An introduction to statistical modelling semantics with higher-order measure theory

Ohad Kammar
University of Edinburgh

Logic of Probabilistic Programming
Logique de la programmation probabiliste
31 January–4 February, 2022
Logic and Interactions — Logique et interactions
CIRM Thematic Month

supported by: Facebook Research NCSC

Spaces Statistical Modelling needs:

Discrete finite

Bool

true false

Animal

cat
dog
Giraffe

Discrete Countable

$\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, String

Continuous

$\mathbb{R}^n$ $\mathbb{R}^{\mathbb{N}}$

$\mathcal{P}\mathbb{R}$,

Weight

standard Borel spaces

Measurable

Recent developments:

Discrete finite

Discrete Countable

Continuous

Boolean-Valued Models [Bacci et al '18]

Regular ordered Banach [Dahlqvist-Kozen '20]

Probabilistic Coherence Spaces & Measurable Cones [Ehrhard-Pagani-Tasson '18]

This talk

Quasi-Borel Spaces [Heunen et al. 17]

~~Measurable~~

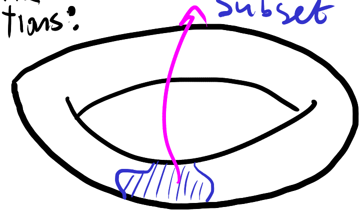
Core idea

Measure Theory	Qbs Theory

sample space Ω

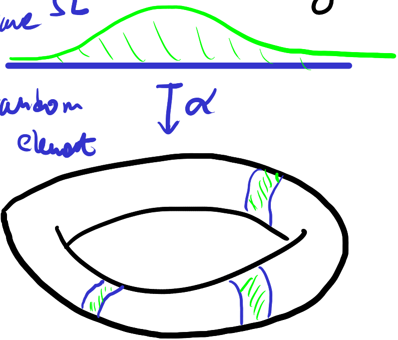
Primitive notions:

measurable subset



random element

$\downarrow \alpha$



Derived notions:

random elements
$\alpha : \Omega \to \text{Space}$

measure

measurable subsets

Conservative extensions:

concrete spaces we "observe"

Standard Borel spaces

Meas

Qbs

abstract spaces

(auxiliary?)

Wide topic:

Variations

Qbs, wQSs, [Vákár et al '19]

QMS, QUS, [Forré '21]

(w)Diff, wPap [Lew et al. '22]

[Vákár et al. '20-'21]

Applications

MC inference

design & verification [Ścibior et al '17, '18a+b]

Network Programming [Vandenbroucke - Schrijvers '19]

Semantics

name generation [Sabok et al. '21]

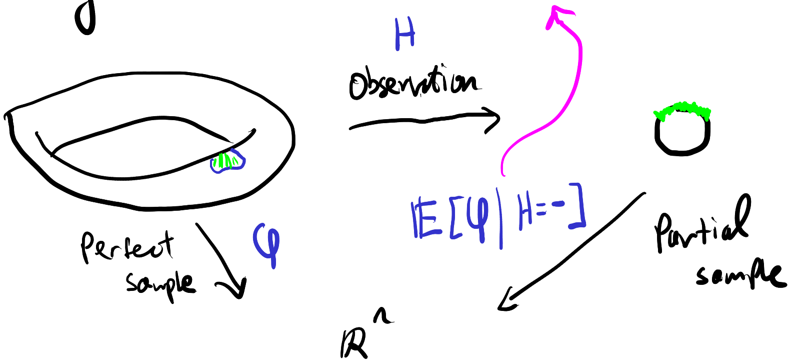
This tutorial:

- Peek behind scenes
- Gain working knowledge

Theme: higher-order measure theory

demonstrated through

Kolmogorov's Conditional Expectation



Kolmogorov's Conditional Expectation

- naturally higher order: $\mathbb{R}^{\Omega} \to \mathbb{R}^{\Theta}$
- behind many modern Probability techniques:
  - existence of Radon-Nikodym derivatives & density
  - existence of disintegration
  - foundation of martingales & stochastic differential equations

