Notes on Polish space

Rui Wang

Monday 27th May, 2019

1 Introduction

This document contains notes about Polish space which play an important role in probability and statistics. The materials are mainly from Cohn (2013), Chapter 8 and Dudley (2002), Chapter 13.

2 Polish space

Exercise 1 (Cohn (2013), Exercise 8.1.3)**.** *Let $(X, \mathscr{A})$ be a measurable space, let Y be a separable metrizable space, and let $f, g : X \to Y$ be measurable with respect to $\mathscr{A}$ and $\mathscr{B}(Y)$. Then $\{x \in X : f(x) = g(x)\} \in \mathscr{A}$.*

Proof. For any $A, B \in \mathscr{B}(Y)$,

$$\{x : (f(x), g(x)) \in A \times B\} = f^{-1}(A) \cap f^{-1}(B) \in \mathscr{A}.$$

Hence the map $F : x \mapsto (f(x), f(x))$ is measurable with respect to $\mathscr{A}$ and $\mathscr{B}(Y) \times \mathscr{B}(Y)$. Since Y is a separable metrizable space, $\mathscr{B}(Y) \times \mathscr{B}(Y) = \mathscr{B}(Y \times Y)$. Thus, the map F is measurable with respect to $\mathscr{A}$ and $\mathscr{B}(Y \times Y)$. Let $\Delta = \{(y_1, y_2) \in Y \times Y : y_1 = y_2\}$. Then Δ is a closed subset of $Y \times Y$ and $\{x \in X : f(x) = g(x)\} = F^{-1}(\Delta)$. It follows that $\{x \in X : f(x) = g(x)\} \in \mathscr{A}$. □

Exercise 2 (Cohn (2013), Exercise 8.2.1)**.** *Let A be an uncountable analytic subset of the Polish space X. Then,*

(a) A has a subset that is homeomorphic to $\{0, 1\}^{\mathbb{N}}$.

(b) A has the cardinality of the continuum.

Proof. From Cohn (2013), Corollary 8.2.8., there is a continuous function f from $\mathscr{N}$ onto A. By the axiom of choice, there is a set $S \subset \mathscr{N}$ such that the restriction of f on S is a bijection of S onto A. As a subspace of $\mathscr{N}$, S is an uncountable separable metrizable space. Let $S_0 \subset S$ be the set of all condensation points of the space S. From Cohn (2013), Lemma 8.2.12, S_0 is uncountable and each point of S_0 is a condensation point of S_0. Let $d_{\mathscr{N}}(\cdot,\cdot)$ be a metric on $\mathscr{N}$ which metrize the topology of $\mathscr{N}$. Let $d_X(\cdot,\cdot)$ be a metric on X which metrize the topology of X.

Now we construct a homeomorphic between a subset of X and $\{0,1\}^{\mathbb{N}}$. First, let x_0 and x_1 be two distinct points in S_0. Since the restriction of f on S_0 is injective, $f(x_0) \neq f(x_1)$. Hence there exists $0 < \epsilon_1 < 1$ such that $\overline{B(x_0,\epsilon_1)} \cap \overline{B(x_1,\epsilon_1)} = \emptyset$ and $f(\overline{B(x_0,\epsilon_1)}) \cap f(\overline{B(x_1,\epsilon_1)}) = \emptyset$. For $i = 0,1$, let $C(i) = B(x_i,\epsilon_1)$. Note that for $i = 0,1$, $C(i) \cap S_0$ is uncountable and each point of $C(i) \cap S_0$ is a condensation point of $C(i) \cap S_0$. Then there exist x_{i0}, $x_{i1} \in C(i) \cap S_0$ $(i = 0,1)$ and $0 < \epsilon_2 < 1/2$ such that for $j = 0,1$, $B(x_{ij},\epsilon_2) \subset B(x_i,\epsilon_1)$, $\overline{B(x_{i0},\epsilon_2)} \cap \overline{B(x_{i1},\epsilon_2)} = \emptyset$ and $f(\overline{B(x_{i0},\epsilon_2)}) \cap f(\overline{B(x_{i1},\epsilon_2)}) = \emptyset$. For $i,j \in \{0,1\}$, let $C(i,j) = B(x_{ij},\epsilon_2)$.

Inductively construct sets $C(n_1,n_2,\ldots,n_k)$, $n_i \in \{0,1\}$, $k \in \mathbb{N}$. Then for $\{n_k\}_{k=1}^{\infty} \in \mathscr{N}$, consider the set $\cap_{k=1}^{\infty}\overline{C(n_1,\ldots,n_k)}$. By the completeness of $\mathscr{N}$, $\cap_{k=1}^{\infty}\overline{C(n_1,\ldots,n_k)} \neq \emptyset$. Also, the diameter of $\overline{C(n_1,\ldots,n_k)}$ tends to 0. Then there exists a unique point in $\cap_{k=1}^{\infty}\overline{C(n_1,\ldots,n_k)}$. Let g be the function from $\mathscr{N}$ to X which maps $\{n_k\}_{k=1}^{\infty}$ to the unique point of $\cap_{k=1}^{\infty}\overline{C(n_1,\ldots,n_k)}$.

