A Formalization of the Lévy-Prokhorov Metric in Isabelle/HOL

Michikazu Hirata

Tokyo Institute of Technology

ITP2024,
September 12, 2024

Tokyo Institute of Technology

Lévy-Prokhorov Metric

The Lévy-Prokhorov Metric
$=$ A metric between finite measures on a metric space.

A measure μ on X

$$\boxed{\mu : \Sigma_X \to [0, \infty]}$$

X : set, $\Sigma_X \subseteq 2^X$: σ-algebra on X.

Intuitively, $\mu(A) =$ the *size* of A
(finite measure $\iff \mu(X) < \infty$)

Ex.

The Lebesgue measure ν on $\mathbb{R}^n$	$\nu\left((a_i, b_i]^n\right) = (b_i - a_i)^n$ infinite measure ($\nu(\mathbb{R}^n) = \infty$)
A probability measure	$P(E) =$ probability E happens finite measure (P(sample space) $= 1$)

Motivation

Let $\mathcal{P}(M) = \{$all finite measures on $M\}$.
$\neq$ power set

X is a Polish space $\Longrightarrow$ So is $\mathcal{P}(X)$

M is a standard Borel space $\Longrightarrow$ So is $\mathcal{P}(M)$

A Polish space / standard Borel space is a topological / measurable space with <u>certain good properties</u>.

Topology on $\mathcal{P}(X)$ / σ-algebra on $\mathcal{P}(M)$ is the least one making $(\lambda\mu.\ \int f \mathrm{d}\mu)$ continuous $\forall f \in \mathrm{C_b}(X)$. / $(\lambda\mu.\ \mu(A))$ measurable $\forall A \in \Sigma_M$.

$(\mathrm{C_b}(X) = \{f : X \to \mathbb{R},\ f$ is bounded countinuous$\})$

Why Polish and Standard Borel Spaces?

X : topological space, M : measurable space

$$X \text{ is a Polish space} \overset{\text{def}}{\iff} X \text{ is separable and completely metrizable}$$

$$M \text{ is a standard Borel space} \overset{\text{def}}{\iff} \exists S : \text{Polish space s.t. } S \text{ generates } M$$

Ex. $\mathbb{R}$, $\mathbb{N}$, and their countable products are Polish (thus standard Borel spaces).

Properties

A Polish space is embedded into a compact space.

A standard Borel space M is either

- countable discrete space or
- $M \cong \mathbb{R}$.

- Large deviation theory

> **3.2.17 Theorem.** (Sanov) *Let μ be a probability measure on the Polish space Σ and let $\tilde{\mu}_n \in \mathbf{M}_1(\mathbf{M}_1(\Sigma))$ be the distribution under μ^n of the function $\mathbf{L}_n$ in (3.2.11). Also define $\mathbf{H}(\cdot|\mu)$ as in (3.2.14). Then $\mathbf{H}(\cdot|\mu)$ is*

Jean-Dominique Deuschel and Daniel W. Stroock, *Large Deviations*, 1989.

- Transportation theory

> **Probability measures**
>
> δ_x is the Dirac mass at point x.
>
> All measures considered in the text are Borel measures on **Polish spaces**, which are complete, separable metric spaces, equipped with their Borel σ-algebra. I shall usually not use the completed σ-algebra

Cédric Villani, *Optimal Transport: Old and New.*, 2008.

- Disintegration theorem

> **Theorem 14.D.10.** Measure disintegration theorem. *Let $(X, \mathcal{X})$ and $(Y, \mathcal{Y})$ be two measurable spaces, and assume that $(Y, \mathcal{Y})$ is Polish. Let λ be a σ-finite measure on $(X \times Y, \mathcal{X} \otimes \mathcal{Y})$ such that the projection λ_X of λ onto $(X, \mathcal{X})$ is also*

Baccelli et al, *Random Measures, Point Processes, and Stochastic Geometry*, 2020

- Quasi-Borel spaces

> where $L(X, M_X) = (X, \Sigma_{M_X})$ and $R(X, \Sigma_X) = (X, M_{\Sigma_X})$. Proposition 15(2) means that the functor R is full and faithful when restricted to standard Borel spaces. Equivalently, $L(R(X, \Sigma_X)) = (X, \Sigma_X)$, that is $\Sigma_X = \Sigma_{M_{\Sigma_X}}$ for standard Borel spaces (X, Σ_X).

[Heunen+, LICS2017]

Lévy-Prokhorov Metric

X is Polish space $\Longrightarrow$ So is $\mathcal{P}(X)$

M is a standard Borel space $\Longrightarrow$ So is $\mathcal{P}(M)$

Need to give a metric on $\mathcal{P}(X)$.