Agenda

Slogan:

Measurable by Type

I
- Borel sets
- Qbs:
  def., constructions,
  partiality, refinement
- Measures & integration

II
- Random variable spaces
- Conditional expectation

Space: all possible states

Subset: all states of current interest

Measure: probability/weight/length assignment to

fine for discrete spaces

e.g. $\{H, T\}^5$

HHTTH

$\frac{1}{32}$

Continuous Caveat:

Thm: no $\lambda : \mathcal{P}\mathbb{R} \to [0, \infty]$:

$\lambda(a,b) = b - a$ (generalises length)

$\lambda(r + A) = \lambda A$ (translation invariant)

$\lambda\left(\biguplus_{n=0}^{\infty} A_n\right) = \sum_{n=0}^{\infty} \lambda A_n$ σ-additive

Workaround: only measure well-behaved subsets

Def: The Borel subsets $\mathcal{B}_{\mathbb{R}} \subseteq \mathbb{R}$:

- open intervals $(a,b) \in \mathcal{B}_{\mathbb{R}}$

Closure under σ-algebra operations:

$$\frac{}{\emptyset \in \mathcal{B}_{\mathbb{R}}} \qquad \frac{A \in \mathcal{B}_{\mathbb{R}}}{A^c := \mathbb{R} \setminus A \in \mathcal{B}} \qquad \frac{\vec{A} \in \mathcal{B}_{\mathbb{R}}^{\mathbb{N}}}{\bigcup_{n=0}^{\infty} A_n \in \mathcal{B}_{\mathbb{R}}}$$

empty set

complements

countable unions

Examples

discrete Countable: $\{r\} = \bigcap_{\varepsilon \in \mathbb{Q}^+} (r-\varepsilon, r+\varepsilon) \in \mathcal{B}_{\mathbb{R}}$

I countable $\Rightarrow$ $I = \bigcup_{r \in I} \{r\} \in \mathcal{B}_{\mathbb{R}}$

closed intervals: $[a,b] = (a,b) \cup \{a,b\}$

Non-examples?

More complicated: analytic, Lebesgue

Def: Measurable space $V = (|V|, B_V)$

Set (carrier)

family of Subsets $B_V \subseteq P|V|$

Closed under σ-algebra operations:

$$\frac{}{\emptyset \in B_V} \qquad \frac{A \in B_{\mathbb{R}}}{A^c := |V| \setminus A \in B} \qquad \frac{\vec{A} \in B_{\mathbb{R}}^{\mathbb{N}}}{\bigcup_{n=0}^{\infty} A_n \in B_V}$$

empty set

complements

countable unions

Idea: Structure all spaces after the worst-case scenario

Examples

- Discrete spaces $\ulcorner X^{\text{meas}} \urcorner = (X, PX)$

- Euclidean spaces $\mathbb{R}^n$ — replace intervals with chests $\prod_{i=1}^{n} (a_i, b_i)$

  $\mathbb{R}^{\mathbb{N}}$ similarly

$\{C \cap A \mid C \in B_V\}$

- Subspaces: $A \in P|V|$ $\quad A := (A, [B_V] \cap A)$

Def: Borel measurable functions $f : V_1 \to V_2$

- functions $f : \lfloor V_1 \rfloor \longrightarrow \lfloor V_2 \rfloor$
- inverse image preserves measurability:

$$f^{-1}[A] \in B_{V_1} \Longleftarrow A \in B_{V_2}$$

Examples

- $(+), (\cdot) : \mathbb{R}^2 \longrightarrow \mathbb{R}$
- $|-|$, $\sin : \mathbb{R} \to \mathbb{R}$
- any continuous function $f : \mathbb{R}^n \to \mathbb{R}^m$
- any function $f : {}^{\ulcorner}X^{\urcorner} \to V$

Category **Meas**

Objects : Measurable spaces

Morphisms : Measurable functions

Identities:

$id : V \to V$

Composition:

$$\frac{f : V_2 \to V_3 \qquad g : V_1 \to V_2}{f \circ g : V_1 \to V_3}$$

Meas Category

Products, coproducts/disjoint union, subspaces
categorical limits, colimits, but:

Thm [Aumann '61] No σ-algebra $B_{\mathbb{R}^{\mathbb{R}}}$ for measurable

$$\mathrm{eval} : (\mathrm{Meas}(\mathbb{R},\mathbb{R}), B_{\mathbb{R}^{\mathbb{R}}}) \times \mathbb{R} \to \mathbb{R}$$
$$(f, r) \mapsto f(r)$$

Questions? skip proof?

Proof (sketch):

Borel hierarchy:

$$
\begin{tikzcd}[column sep=small]
\Sigma^0_0 \arrow[dr, phantom, "\subsetneq"] & & \Sigma^0_1 \arrow[dr, phantom, "\subsetneq"] & & & & \Sigma^0_\omega \arrow[dr, phantom, "\subsetneq"] & & \\
& \Delta^0_1 \arrow[ur, phantom, "\subsetneq"] \arrow[dr, phantom, "\subsetneq"] & & \Delta^0_2 \arrow[r, phantom, "\subsetneq"] & \cdots & \Delta^0_\omega \arrow[ur, phantom, "\subsetneq"] \arrow[dr, phantom, "\subsetneq"] & & \Delta^0_{\omega+1} \arrow[r, phantom, "\subsetneq"] & {} \\
\Pi^0_0 \arrow[ur, phantom, "\subsetneq"] & & \Pi^0_1 \arrow[ur, phantom, "\subsetneq"] & & & & \Pi^0_\omega \arrow[ur, phantom, "\subsetneq"] & &
\end{tikzcd}
$$

Stabilises at $\Delta^0_{\omega_1} = \mathcal{B}(\Sigma^0_0) = \Delta^0_{\omega_1+1}$

$$\text{rank } A := \min\{\alpha < \omega_1 \mid A \in \Delta^0_\alpha\}$$

even for $B_{\mathbb{R}^{\mathbb{R}}} = P(\mathrm{Meas}(\mathbb{R},\mathbb{R}))$

$$\mathrm{eval} : (\mathrm{Meas}(\mathbb{R},\mathbb{R}), B_{\mathbb{R}^{\mathbb{R}}}) \times \mathbb{R} \to \mathbb{R}$$
$$(f, r) \mapsto f(r)$$

$B_{V\times U} = B([B_V] \times [B_U])$

If measurable:

$$\alpha := \sup\{\mathrm{rank}(\mathrm{eval}^{-1}[(p,q)]) \mid p,q \in \mathbb{Q}\} < \omega_1$$

Take $A \in B_{\mathbb{R}}$, $\mathrm{rank}\,A > \alpha$

$f : \mathbb{R} \to \mathbb{R}$ $\quad f := [- \in A] := \lambda x. \begin{cases} x \in A : 1 \\ x \notin A : 0 \end{cases}$

But:

$$\alpha < \mathrm{rank}\,A = \mathrm{rank}(\langle f, -\rangle^{-1}[\mathrm{eval}^{-1}[\{1\}]] \leq \mathrm{rank}(\mathrm{eval}^{-1}[\{1\}]) \leq \alpha$$

\#

Sequential Higher-order structure:

I Countable : $V^{\ulcorner I \urcorner} = \prod_{i \in I} V$

$\Longrightarrow$ Some higher-order structure in **Meas**:

$\text{Cauchy} \in B_{[-\infty,\infty]^{\mathbb{N}}}$

$$\text{Cauchy} := \bigcap_{\varepsilon \in \mathbb{Q}^+} \bigcup_{k \in \mathbb{N}} \bigcap_{\substack{m,n \in \mathbb{N} \\ m,n \geq k}} \{ \vec{y} \in [-\infty,\infty]^{\mathbb{N}} \mid |y_m - y_n| < \varepsilon \}$$