By the construction of $C(n_1,\ldots,n_k)$, g is continuous and injective. Then $f \circ g$ is continuous. To see that $f \circ g$ is injective, let $\{n_k\}_{k=1}^{\infty}$ and $\{m_k\}_{k=1}^{\infty}$ be two distinct points of $\{0,1\}^{\mathscr{N}}$. Let k_0 be the first k such that $n_k \neq m_k$. By the construction of $C(\cdot,\ldots,\cdot)$, $f(\overline{C(n_1,\ldots,n_{k_0})}) \cap f(\overline{C(m_1,\ldots,m_{k_0})}) = \emptyset$. Since $g(\{n_k\}_{k=1}^{\infty}) \subset \overline{C(n_1,\ldots,n_{k_0})}$, $g(\{m_k\}_{k=1}^{\infty}) \subset \overline{C(m_1,\ldots,m_{k_0})}$. Then $f \circ g(\{n_k\}_{k=1}^{\infty}) \neq f \circ g(\{m_k\}_{k=1}^{\infty})$.

Since $\{0,1\}^{\mathscr{N}}$ is compact, the inverse of $f \circ g$ is also continuous. This completes the proof of (a).

(a) implies that $\text{card}(A) \geq \mathfrak{c}$. On the other hand, Cohn (2013), Corollary 8.2.8. implies that $\text{card}(A) \leq \mathfrak{c}$. Thus, $\text{card}(A) = \mathfrak{c}$.

□

Exercise 3 (Cohn (2013), Exercise 8.2.2)**.** *Let X be an uncoutable Polish space. Then the collection of analytic subsets of X and the collection of Borel subsets of X have the cardinality of the continuum.*

Proof. Exercise 2 implies that the cardinality of X is $\mathfrak{c}$. Since each single point of X is a Borel set, the cardinality of the collection of Borel subsets of X is at least $\mathfrak{c}$. We only need to prove that the cardinality of the collection of analytic subsets of X is at most $\mathfrak{c}$.

Cohn (2013), Proposition 8.2.9 implies that it suffices to upper bound the cardinality of the collection of closed subsets of the Polish space $\mathscr{N} \times X$. Let $\{U_i\}_{i=1}^{\infty}$ be a countable base of the topology of $\mathscr{N} \times X$. Then every closed subset of $\mathscr{N} \times X$ is the intersection of certain $U_i^{\complement}$, that is, $\cap_{i\in S}U_i^{\complement}$ where S is a subset of $\mathbb{N}$. Hence there is an injective map from the collection of closed subsets of $\mathscr{N} \times X$ to $2^{\mathbb{N}}$. Thus, the cardinality of the collection of closed subsets of $\mathscr{N} \times X$ is at most $\mathfrak{c}$.

□

Exercise 4 (Cohn (2013), Exercise 8.2.3)**.**

(a) Let X be a nonempty zero-dimensional Polish space such that each nonempty open subset of X is not compact. Then X is homeomorphic to $\mathscr{N}$.

(b) the Space $\mathscr{I}$ of irrational numbers in the interval $(0, 1)$ is homeomorphic to $\mathscr{N}$.

Proof. Let $d(\cdot, \cdot)$ be a complete metric for X. We begin by constructing a family $\{C(n_1, \ldots, n_k)\}$ of subsets of X, indexed by the set of all finite sequences $\{(n_1, \ldots, n_k)\}$ of positive integers, in such a way that

1. $C(n_1, \ldots, n_k)$ is nonempty, open, closed and noncompact,
2. the diameter of $C(n_1, \ldots, n_k)$ is at most $1/k$,
3. $\{C(n_1, \ldots, n_{k-1}, n_k)\}_{n_k=1}^{\infty}$ are disjoint and $C(n_1, \ldots, n_{k-1}) = \cup_{n_k=1}^{\infty} C(n_1, \ldots, n_k)$,
4. $X = \cup_{n_1=1}^{\infty} C(n_1)$.

We do this by induction on k.

First, suppose that $k = 1$. Since X is assumed to be not compact, Cohn (2013), Lemma 8.2.11 gives a sequnce $\{C(n_1)\}_{n_1=1}^{\infty}$ where terms are nonempty, open, closed and with diameter at most 1. By assumption, each $C(n_1)$ is not compact.

Now suppose that $k > 1$ and that $C(n_1, \ldots, n_{k-1})$ has already been chosen. It is easy to use a modification of the construction of the $C(n_1)$'s, now applied to $C(n_1, \ldots, n_{k-1})$ rather than to X, to produce sets $C(n_1, \ldots, n_k)$, $n_k = 1, 2, \ldots$ that satisfy conditions 1 to 4. With this, the induction step in our construction is complete.

We turn to the construction of a homeomorphic between $\mathscr{N}$ and X. Let $\mathbf{n} = \{n_k\}$ be an element of $\mathscr{N}$. Then the sets $C(n_1)$, $C(n_1, n_2)$, $\ldots$ are decreasing nonempty closed sets whose diameters approach to 0. Since X is complete, there is a unique element in $\cap_{k=1}^{\infty} C(n_1, \ldots, n_k)$. We can define a function $f : \mathscr{N} \to X$ by letting $f(\mathbf{n})$ be the unique member of $\cap_{k=1}^{\infty} C(n_1, \ldots, n_k)$. Note that if $\mathbf{m}$ and $\mathbf{n}$ are elements of $\mathscr{N}$ such that $m_i = n_i$ holds for $k = 1, \ldots, k$, then $d(\mathbf{m}, \mathbf{n}) \leq 1/k$. It follows that f is continuous. Also, it is obvious that f is bijective. It remain to prove that the inverse of f is continuous. Suppose $f(\mathbf{n}^{(l)}) \to f(\mathbf{n})$. Fix $k > 0$. Then if l is large enough, $f(\mathbf{n}^{(l)}) \in C(n_1, \ldots, n_k)$. By the construction of f, this implies that $n_i^{(l)} = n_i$ for $i = 1, \ldots, k$. Thus, $\mathbf{n}^{(l)} \to \mathbf{n}$ as $l \to \infty$. This completes the proof of (a).