"Metrics" on measures (or distributions)

- Lévy-Prokhorov metric
- Wasserstein metric
- Relative entropy
- etc.

The Lévy-Prokhorov metric

- always being a metric in the mathematical sense
- Detailed lecture notes by Gaans (2002/2003)

Contribution: Polishness of $\mathcal{P}(X)$

X is Polish space $\Longrightarrow$ So is $\mathcal{P}(X)$

$$X \text{ is a Polish space} \overset{\text{def}}{\Longleftrightarrow} X \text{ is separable and completely metrizable}$$

Step 1. Separability of $\mathcal{P}(X)$

Step 2. Completeness of $(\mathcal{P}(X), d_{\mathcal{P}(X)})$ Tough!!

- Prokhorov's theorem
  - Riesz representation theorem (2.1K lines of proofs)
  - Alaoglu's theorem

Step 3. $(\mathcal{P}(X), d_{\mathcal{P}(X)})$ metrizes $\mathcal{P}(X)$

M is a standard Borel space $\Longrightarrow$ So is $\mathcal{P}(M)$

$$M \text{ is a standard Borel space} \overset{\text{def}}{\Longleftrightarrow} \exists S\colon \text{Polish space s.t. } S \text{ generates } M$$

Show $(\mathcal{P}(M), d_{\mathcal{P}(M)})$ generates $\mathcal{P}(M)$

We construct the proof

Contribution: Prokhorov's Theorem

Prokhorov's Theorem

X : a Polish space

$\Gamma \subseteq \mathcal{P}(X) \cap \{\mu.\ \mu(X) \leq r\}$ for some $r < \infty$.

$\overline{\Gamma}$ is compact in $\mathcal{P}(X) \iff \Gamma$ is tight: i.e.,

$$\forall \varepsilon > 0.\ \exists K : \text{compact in } X,\ \text{s.t.} \forall \mu \in \Gamma.\ \mu(X - K) \leq \varepsilon$$

Used for

- Completeness of the Lévy-Prokhorov metric
- Central limit theorem
- Sanov's theorem in large deviation theory
- Existence of optimal coupling in transportation theory

depending on

- Riesz representation theorem
- Alaoglu's theorem

Results

Formalization in Isabelle/HOL of

- Lévy-Prokhorov metric
  - weak convergence
  - Portmanteau theorem
  - separability and completeness
- Prokhorov's theorem
  - a special case of Alaoglu's theorem
  - Riesz representation theorem
- $\mathcal{P}(M)$ is a Polish and standard Borel space
  - equivalence of measurable spaces of finite measures

Archive of formal proofs

- *The Lévy-Prokhorov Metric*, June 2024 (6.6K lines)
- *The Riesz representation theorem*, June 2024 (4.4K lines)

Results

Formalization in Isabelle/HOL of

- Lévy-Prokhorov metric
  - weak convergence
  - Portmanteau theorem
  - separability and completeness
- Prokhorov's theorem
  - a special case of Alaoglu's theorem
  - Riesz representation theorem
- $\mathcal{P}(M)$ is a Polish and standard Borel space
  - equivalence of measurable spaces of finite measures

Archive of formal proofs

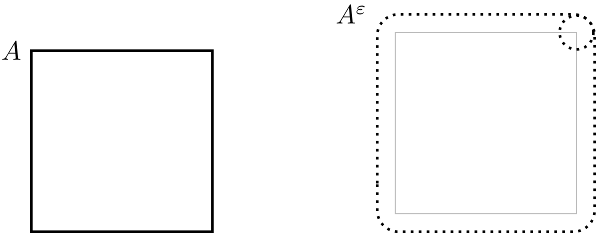
- *The Lévy-Prokhorov Metric*, June 2024 (6.6K lines)
- *The Riesz representation theorem*, June 2024 (4.4K lines)

Lévy-Prokhorov Metric

Let X be a metric space.