$\limsup : [-\infty,\infty]^{\mathbb{N}} \to [-\infty,\infty]$ $\quad$ $\lim : \text{Cauchy} \to \mathbb{R}$

Compose higher-order building blocks:

lim is measurable!

$$\mathrm{VanishingSeq}(\mathbb{R}) := \left\{ \vec{r} \in \mathbb{R}^{\mathbb{N}} \,\middle|\, \lim_{n\to\infty} r_n = 0 \right\} \in B_{\mathbb{R}^{\mathbb{N}}}$$

$$\mathrm{approx}\_ : \mathrm{VanishingSeq}(\mathbb{R}^+) \times \mathbb{R} \longrightarrow \mathbb{Q}^{\mathbb{N}}$$

$$\text{s.t.:} \quad \left| \left(\mathrm{approx}_{\vec{\Delta}}\, r \right)_n - r \right| < \Delta_n$$

Slogan: Measurable by Type!

Not all operations of interest fit:

Intrinsically higher-order!

$$\limsup : \left([-\infty, \infty]^{\mathbb{R}} \right)^{\mathbb{N}} \longrightarrow [-\infty, \infty]^{\mathbb{R}}$$

$$\limsup := \lambda \vec{f}.\lambda x. \limsup_{n\to\infty} f_n\, x$$

Want

Slogan: Measurable by Type !

But

For higher-order building-blocks, must

defer measurability proofs until we're

1st order again. ⇒ non-compositionality

Plan

Def: $V \in \mathbf{Meas}$ is **Standard Borel** when

$$V \cong A \quad \text{for some } A \in \mathcal{B}_{\mathbb{R}}$$

the "good part" of **Meas** — the subcategory

$$\mathbf{Sbs} \hookrightarrow \mathbf{Meas}$$

Sbs includes

- Discrete $\ulcorner I \urcorner$, I countable
- Countable products of Sbs:

$$\mathbb{R}^n, \mathbb{R}^{\mathbb{N}}, 2^{\mathbb{N}}, \mathbb{N}^{\mathbb{N}}$$

- Borel subspaces of Sbs:

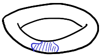
$\mathbb{I} := [0,1]$ $\mathbb{R}^+ := (0,\infty)$ $\mathbb{R}_{\geq 0} := [0,\infty]$

- Countable coproducts of Sbs:

$\mathbb{W} := [0,\infty]$ $\overline{\mathbb{R}} := [-\infty,\infty]$

Agenda

Slogan: Measurable by Type

- Borel sets 
- Qbs:
  def., constructions,
  partiality, refinement
- Measures & integration
- Random variable spaces
- Conditional expectation

Def: Quasi-Borel space $X = (\lfloor X \rfloor, R_X)$

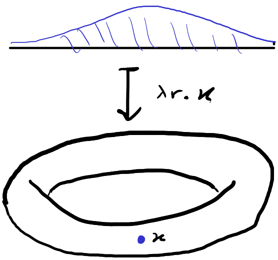
set "carrier"

set of functions $\alpha : \mathbb{R} \to \lfloor X \rfloor$ "random elements"

$R_X \subseteq \lfloor X \rfloor^{\lfloor \mathbb{R} \rfloor}$ closed under:

- Constants:

$$\frac{x \in \lfloor X \rfloor}{(\lambda r.x) \in R_X}$$



- precomposition:

- recombination

Def: Quasi-Borel space $X = (\lfloor X \rfloor, R_X)$

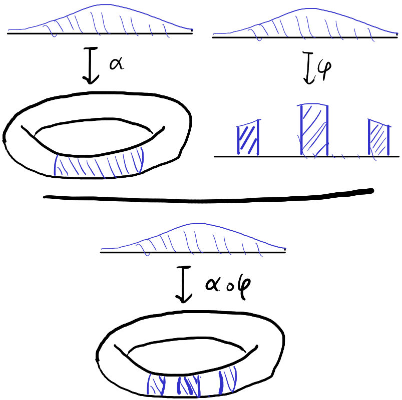
set "carrier"

set of functions $\alpha : \mathbb{R} \to \lfloor X \rfloor$ "random elements"