We turn to the proof of (b). The space $\mathscr{I}$ is a G_δ set of $[0, 1]$, and hence is a Polish space. The family of intervels (a_i, b_i) where a_i and b_i is rational is a base that consists of sets that are both open and closed. It follows that $\mathscr{I}$ is zero-dimensional. Each interval (a, b) is the union of $\{(a_i, b_i)\}_{i=1}^{\infty}$ where a_i, b_i are rational and $a_i \downarrow a$ and $b_i \uparrow b$. Hence each interval of $\mathscr{I}$ is not compact. Then the conclusion follows from (a).

$\square$

Exercise 5 (Cohn (2013), Exercise 8.2.3)**.** *Each nonempty Polish space is the image of $\mathscr{N}$ under a continuous open map.*

Proof. We mimic the proof of Cohn (2013), Proposition 8.2.7.

Let X be a nonempty Polish space, and let d be a complete metric for X. We begin by constructing a family $\{C(n_1, \ldots, n_k)\}$ of subsets of X, indexed by the set of all finite sequences $\{n_1, \ldots, n_k\}$ of positive integers, in such a way that

1. $C(n_1, \ldots, n_k)$ is nonempty and open,
2. the diameter of $C(n_1, \ldots, n_k)$ is at most $1/k$,
3. $\overline{C(n_1, \ldots, n_{k-1}, n_k)} \subset C(n_1, \ldots, n_{k-1})$ and $C(n_1, \ldots, n_{k-1}) = \cup_{n_k=1}^{\infty} C(n_1, \ldots, n_k)$,
4. $X = \cup_{n_1=1}^{\infty} C(n_1)$.

We do this by induction on k.

First, suppose that $k = 1$, and let $\{x_i\}_{i=1}^{\infty}$ be a sequence whose terms form a dense subset of X. The sequence $\{X_i\}_{i=1}^{\infty}$ may have duplicated elements. Let $\{C(n_1)\}_{n1=1}^{\infty}$ be the collection of open balls which center at certain x_i and with rational radius not larger than $1/2$. Certainly each $C(n_1)$ is open and nonempty and has diameter at most 1. Furthermore, $X = \cup_{n_1} C(n_1)$.

Now suppose that $k > 1$ and that $C(n_1, \ldots, n_{k-1})$ has already been chosen. Let $\{C\,(n_1, \ldots, n_{k-1}, n_k)\}_{n_k=1}^{\infty}$ be the collection of open balls which center at centain x_i and with rational radius not larger than $1/(2k)$ and whose closure is contained in $C(n_1, \ldots, n_{k-1})$. Certainly each $C(n_1, \ldots, n_k)$ is open and nonempty and has diameter at most $1/k$. Now we prove that $C(n_1, \ldots, n_{k-1}) = \cup_{n_k=1}^{\infty} C(n_1, \ldots, n_k)$. Suppose $x \in C(n_1, \ldots, n_{k-1})$. Since $C(n_1, \ldots, n_{k-1})$ is open, there is a open ball $B(x, r) \subset C(n_1, \ldots, n_{k-1})$ where r is rational and $r < 1/k$. Since $\{x_i\}_{i=1}^{\infty}$ is dense in X, there is an x_i such that $d(x, x_i) < r/3$. Then the ball $B(x_i, r/2)$ contains x. Also, the Closure of $B(x_i, r/2)$ has radius not larger than $1/(2k)$ and is contained in $C(n_1, \ldots, n_{k-1})$. Thus, $B(x_i, r/2) = C(n_1, \ldots, n_k)$ for some n_k. With this, the induction step in our construction is complete.

We turn to the construction of a continuous function that maps $\mathscr{N}$ onto X. Let $\mathbf{n} = \{n_k\}$ be an element of $\mathscr{N}$. It follows from 3 that $\cap_{k=1}^{\infty} C(n_1, \ldots, n_k) = \cap_{k=1}^{\infty} \overline{C(n_1, \ldots, n_k)}$ which is intersection of a decreasing sequence of nonempty closd subsets of X whose diameters approach 0. Thus there is a unique element in the intersection of these sets, and we can define a function $f : \mathscr{N} \to X$ by letting $f(\mathbf{n})$ be the unique member of $\cap_k C(n_1, \ldots, n_k)$. Note that if $\mathbf{m}$ and $\mathbf{n}$ are elements on $\mathscr{N}$ such that $m_i = n_i$ holds for $i = 1, \ldots, k$, then $d(f(\mathbf{m}, \mathbf{n})) \leq 1/k$. It follows that f is continuous. Also, 3 and 4 above imply that for each x in X there is an element $\mathbf{n} = \{n_k\}$ of $\mathscr{N}$ such that $x \in \cap_k C(n_1, \ldots, n_k)$ and hence such that $x = f(\mathbf{n})$. Thus f is surjective.

