  (all linear combinations of elements in  $K$ ), and let  $\overline{\text{lin } K}$  the closure of  $\text{lin } K$ .

Note that both  $\text{lin } K$  and  $\overline{\text{lin } K}$  are now vector spaces and that  $\overline{\text{lin } K}$  is also a Banach space.

We now present several specific examples of metric spaces.//[0.5ex]

The Euclidean space  $\mathbb{R}^d$  is a Banach space with squared norm

$$\|x\|^2 = \sum_{j=1}^d x_j^2.$$

This space is equivalent under several other norms, including

$$\|x\| = \max_{1 \leq j \leq d} |x_j| \text{ and } \|x\| = \sum_{j=1}^d |x_j|.$$

A Euclidean space is separable with a countably dense subset consisting of all vectors with rational coordinates.

By the Heine-Borel theorem, a subset in a Euclidean space is compact if and only if it is closed and bounded.

The Borel  $\sigma$ -field is generated by the intervals of the type  $(-\infty, x]$ , for rational  $x$ , where the interval is defined as follows:  $y \in (-\infty, x]$  if and only if  $y_j \in (-\infty, x_j]$  for all coordinates  $j = 1, \dots, d$ .

For one-dimensional Euclidean space,  $\mathbb{R}$ , the norm is  $\|x\| = |x|$  (absolute value).

The extended real line  $\bar{\mathbb{R}} = [-\infty, \infty]$  is a metric space with respect to the metric  $d(x, y) = |G(x) - G(y)|$ , where  $G : \bar{\mathbb{R}} \mapsto \mathbb{R}$  is any strictly monotone increasing, continuous and bounded function, such as the arctan function.

For any sequence  $\{x_n\} \in \bar{\mathbb{R}}$ ,  $|x_n - x| \rightarrow 0$  implies  $d(x_n, x) \rightarrow 0$ , while divergence of  $d(x_n, x)$  implies divergence of  $|x_n - x|$ .

In addition, it is possible for a sequence to converge, with respect to  $d$ , to either  $-\infty$  or  $\infty$ .

This makes  $\bar{\mathbb{R}}$  compact.

Another important example is the set of bounded real functions

$f : T \mapsto \mathbb{R}$ , where  $T$  is an arbitrary set.

This is a vector space if sums  $z_1 + z_2$  and products with scalars,  $\alpha z$ , are defined pointwise for all  $z, z_1, z_2 \in \ell^\infty(T)$ .

Specifically,  $(z_1 + z_2)(t) = z_1(t) + z_2(t)$  and  $(\alpha z)(t) = \alpha z(t)$ , for all  $t \in T$ .

This space is denoted  $\ell^\infty(T)$ .

The *uniform norm*  $\|x\|_T \equiv \sup_{t \in T} |x(t)|$  makes  $\ell^\infty(T)$  into a Banach space consisting exactly of all functions  $z : T \mapsto \mathbb{R}$  satisfying  $\|z\|_T < \infty$ .

Note that  $\ell^\infty(T)$  is separable if and only if  $T$  is countable.

Two useful subspaces of  $\ell^\infty([a, b])$ , where  $a, b \in \bar{\mathbb{R}}$ , are  $C[a, b]$  and  $D[a, b]$ .

The space  $C[a, b]$  consists of continuous functions  $z : [a, b] \mapsto \mathbb{R}$ , and  $D[a, b]$  is the space of *cadlag* functions which are right-continuous with left-hand limits (cadlag is an abbreviation for continue à droite, limites à gauche).

We usually equip these spaces with the uniform norm  $\| \cdot \|_{[a, b]}$  inherited from  $\ell^\infty([a, b])$ .

Note that  $C[a, b] \subset D[a, b] \subset \ell^\infty([a, b])$ .

Relative to the uniform norm,  $C[a, b]$  is separable, and thus also Polish by the completeness established in Exercise 6.5.5(a), but  $D[a, b]$  is not separable.

Sometimes,  $D[a, b]$  is called the *Skorohod space*, although Skorohod equipped  $D[a, b]$  with a special metric—quite different than the uniform metric—resulting in a separable space.



# Arzelà-Ascoli Theorem

An important subspace of  $\ell^\infty(T)$  is the space  $UC(T, \rho)$ , where  $\rho$  is a semimetric on  $T$ .

$UC(T, \rho)$  consists of all bounded function  $f : T \mapsto \mathbb{R}$  which are uniformly  $\rho$ -continuous, i.e.,

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

When  $(T, \rho)$  is totally bounded, the boundedness requirement for functions in  $UC(T, \rho)$  is superfluous since a uniformly continuous function on a totally bounded set must necessarily be bounded.

We denote  $C(T, \rho)$  to be the space of  $\rho$ -continuous (not necessarily continuous) function on  $T$ .