$R_X \subseteq \lfloor X \rfloor^{\lfloor \mathbb{R} \rfloor}$ closed under:

- precomposition:

$$\frac{\alpha \in R_X \qquad \varphi : \mathbb{R} \to \mathbb{R} \text{ in Sbs}}{\varphi \circ \alpha : \mathbb{R} \xrightarrow{\varphi} \mathbb{R} \xrightarrow{\alpha} \lfloor X \rfloor \in R_X}$$



Def: Quasi-Borel space $X = (\lfloor X \rfloor, R_X)$

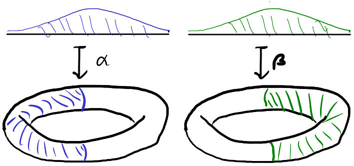
Set "carrier"

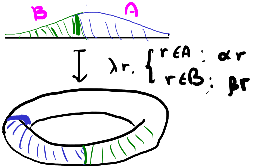
Set of functions $\alpha : \mathbb{R} \to \lfloor X \rfloor$ "random elements"

$R_X \subseteq \lfloor X \rfloor^{\lfloor \mathbb{R} \rfloor}$ closed under:

- recombination

$$\frac{\vec{\alpha} \in R_X^{\mathbb{N}} \qquad \mathbb{R} = \biguplus_{n=0}^{\infty} \underbrace{A_n}_{\in \mathcal{B}_{\mathbb{R}}}}{\lambda r. \begin{cases} \vdots \\ r \in A_n : \alpha_n r \\ \vdots \end{cases}}$$



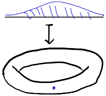


Def: Quasi-Borel space $X = (\lfloor X \rfloor, R_X)$

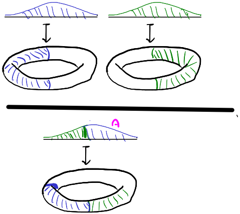
set "carrier"

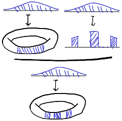
set of functions $\alpha : \mathbb{R} \to \lfloor X \rfloor$ "random elements"

$R_X \subseteq \lfloor X \rfloor^{\lfloor \mathbb{R} \rfloor}$ closed under:

- Constants: 

- recombination



- Precomposition: 

Examples

- $\mathbb{R} = (\lfloor\mathbb{R}\rfloor, \mathrm{Meas}(\mathbb{R},\mathbb{R}))$

  qbs underlying $\mathbb{R}$

- $X \in \mathrm{Set}$, $\ulcorner X \urcorner^{\mathrm{Qbs}} := (X, \sigma\text{-Simple}(\mathbb{R}, X))$

  recombination of constants

  $$\lambda r. \begin{cases} \vdots \\ r \in A_n : x_n \\ \vdots \end{cases}$$

  discrete qbs on X

- " $\lfloor X \rfloor_{\mathrm{Qbs}} := (X, X^{\lfloor\mathbb{R}\rfloor})$

  all functions

  Indiscrete qbs on X

Qbs morphism $f : X \to Y$

- function $f : \lfloor X \rfloor \to \lfloor Y \rfloor$
- $$\frac{\mathbb{R} \xrightarrow{\alpha} \lfloor X \rfloor \in R_X}{\mathbb{R} \xrightarrow{\alpha} \lfloor X \rfloor \xrightarrow{f} \lfloor Y \rfloor \in R_Y}$$

Example

- Constant functions are qbs morphisms
- σ-simple functions are qbs morphisms

Category Qbs $\Leftarrow$ - identity, composition

Useful adjunctions:

$$
\begin{tikzcd}
\mathbf{Meas} \arrow[d, "\lfloor - \rfloor_{Qbs}"'] \arrow[d, phantom, bend right=60, "\dashv"] \\
\mathbf{Qbs} \arrow[u, bend left=80, "\lceil - \rceil^{Meas}"] \arrow[d, "\lfloor - \rfloor"] \\
\mathbf{Set} \arrow[u, bend left=60, "\lceil - \rceil^{Qbs}"] \arrow[u, bend right=60, "\lfloor \overline{-} \rfloor_{Qbs}"']
\end{tikzcd}
$$