It remains to prove that f is an open map. Note that the sets of the form $\{n_1\} \times \cdots \times \{n_k\} \times \mathbb{N} \times \ldots$ is a base for the topology of $\mathscr{N}$. By the construction of f, for any $n_1, \ldots, n_k$, $f(\{n_1\} \times \cdots \times \{n_k\} \times \mathbb{N} \times \ldots) = C(n_1, \ldots, n_k)$ is an open set. This completes the proof.

$\square$

Exercise 6 (Cohn (2013), Exercise 8.2.5)**.** *Each Borel subet of a Polish space is the image under a continuous injective map of some Polish space.*

Proof. Let X be a Polish space. Let $\mathcal{A}$ be the collection of Borel subsets of X which are the image under continuous injective maps of some Polish spaces. Then all open and closed subsets of X belong to $\mathcal{A}$ since they are themselves Polish spaces.

Assume $A_1, \ldots, A_n, \cdots \in \mathcal{A}$ and $A_1, \ldots, A_n, \ldots$ are disjoint. For each A_i, there is a Polish space X_i and a continuous infective map $f_i(\cdot)$ such that $f_i(X_i) = A_i$. Define $f : \cup_{i=1}^{\infty} X_i \mapsto \cup_{i=1}^{\infty} A_i$ by $f(x) = f_i(x)$ if $x \in X_i$. Here $\cup_{i=1}^{\infty} X_i$ is the disjoint union of X_i. Then $\cup_{i=1}^{\infty} X_i$ is Polish and f is injective and continuous. Then $\cup_{i=1}^{\infty} A_i \in \mathcal{A}$.

Assume $A_1, \ldots, A_n, \cdots \in \mathcal{A}$. For each A_i, there is a Polish space X_i and a continuous infective map $f_i(\cdot)$ such that $f_i(X_i) = A_i$. Define $f : \prod_{i=1}^{\infty} X_i \mapsto \prod_{i=1}^{\infty} X$ by $f(\{x_i\}_{i=1}^{\infty}) = \{f_i(x_i)\}_{i=1}^{\infty}$. Then f is injective and continuous onto $\prod_{i=1}^{\infty} A_i \subset \prod_{i=1}^{\infty} X$. Let $D = \{(x, x, \ldots) : x \in X\}$. Define $g : D \mapsto X$ by $g(x, x, \ldots) = x$. Then g is a homeomorphism between D and X. Consider $g \circ f$ defined on $f^{-1}(D)$. Then $g \circ f$ is injective and continuous from $f^{-1}(D)$ onto $\cap_{i=1}^{\infty} A_i$. Since $f^{-1}(D)$ is a closed subset of $\prod_{i=1}^{\infty} X_i$, it is Polish. Thus, $\cap_{i=1}^{\infty} A_i \in \mathcal{A}$.

From Cohn (2013), Lemma 8.2.4, $\mathcal{A}$ contains all Borel subset of X. This completes the proof. $\square$

Exercise 7 (Cohn (2013), Exercise 8.2.6)**.** *If X is an uncountable Polish space, then there is an analytic subset of X that is not a Borel set.*

Proof. Let X be an uncountable Polish space. From Cohn (2013), Proposition 8.2.13, there is a continuous injective map $f : \mathscr{N} \to X$ such that $X - f(\mathscr{N})$ is countable. From Cohn (2013), Corollary 8.2.17, there is an analytic set $A \in \mathscr{N}$ that is not a Borel set. Then $f(A)$ is not a Borel set of X, or else $A = f^{-1}(f(A))$ would be a Borel set, a contradiction. On the other hand, $f(A)$ is analytic. This completes the proof.

$\square$

Exercise 8 (Cohn (2013), Exercise 8.3.1)**.** *Let X and Y be Polish spaces, and let $f : X \to Y$ be a function whose graph is an analytic subset of $X \times Y$. Then f is Borel measurable.*

Remark 1. It follows from this conclution and Cohn (2013), Proposition 8.1.8 that f is Borel measurable iif the graph of f is a Borel subset of $X \times Y$. Then the graph of f can not be an analytic set which is not a Borel set.

Proof. Let $G = \{(x, f(x)) : x \in X\}$ denote the graph of f. For any Borel subset B of Y, the sets $G \cap (X \times B)$ and $G \cap (X \times B^{\complement})$ are analytic. Then the projection of these two sets on X, i.e. $f^{-1}(B)$ and $f^{-1}(B^{\complement})$, are also analytic. From separation theorem, i.e. Cohn (2013), Theorem 8.3.1, B and $B^{\complement}$ are Borel sets. Hence f is Borel measurable. $\square$

Exercise 9 (Cohn (2013), Exercise 8.3.2). *Let X and Y be uncountable Polish spaces. Then the cardinality of the collection of Borel measurable functions from X to Y is that of the continuum.*

Proof. The cardinalities of X and Y are both $\mathfrak{c}$. For each $y \in Y$, the constant function $f(x) \equiv y$ is Borel measurable. Hence the cardinality of the collection of Borel measurable functions from X to Y is at least $\mathfrak{c}$.