The spaces  $C[a, b]$ ,  $D[a, b]$ ,  $UC(T, \rho)$ ,  $C(\overline{T}, \rho)$ , when  $(T, \rho)$  is a totally bounded semimetric space, and  $UC(T, \rho)$  and  $\ell^\infty(T)$ , for an arbitrary set  $T$ , are all complete with respect to the uniform metric.

When  $(T, \rho)$  is a compact semimetric space,  $T$  is totally bounded, and a  $\rho$ -continuous function in  $T$  is automatically uniformly  $\rho$ -continuous.

Thus, when  $T$  is compact,  $C(T, \rho) = UC(T, \rho)$ .

Actually, every space  $UC(T, \rho)$  is equivalent to a space  $C(\overline{T}, \rho)$ , because the completion  $\overline{T}$  of a totally bounded space  $T$  is compact and, furthermore, every uniformly continuous function on  $T$  has a unique continuous extension to  $\overline{T}$ .

The foregoing structure makes it clear that  $UC(T, \rho)$  is a Polish space that is made complete by the uniform norm (and hence is also separable).

Moreover, all compact sets in  $\ell^\infty(T)$  have a specific form:

THEOREM 1. (*Arzelà-Ascoli*)

(a) *The closure of  $K \subset \ell^\infty(T)$  is compact if and only if*

(i)  $\sup_{x \in K} |x(t_0)| < \infty$ , for some  $t_0 \in T$ ; and

(ii) *for some semimetric  $\rho$  making  $T$  totally bounded,*

$$\lim_{\delta \downarrow 0} \sup_{x \in K} \sup_{s, t \in T: \rho(s, t) < \delta} |x(s) - x(t)| = 0.$$

(b) *The set  $K \subset \ell^\infty(T)$  is  $\sigma$ -compact if and only if  $K \subset LC(T, \rho)$  for some semimetric  $\rho$  making  $T$  totally bounded, where  $LC(T, \rho)$  is the subset of  $UC(T, \rho)$  consisting of all bounded functions  $x$  with*

$$\sup_{\delta > 0} \sup_{s, t \in T: \rho(s, t) < \delta} \frac{|x(s) - x(t)|}{\delta} < \infty.$$

(c) *The closure of  $K \subset \ell^\infty(T)$  is separable if and only if  $K \subset UC(T, \rho)$  for some semimetric  $\rho$  making  $T$  totally bounded.*

Since all compact sets are trivially  $\sigma$ -compact, Theorem 1 implies that any compact set in  $\ell^\infty(T)$  is actually contained in  $LC(T, \rho)$  for some semimetric  $\rho$  making  $T$  totally bounded.

Another important class of metric spaces are product spaces.

For a pair of metric spaces  $(\mathbb{D}, d)$  and  $(\mathbb{E}, e)$ , the *Cartesian product*  $\mathbb{D} \times \mathbb{E}$  is a metric space with respect to the metric

$$\rho((x_1, y_1), (x_2, y_2)) \equiv d(x_1, x_2) \vee e(y_1, y_2),$$

for  $x_1, x_2 \in \mathbb{D}$  and  $y_1, y_2 \in \mathbb{E}$ .

The resulting topology is the *product topology*.

In this setting, convergence of  $(x_n, y_n) \rightarrow (x, y)$  is equivalent to convergence of both  $x_n \rightarrow x$  and  $y_n \rightarrow y$ .

There are two natural  $\sigma$ -fields for  $\mathbb{D} \times \mathbb{E}$  that we can consider.

The first is the Borel  $\sigma$ -field for  $\mathbb{D} \times \mathbb{E}$  generated from the product topology.

The second is the product  $\sigma$ -field generated by all sets of the form  $A \times B$ , where  $A \in \mathcal{A}$ ,  $B \in \mathcal{B}$ , and  $\mathcal{A}$  and  $\mathcal{B}$  are the respective  $\sigma$ -fields for  $\mathbb{D}$  and  $\mathbb{E}$ .

These two are equal when  $\mathbb{D}$  and  $\mathbb{E}$  are separable, but they may be unequal otherwise, with the first  $\sigma$ -field larger than the second.

Suppose  $X : \Omega \mapsto \mathbb{D}$  and  $Y : \Omega \mapsto \mathbb{E}$  are Borel-measurable maps defined on a measurable space  $(\Omega, \mathcal{A})$ .

Then  $(X, Y) : \Omega \mapsto \mathbb{D} \times \mathbb{E}$  is a measurable map for the product of the two  $\sigma$ -fields by the definition of a measurable map.

Unfortunately, when the Borel  $\sigma$ -field for  $\mathbb{D} \times \mathbb{E}$  is larger than the product  $\sigma$ -field, then it is possible for  $(X, Y)$  to not be Borel-measurable.