The Lévy-Prokhorov metric $d_{\mathcal{P}(X)}$ on $\mathcal{P}(X)$

$$d_{\mathcal{P}(X)}(\mu, \nu) = \inf\{\varepsilon > 0 \mid \forall A \in \Sigma_X.\ \mu(A) \leq \nu(A^\varepsilon) + \varepsilon \\ \wedge \nu(A) \leq \mu(A^\varepsilon) + \varepsilon\}.$$



$A^\varepsilon = \bigcup_{x \in A} ball_X(x, \varepsilon)$

Lévy-Prokhorov Metric

Weak Convergence

For $\{\mu_n\}_{n\in\mathbb{N}} \subseteq \mathcal{P}(X)$ and $\mu \in \mathcal{P}(X)$,

$$\mu_n \Rightarrow_{\mathrm{wc}} \mu \overset{\text{def}}{\iff} \forall f : X \to \mathbb{R} \text{ bounded continuous,}$$
$$\int f \mathrm{d}\mu_n \longrightarrow \int f \mathrm{d}\mu$$

The Lévy-Prokhorov metric *metrizes* weak convergence.

Theorem

1. $\mu_n \longrightarrow \mu$ in $(\mathcal{P}(X), d_{\mathcal{P}(X)})$ implies $\mu_n \Rightarrow_{\mathrm{wc}} \mu$
2. If X is separable, then $\mu_n \longrightarrow \mu$ in $(\mathcal{P}(X), d_{\mathcal{P}(X)})$ iff $\mu_n \Rightarrow_{\mathrm{wc}} \mu$

Lévy-Prokhorov Metric in Isabelle/HOL

Metric spaces in Isabelle/HOL

- Type-based (**class**) by [Hölzl+, ITP2013]
  - work better in automation
  - carrier sets must be *UNIV* (all elements of the type)

    Ex. Metric on

    ☺ $\mathbb{R}$, $\mathbb{N}$

    ☹ $[0, 1]$, $\mathrm{C_b}(X)$ (bounded continuous functions)
- Set-based (**locale**, **typedef**) by Paulson* since Isabelle2023
  - Any carrier sets

We use the set-based library because

$$\mathcal{P}(X) = \{\mu.\ \mu \text{ is a finite measures on } X\} \neq UNIV.$$

* Lawrence C. Paulson, Porting the HOL Light metric space library,

https://lawrencecpaulson.github.io/2023/07/12/Metric_spaces.html

Set-Based Metric Spaces in Isabelle/HOL

`HOL/Analysis/Abstract_Metric_Spaces.thy`

locale *Metric-space* $=$
  fixes M :: *'a set* **and** d :: $'a \Rightarrow\ 'a \Rightarrow$ *real*
  assumes *nonneg*: $\bigwedge x\ y.\ 0 \leq d\ x\ y$
  assumes *commute*: $\bigwedge x\ y.\ d\ x\ y = d\ y\ x$
  assumes *zero*: $\bigwedge x\ y.\ [\![x \in M;\ y \in M]\!] \Longrightarrow d\ x\ y = 0 \longleftrightarrow x{=}y$
  assumes *triangle*: $\bigwedge x\ y\ z.\ [\![x \in M;\ y \in M;\ z \in M]\!] \Longrightarrow d\ x\ z \leq d\ x\ y + d\ y\ z$

- (M,d) forms a metric space.
- Non-negativity and commutativity must hold on the whole type.

Set-Based Metric Spaces in Isabelle/HOL

`HOL/Analysis/Abstract_Metric_Spaces.thy`

locale *Metric-space* =
  fixes M :: $'a$ set **and** d :: $'a \Rightarrow {}'a \Rightarrow real$
  assumes *nonneg*: $\bigwedge x\ y.\ 0 \le d\ x\ y$
  assumes *commute*: $\bigwedge x\ y.\ d\ x\ y = d\ y\ x$
  assumes *zero*: $\bigwedge x\ y.\ [\![x \in M;\ y \in M]\!] \Longrightarrow d\ x\ y = 0 \longleftrightarrow x{=}y$
  assumes *triangle*: $\bigwedge x\ y\ z.\ [\![x \in M;\ y \in M;\ z \in M]\!] \Longrightarrow d\ x\ z \le d\ x\ y + d\ y\ z$

- (M,d) forms a metric space.
- Non-negativity and commutativity must hold <u>on the whole type</u>.

Lévy-Prokhorov Metric in Isabelle/HOL

```
locale Levy-Prokhorov = Metric-space
begin
                N is a measure on (M, d)
definition 𝒫 ≡ {N. sets N = sets (borel-of mtopology) ∧ finite-measure N}

definition LPm :: 'a measure ⇒ 'a measure ⇒ real where
LPm N L ≡
  if N ∈ 𝒫 ∧ L ∈ 𝒫 then
     (⨅ {e. e > 0 ∧ (∀A∈sets (borel-of mtopology).
                        measure N A ≤ measure L (⋃a∈A. mball a e) + e ∧
                        measure L A ≤ measure N (⋃a∈A. mball a e) + e)})
  else 0    for ⋀x y. 0 ≤ d x y (because ⨅∅ = ?)

sublocale LPm: Metric-space 𝒫 LPm
⟨proof⟩

end
```

c.f.

$$d_{\mathcal{P}(X)}(\mu,\nu) = \inf\{\varepsilon > 0 \mid \forall A \in \Sigma_X.\, \mu(A) \le \nu(A^\varepsilon) + \varepsilon \wedge \nu(A) \le \mu(A^\varepsilon) + \varepsilon\},$$

where $A^\varepsilon = \bigcup_{a\in A} ball_X(a,\varepsilon)$.