On the other hand, since the graph of Borel measurable function f is a Borel subset of $X \times Y$, and the collection of Borel subsets of an uncontable Polish space has cardinality $\mathfrak{c}$, the cardinality of the collection of Borel measurable functions from X to Y is as most $\mathfrak{c}$. This completes the proof. □

Exercise 10 (Cohn (2013), Exercise 8.3.2). *There is a Lebesgue measurable function $f : \mathbb{R} \to \mathbb{R}$ such that no real-valued Borel measurable function f_1 satisfies $f(x) \leq f_1(x)$ at each x in $\mathbb{R}$.*

Proof. Let K be the Cantor set. K is an uncontable Polish space. According to the preceding ecercise, there is a bijection $x \mapsto g_x$ of K onto the set of real-valued Borel functions on K. Define $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} g_x(x) + 1 & \text{if } x \in K, \\ 0 & \text{otherwise.} \end{cases}$$

Since $f(x) = 0$ a.e., f is Lebesgue measurable. Suppose there is a real-valued Borel measurable function f_1 satisfying $f(x) \leq f_1(x)$ at each $x \in \mathbb{R}$. Then the restriction of f_1 on K is still Borel measurable. Hence there is an $x_1 \in K$ such that $f_1(x) = g_{x_1}(x)$ at each $x \in K$. Then $g_{x_1}(x) \geq g_x(x) + 1$ at each $x \in K$. But this is impossible when $x = x_1$. This completes the proof. □

Exercise 11 (Cohn (2013), Exercise 8.3.3). *Let X be a Polish space, let μ be a Borel measure on X such that $\mu(X) = 1$, and let λ be Lebesgue measure on the Borel subsets of $[0, 1]$. Then there is a Borel measurable function $f : [0, 1] \to X$ such that $\mu = \lambda f^{-1}$.*

Proof. If X is countably infinite, let $\{x_i\}_{i=1}^{\infty}$ be an enumeration of the points of X. Then $\sum_{i=1}^{\infty} \mu(\{x_i\}) = \mu(\cup_{i=1}^{\infty}\{x_i\}) = 1$. We can construct f by letting $f(t) = x_i$ if $t \in \left[\sum_{j=1}^{i-1} x_j, \sum_{j=1}^{i} x_j\right)$, and $f(1) = x_1$. Then $\lambda f^{-1}(\{x_i\}) = \lambda\left([\sum_{j=1}^{i-1} x_j, \sum_{j=1}^{i} x_j)\right) = x_i$. Hence $\mu = \lambda f^{-1}$. If X is finite, the construction of f is similar.

Now suppose X is uncountable, from Cohn (2013), Theorem 8.3.6, there is a bijection g from $\mathbb{R}$ onto X which is Borel isomorphism. Then the measure μg is a probability measure on the Borel sets of $\mathbb{R}$. From Cohn (2013), Proposition 10.1.15, there is a Borel measurable function h from $[0, 1]$ to $\mathbb{R}$ such that $\mu g = \lambda h^{-1}$. Thus, for any $A \in \mathcal{B}(X)$,

$$\mu(A) = \mu(gg^{-1}A) = \mu g(g^{-1}A) = \lambda h^{-1}(g^{-1}A) = \lambda(h^{-1}g^{-1}A) = \lambda((g \circ h)^{-1}A) = \lambda(g \circ h)^{-1}(A).$$

The conclusion follows by letting $f = g \circ h$. □

Exercise 12 (Cohn (2013), Exercise 8.4.1)**.** *Let $(X, \mathscr{A})$ be a measurable space.*

(a) A function $f : X \to [-\infty, +\infty]$ is $\mathscr{A}_*$-measurable if and only if for each finite measure μ on $(X, \mathscr{A})$ there are $\mathscr{A}$-measurable functions $f_0, f_1 : X \to [-\infty, +\infty]$ that satisfy $f_0 \leq f \leq f_1$ everywhere on X and $f_0 = f_1$ μ-almost surely.

(b) If $f : X \to [-\infty, +\infty]$ is $\mathscr{A}_*$-measurable and if the functions f_0 and f_1 in part (a) can be chosen independently of μ, then f is $\mathscr{A}$-measurable.

Proof. It is understood that $\mathscr{A}_*$-measurable means $\mathscr{A}_*/\mathscr{B}([-\infty, +\infty])$ measurable.

Suppose $f : X \to [-\infty, +\infty]$ is $\mathscr{A}_*$-measurable. Then for each finite measure μ on $(X, \mathscr{A})$, since $\mathscr{A} \subset \mathscr{A}_\mu$, f is $\mathscr{A}_\mu$ measurable. Then the existence of f_0, f_1 is implied by Cohn (2013), Proposition 2.2.5. Conversely, suppose for each finite measure μ, such f_0, f_1 exist. Then Cohn (2013), Proposition 2.2.5 implies that f is $\mathscr{A}_\mu$ measurable. Thus, for any $A \in \mathscr{B}([-\infty, +\infty])$, $f^{-1}(A) \in \cap_\mu \mathscr{A}_\mu = \mathscr{A}_*$. This completes the proof of (a).

We turn to (b). Consider the set $A = \{x : f_0(x) \neq f_1(x)\}$. By assumption, $\mu^*(A) = 0$ for any finite μ on $\mathscr{A}$. If A is not empty, let x be a point of A and δ_x be the point mass concentrated on x. Then $\delta_x^*(A) = 1$, a contradiction. It follows that $A = \emptyset$. Thus $f = f_0$, hence f is $\mathscr{A}$-measurable. □

Exercise 13 (Cohn (2013), Exercise 8.4.2)**.** *Let $(X, \mathscr{A})$ be a measurable space. Then*

(a) $(\mathscr{A}_*)_* = \mathscr{A}_*$.

(b) If μ is a finite measure on $(X, \mathscr{A})$, then $(\mathscr{A}_\mu)_* = \mathscr{A}_\mu$.

Proof. First we proof (b). It is clear that $(\mathscr{A}_\mu)_* \supset \mathscr{A}_\mu$. On the other hand, $(\mathscr{A}_\mu)_* \subset (\mathscr{A}_\mu)_\mu = \mathscr{A}_\mu$. This completes the proof of (b).

We turn to (a). It is clear that $(\mathscr{A}_*)_* \supset \mathscr{A}_*$. On the other hand, let μ be any finite measure on $\mathscr{A}$. Then $(\mathscr{A}_*)_* \subset (\mathscr{A}_\mu)_* = \mathscr{A}_\mu$. Hence $(\mathscr{A}_*)_* \subset \cap_\mu \mathscr{A}_\mu = \mathscr{A}_*$. □

The following lemma is useful in proving the next two lemmas.

Lemma 1. *Let $(X, \mathscr{A})$ and $(Y, \mathscr{B})$ are two measurable spaces. Let f be an isomorphism between $(X, \mathscr{A})$ and $(Y, \mathscr{B})$, that is, f is a bijection and f and f^{-1} are both measurable. Then for any finite measure μ on $\mathscr{A}$, f is also an isomorphism between $(X, \mathscr{A}_\mu)$ and $(Y, \mathscr{B}_{\mu f^{-1}})$. Furthermore, f is an isomorphism between $(X, \mathscr{A}_*)$ and $(Y, \mathscr{B}_*)$.*

Proof. Let μ be any finite measure on $\mathscr{A}$. For any $A \in \mathscr{A}_\mu$, there exist $A_0, A_1 \in \mathscr{A}$ such that $A_0 \subset A \subset A_1$ and $\mu(A_1/A_0) = 0$. Then $f(A_0), f(A_1) \in \mathscr{B}$, $f(A_0) \subset f(A) \subset f(A_1)$ and $\mu f^{-1}(f(A_1)/f(A_0)) = \mu(A_1/A_0) = 0$. Then $f(A) \in \mathscr{B}_{\mu f^{-1}}$. Similarly, if $B \in \mathscr{B}_{\mu f^{-1}}$, then $f^{-1}(B) \in \mathscr{A}_\mu$. Thus, f is an isomorphism between $(X, \mathscr{A}_\mu)$ and $(Y, \mathscr{B}_{\mu f^{-1}})$.

This is true for any finite measure μ on $\mathscr{A}$. Hence f is an isomorphism between $(X, \cap_\mu \mathscr{A}_\mu)$ and $(Y, \cap_\mu \mathscr{B}_{\mu f^{-1}})$. By definition, $\cap_\mu \mathscr{A}_\mu = A_*$. On the other hand, for any finite measure ν on $\mathscr{B}$, $\nu(B) = \nu(f \circ f^{-1} B) = \nu f(f^{-1} B) = (\nu f) f^{-1}(B)$. Note that νf is a finite measure on $\mathscr{A}$. It follows that $\cap_\mu \mathscr{B}_{\mu f^{-1}} = \cap_\nu \mathscr{B}_\nu = \mathscr{B}_*$. This complemtes the proof. □

Exercise 14 (Cohn (2013), Exercise 8.4.3). *There is a Lebesgue measurable subset of* $\mathbb{R}$ *that is not universally measurable.*

Proof. Let C be the Cantor set on $[0,1]$ (see Cohn (2013), example 1.4.6), let $f:[0,1]\to[0,1]$ be the Cantor function (see Cohn (2013), example 2.1.10). It is known that f is nondecreasing and continuous.

Let λ be the Lebesgue measure on $\mathbb{R}$. Let $\lambda_{[0,1]}$ be the Lebesgue measure on $[0,1]$. Let μ be the measure on $[0,1]$ with distribution function f. It is known that $\mu(C)=1$. Define $\mu_1=\lambda+\mu$. Then for $x\in[0,1]$, $\mu_1([0,x])=f(x)+x$. Define $g:[0,1]\to[0,2]$ by $g(x)=f(x)+x$. Then g is strict increasing and continuous. Thus, g is a Borel isomorphism from $([0,1],\mathscr{B}([0,1]),\mu_1)$ to $([0,2],\mathscr{B}([0,2]),\lambda)$. Also, $\lambda\circ g=\mu_1$. From Lemma 1, g is also an isomorphism between the completion spaces $([0,1],\mathscr{B}([0,1])_{\mu_1},\mu_1)$ and $([0,2],\mathscr{B}([0,2])_{\lambda},\lambda)$. It can be shown that $\lambda(g(C^{\complement}))=1$. Hence $\lambda(g(C))=2-1=1$. Then There is a Lebesgue **non**measurable subset D of $g(C)$ (see Cohn (2013), exercise 1.4.6). Then $g^{-1}(D)$ is a $\mathscr{B}([0,1])_{\mu_1}$-nonmeasurable subset of C. Note also that $g^{-1}(D)$ is Lebesgue measurable since $g^{-1}(D)$ is a subset of C which has Lebesgue measure 0. Thus, $g^{-1}(D)$ is a Lebesgue measurable subset of $\mathbb{R}$ which is not universally measurable.