Convergence w.r.t. Filters

Limit is defined using filters in Isabelle/HOL.

$$x_n \xrightarrow{n\to\infty} x \iff (x_n \longrightarrow x)\ \mathcal{F}_{\mathrm{seq}} \text{ in } \mathbb{R}$$
$$f(x) \xrightarrow{x\to a} L \iff (f \longrightarrow L)\ (\mathrm{at}\ a) \text{ in } \mathbb{R}$$

F_{seq} : a filter on $\mathbb{N}$, $(\mathrm{at}\ a)$: a filter on $\mathbb{R}$

A *filter* $\mathcal{F}$ on I : $\mathcal{F} \subseteq 2^I$ satisfying certain conditions.

In Isabelle/HOL, limit is

- *tendsto* (type-based topological spaces)
- *limitin* (set-based topological spaces)

Weak Convergence with Filters

Weak Convergence

For $\{\mu_n\}_{n\in\mathbb{N}} \subseteq \mathcal{P}(X)$, $\mu \in \mathcal{P}(X)$,

$$\mu_n \Rightarrow_{\mathrm{wc}} \mu \qquad \overset{\text{def}}{\iff} \forall f : X \to \mathbb{R} \text{ bounded continuous,}$$

$$\int f \mathrm{d}\mu_n \longrightarrow \int f \mathrm{d}\mu$$

The Lévy-Prokhorov metric *metrizes* weak convergence.

Theorem

1. $\mu_n \longrightarrow \mu \qquad$ in $(\mathcal{P}(X), d_{\mathcal{P}(X)})$ implies $\quad \mu_n \Rightarrow_{\mathrm{wc}} \mu$
2. If X is separable, then

$$\mu_n \longrightarrow \mu \qquad \text{in } (\mathcal{P}(X), d_{\mathcal{P}(X)}) \text{ iff } \quad \mu_n \Rightarrow_{\mathrm{wc}} \mu$$

Weak Convergence with Filters

Weak Convergence

For $\{\mu_i\}_{i\in I} \subseteq \mathcal{P}(X)$, $\mu \in \mathcal{P}(X)$, and a filter $\mathcal{F}$ on I,

$$(\mu_i \Rightarrow_{\mathrm{wc}} \mu)\ \mathcal{F} \overset{\text{def}}{\iff} \forall f : X \to \mathbb{R} \text{ bounded continuous,}$$

$$\left(\int f \mathrm{d}\mu_i \longrightarrow \int f \mathrm{d}\mu\right)\ \mathcal{F} \text{ in } \mathbb{R}$$

The Lévy-Prokhorov metric *metrizes* weak convergence.

Theorem

1. $(\mu_i \longrightarrow \mu)\ \mathcal{F}$ in $(\mathcal{P}(X), d_{\mathcal{P}(X)})$ implies $(\mu_i \Rightarrow_{\mathrm{wc}} \mu)\ \mathcal{F}$
2. If X is separable, then

$$(\mu_i \longrightarrow \mu)\ \mathcal{F} \text{ in } (\mathcal{P}(X), d_{\mathcal{P}(X)}) \text{ iff } (\mu_i \Rightarrow_{\mathrm{wc}} \mu)\ \mathcal{F}$$

Topology of Weak Convergence

Topology of Weak Convergence

The topology of weak convergence $\mathcal{O}_{\mathrm{WC}(X)}$ on $\mathcal{P}(X)$
$\mathcal{O}_{\mathrm{WC}(X)} =$ the coarsest topology making $(\lambda\mu.\ \int f \mathrm{d}\mu)$ continuous $\forall f \in \mathrm{C_b}(X)$.

Lemma

$(\mu_i \longrightarrow \mu)\ \mathcal{F}$ in $(\mathcal{P}(X), \mathcal{O}_{\mathrm{WC}(X)})$ iff $(\mu_i \Rightarrow_{\mathrm{wc}} \mu)\ \mathcal{F}$.