□

Exercise 15 (Cohn (2013), Exercise 8.4.4). *Each uncountable Polish space has a subset that is not universally measurable.*

Proof. Let X be an uncountable Polish space. From Cohn (2013), Theorem 8.3.6, there is a Borel isomorphism f between $(X,\mathscr{B}(X))$ and $(\mathbb{R},\mathscr{B}(\mathbb{R}))$. Form Lemma 1, f is an isomorphism between $(X,\mathscr{B}(X)_*)$ and $(\mathbb{R},\mathscr{B}(\mathbb{R})_*)$. But there is an $B\notin\mathscr{B}(\mathbb{R})_*$, e.g., a Lebesgue nonmeasurable set. Then $f^{-1}(B)\notin\mathscr{B}(X)_*$. This complemtes the proof. □

Exercise 16 (Cohn (2013), Exercise 8.4.5). *There is a measurable space* $(X,\mathscr{A})$ *and an outer measure* μ^* *on it such that there exists an increasing sequence* $\{A_n\}$ *of subsets of* X,

$$\mu^*(\cup_n A_n)=\lim_n \mu^*(A_n).$$

Remark 2. An outher measure μ^* on $(X,\mathscr{A})$ is a function from $\mathscr{A}$ to $[0,+\infty]$ such that

(a) $\mu^*(\emptyset)=0$.

(b) if $A\subset B\subset X$, then $\mu^*(A)\le\mu^*(B)$, and

(c) if $\{A_n\}$ is an infinite sequence of subsets of X, then $\mu^*(\cup_n A_n)\le\sum_n\mu^*(A_n)$.

Proof. Let $X=\mathbb{N}$, $\mathscr{A}=2^X$. Let $\mu^*(\emptyset)=0$, $\mu^*(X)=2$ and $\mu^*(A)=1$ for $A\neq\emptyset,X$. It is an easy task to check μ^* is an outer measure. Now consider $A_n=\{0,1,\dots,n\}$. Then $\cup_n A_n=X$. And $\mu^*(\cup_n A_n)=2>1=\lim_n\mu^*(A_n)$. □

Definition 1. Let $(X, \mathscr{A})$ and $(Y, \mathscr{B})$ be measurable spaces. A function $K : X \times \mathscr{B} \to [0, +\infty]$ is called a *kernel* from $(X, \mathscr{A})$ to $(Y, \mathscr{B})$ if

(i) for each $x \in X$ the function $B \mapsto K(x, B)$ is a measure on $(Y, \mathscr{B})$, and

(ii) for each $B \in \mathscr{B}$ the function $x \mapsto K(x, B)$ is $\mathscr{A}$-measurable.

Exercise 17 (Cohn (2013), Exercise 2.4.7)**.** *Suppose that K is a kernel from $(X, \mathscr{A})$ to $(Y, \mathscr{B})$, that μ is a measure on $(X, \mathscr{A})$, and that f is a $[0, +\infty]$-valued $\mathscr{B}$-measurable function on Y. Then*

(a) $B \mapsto \int K(x, B)\mu(\mathrm{d}x)$ is a measure on $(Y, \mathscr{B})$,

(b) $x \mapsto \int f(y)K(x, \mathrm{d}y)$ is an $\mathscr{A}$-measurable function on X, and

(c) if ν is the measure on $(Y, \mathscr{B})$ defined in part (a), then $\int f(y)\nu(\mathrm{d}y) = \int(\int f(y)K(x, \mathrm{d}y))\mu(\mathrm{d}x)$.

Proof.
(a): As a function of x, $K(x, B)$ is $\mathscr{A}$-measurable. Hence $\int K(x, B)\mu(\mathrm{d}x)$ is well defined for each B. Clearly, $\int K(x, \emptyset)\mu(\mathrm{d}x) = \int 0\mu(\mathrm{d}x) = 0$. Suppose $\{A_i\}_{i=1}^{\infty}$ is an infinite sequence of disjoint sets that belongs to $\mathscr{A}$. Then

$$\int K(x, \cup_{i=1}^{\infty} A_i)\mu(\mathrm{d}x) = \int \sum_{i=1}^{\infty} K(x, A_i)\mu(\mathrm{d}x) = \sum_{i=1}^{\infty} \int K(x, A_i)\mu(\mathrm{d}x),$$

where the last equality follows from the monotone convergence theorem.
(b): If $f = \mathbf{1}_A$ for $A \in \mathscr{B}$, then $\int f(y)K(x, \mathrm{d}y) = K(x, B)$ is $\mathscr{A}$-measurable by the definition of kernel. It follows that $\int f(y)K(x, \mathrm{d}y) = K(x, B)$ is $\mathscr{A}$-measurable for every simple $\mathscr{B}$-measurable function f. Finally, let $f : Y \to [0, +\infty]$ be an arbitrary $\mathscr{B}$-measurable function, and choose a sequence $\{g_n\}$ of simple $\mathscr{B}$-measurable functions from Y to $[0, +\infty)$ such that $g_n(y) \uparrow f(y)$ for each $y \in Y$. Then the monotone convergence theorem implies that $\int f(y)K(x, \mathrm{d}y) = \int \lim_n g_n(y)K(x, \mathrm{d}y) = \lim_n \int g_n(y)K(x, \mathrm{d}y)$. It follows that $\int f(y)K(x, \mathrm{d}y)$ is $\mathscr{A}$-measurable.
(c): If $f = \mathbf{1}_A$ for $A \in \mathscr{B}$, then

$$\int f(y)\nu(\mathrm{d}y) = \nu(A) = \int K(x, A)\mu(\mathrm{d}x) = \int\int f(y)K(x, \mathrm{d}y)\mu(\mathrm{d}x).$$

By the linearity of integral, the conclustion holds for any $[0, +\infty]$ valued simple $\mathscr{B}$-measurable function. Finally, the conclusion holds for any $[0, +\infty]$-valued $\mathscr{B}$-measurable function on Y by the monotone convergence theorem.

□

Exercise 18 (Cohn (2013), Exercise 8.4.6)**.** *Let $(X, \mathscr{A})$ and $(Y, \mathscr{B})$ be measurable spaces, and let K be a kernel from $(X, \mathscr{A})$ to $(Y, \mathscr{B})$ such that $\sup\{K(x, Y), : x \in X\}$ is finite. For each x in X let $B \mapsto \overline{K}(x, B)$ be the restriction to $\mathscr{B}_*$ of the completion of the measure $B \mapsto K(x, B)$. Finally, for each finite measure μ on $(X, \mathscr{A})$ let μK be the measure on $(Y, \mathscr{B})$ defined by $(\mu K)(B) = \int K(x, B)\mu(\mathrm{d}x)$.*

(a) $x, B \to \overline{K}(x, B)$ *is a kernel from* $(X, \mathscr{A}_*)$ *to* $(Y, \mathscr{B}_*)$*.*

(b) Suppose that μ *is a finite measure on* $(X, \mathscr{A})$ *and that* $\overline{\mu}$ *and* $\overline{\mu K}$ *are the restrictions to* $\mathscr{A}_*$ *and* $\mathscr{B}_*$ *of the completions of* μ *and* μK*. Then* $\overline{\mu K} = \overline{\mu}\overline{K}$*, that is,*

$$\overline{\mu K}(B) = \int \overline{K}(x, B)\overline{\mu}(\mathrm{d}x)$$

holds for each B *in* $\mathscr{B}_*$*.*

Proof.
(a): By definition, for any fixed $x \in X$, $B \mapsto \overline{K}(x, B)$ is a measure on $(Y, \mathscr{B})$. We only need to prove that for each $B \in \mathscr{B}_*$, the function $x \mapsto \overline{K}(x, B)$ is $\mathscr{A}_*$-measurable. We apply [Cohn (2013), Exercise 8.4.1(a)] to prove this claim. Fix $B \in \mathscr{B}_*$. We shall prove that for each finite measure μ on $(X, \mathscr{A})$ there are $\mathscr{A}$-measurable functions $f_0, f_1 : X \to [0, +\infty]$ that satisfy $f_0 \leq \overline{K}(x, B) \leq f_1$ everywhere on X and $f_0 = f_1$ μ-almost everywhere.

From Cohn (2013), Exercise 2.4.7(a), the measure $\mu K : \mathscr{B} \to [0, +\infty]$ defined as $A \mapsto \int K(x, A)\mu(\mathrm{d}x)$ is well defined. Also $\int K(x, Y)\mu(\mathrm{d}x) \leq \int \sup_{x \in X} K(x, Y)\mu(\mathrm{d}x) = \sup_{x \in X} K(x, Y)\mu(X) < \infty$. Hence μK is a finite measure on $\mathscr{B}$. Since B is universally measurable, there exists $B_0, B_1 \in \mathscr{B}$ such that $B_0 \subset B \subset B_1$ and $\mu K(B_0) = \mu K(B_1)$. It follows that

$$\int K(x, B_1) - K(x, B_0)\mu(\mathrm{d}x) = 0.$$

Hence $K(x, B_1) = K(x, B_0)$ μ-almost everywhere. Note that $K(x, B_0) \leq \overline{K}(x, B) \leq K(x, B_1)$. This completes the proof.
(b): For $B \in \mathscr{B}$, $\overline{\mu K}(B) = \mu K(B)$ by definition, and

$$\overline{\mu}\overline{K}(B) = \int \overline{K}(x, B)\overline{\mu}(\mathrm{d}x) = \int K(x, B)\overline{\mu}(\mathrm{d}x) = \int K(x, B)\mu(\mathrm{d}x) = \mu K(B).$$

Thus $\overline{\mu K} = \overline{\mu}\overline{K}$ on $\mathscr{B}$. Hence they agree on $\mathscr{B}_*$.

□

References

Cohn, D. L. (2013). *Measure Theory.* Birkhäuser, New York, 2nd edition.

Dudley, R. M. (2002). *Real Analysis and Probability.* Cambridge University Press.