Proposition

If X is a separable metric space, then

$$(\mathcal{P}(X), \mathcal{O}_{\mathrm{WC}(X)}) = (\mathcal{P}(X), \mathcal{O}_{d_{\mathcal{P}(X)}})$$

This proposition is a well-known result, but few books contain its proof.

Our new proof uses <u>limit w.r.t. filters</u>.

Metrizability of Topology of Weak Convergence

Proposition

If X is a separable metric space, then

$$(\mathcal{P}(X), \mathcal{O}_{\mathrm{WC}(X)}) = (\mathcal{P}(X), \mathcal{O}_{d_{\mathcal{P}(X)}})$$

Claim

$\mathcal{O}$, $\mathcal{O}'$: topologies on X

$$\mathcal{O} = \mathcal{O}' \iff (\forall \mathcal{F}.\ \text{convergence in } \mathcal{O} \text{ w.r.t. } \mathcal{F} = \text{convergence in } \mathcal{O}' \text{ w.r.t. } \mathcal{F})$$

($\Longrightarrow$) trivial, ($\Longleftarrow$) this work

Proof of Proposition: For any filter $\mathcal{F}$,

$$\begin{aligned}\text{convergence in } \mathcal{O}_{\mathrm{WC}(X)} \text{ w.r.t. } \mathcal{F} &= \text{weak convergence w.r.t. } \mathcal{F} \\ &= \text{convergence in } (\mathcal{P}(X), d_{\mathcal{P}(X)}) \text{ w.r.t. } \mathcal{F}\end{aligned}$$

Metrizability of Topology of Weak Convergence

Proposition

If X is a separable metric space, then

$$(\mathcal{P}(X), \mathcal{O}_{\mathrm{WC}(X)}) = (\mathcal{P}(X), \mathcal{O}_{d_{\mathcal{P}(X)}})$$

The textbook by Billingsley (1968)

- By examining neighborhoods

The textbook by Deuschel and Stroock (1989)

- By showing

  convergence in $\mathcal{O}_{\mathrm{WC}(X)}$ w.r.t. $\mathcal{F}_{\mathrm{seq}}$ = convergence in $(\mathcal{P}(X), d_{\mathcal{P}(X)})$ w.r.t. $\mathcal{F}_{\mathrm{seq}}$

  (sequences are on $\mathbb{N}$)
- Need to show the first-countability of $(\mathcal{P}(X), \mathcal{O}_{\mathrm{WC}(X)})$

Our proof

- By using equivalence of limit w.r.t. filters

Related Works

In Isabelle/HOL by Avigad et al. (2017)

- weak convergence for ~~probability measures~~ on ~~$\mathbb{R}$~~
  finite measures any topological spaces
- ~~a special case~~ of Prokhorov's theorem
  the general form
- applied them to prove the central limit theorem

In Lean by Kytölä (ongoing, 2023–)

- weak convergence on finite measures using filters
  same as ours
- X is a pseudometric space $\Longrightarrow$ The Lévy-Prokhorov metric is a pseudometric
  X is a metric space $\Longrightarrow$ The Lévy-Prokhorov metric is a metric
  "finer" than our work
  metric = pseudometric + $(d(x, y) = 0 \iff x = y)$
- The Lévy-Prokhorov metric on ~~probability measures~~ metrizes the topology of weak convergence
  finite measures

  $+$ more results (e.g. Prokhorov's theorem)

Efforts

Limit using filters
$\Longrightarrow$ Similar to the sequential cases

Sequential $(\forall n \geq N_1.\, x_n < \varepsilon/2) \wedge (\forall n \geq N_2.\, y_n < \varepsilon/2)$
$\Longrightarrow \forall n \geq \max\{N_1, N_2\}.\, x_n + y_n < \varepsilon$

Filter *eventually* $x_i < \varepsilon/2$ w.r.t. $\mathcal{F}$ $\wedge$ *eventually* $y_i < \varepsilon/2$ w.r.t. $\mathcal{F}$
$\Longrightarrow$ *eventually* $x_i + y_i < \varepsilon$ w.r.t. $\mathcal{F}$

Riesz representation theorem

- 9 pages in the book by Rudin "*Real and Complex Analysis*" (1987)
- 2.1K– lines in Isabelle/HOL

Conclusion

We formalized

- Lévy-Prokhorov metric
- Prokhorov's theorem
- $\mathcal{P}(M)$ is a Polish and standard Borel space

in Isabelle/HOL

Future Works

- Large deviation theory
- Transportation theory
- Point process