$\lceil - \rceil^{Qbs} \dashv \lfloor - \rfloor \dashv \lfloor \overline{-} \rfloor_{Qbs}$

$$\lfloor V \rfloor_{Qbs} := (\lfloor V \rfloor, \mathbf{Meas}(\mathbb{R}, V)) \quad (V \in \mathbf{Meas})$$

$$\lceil X \rceil^{Meas} := \{ A \subseteq \lfloor X \rfloor \mid \forall \alpha \in R_X .\ \alpha^{-1}[A] \in \mathcal{B}_{\mathbb{R}} \}$$

- limits (products, subspaces) and colimits (coproducts, quotients) as in **Set**
- Slogan: every measurable space is carried by a qbs

Example

Product $(X \times Y, \pi_1, \pi_2)$:

necessarily!

- $\lfloor X \times Y \rfloor = \lfloor X \rfloor \times \lfloor Y \rfloor$
- $R_{X \times Y} = \{ \lambda r. (\alpha r, \beta r) \mid \alpha \in R_X, \beta \in R_Y \}$

correlated random elements

rest of structure as in **Set**.

Function Spaces

Straightforward!

- $\lfloor Y^X \rfloor := \mathbf{Qbs}(X, Y)$
- $R_{Y^X} := \mathrm{uncurry}[\mathbf{Qbs}(\mathbb{R} \times X, Y)]$

  $= \{ \alpha : \mathbb{R} \to \lfloor Y^X \rfloor \mid \lambda(r, x).\, \alpha\, r\, x : \mathbb{R} \times X \to Y \}$
- $\mathrm{eval} : Y^X \times X \to Y$

  $\mathrm{eval}(f, x) := f\, x$

Meas vs Qbs

By generalities:

σ-algebra on $\mathrm{Meas}(\mathbb{R},\mathbb{R})$

$$\begin{tikzcd} \ulcorner \mathbb{R}^{\mathbb{R}}_{\mathrm{Meas}} \urcorner \times \mathbb{R} \arrow[r] \arrow[rr, bend right, "\ulcorner \mathrm{eval}^{\mathrm{Meas}} \urcorner"'] & \ulcorner \mathbb{R}^{\mathbb{R}}_{\mathrm{Meas}} \urcorner \times \mathbb{R} \arrow[r, dashed, "\times"] & \ulcorner \mathbb{R} \urcorner = \mathbb{R} \end{tikzcd}$$

$$\begin{tikzcd} \mathrm{Meas} \arrow[d, bend right, "\lfloor - \rfloor_{\mathrm{Qbs}}"'] \\ \mathrm{Qbs} \arrow[u, bend left, "\ulcorner - \urcorner^{\mathrm{Meas}}"] \end{tikzcd} \qquad \ulcorner - \urcorner^{\mathrm{Meas}} \dashv \lfloor - \rfloor_{\mathrm{Qbs}}$$

No factorisation by Aumann's Theorem.

So:

$$\left(\ulcorner \mathbb{R}^{\mathbb{R}} \times \mathbb{R} \urcorner \ncong \ulcorner \mathbb{R}^{\mathbb{R}} \urcorner \times \ulcorner \mathbb{R} \urcorner \right)$$

Random element Space

$R_X := X^{\mathbb{R}}$ since $\lfloor X^{\mathbb{R}} \rfloor = R_X$ as sets.

Why?

($\subseteq$) $\alpha \in \lfloor X \rfloor^{\mathbb{R}} \Rightarrow \alpha : \mathbb{R} \to X$ in Qbs.

$\mathrm{id}_{\mathbb{R}} : \mathbb{R} \to \mathbb{R}$ measurable $\Rightarrow \mathrm{id} \in R_{\mathbb{R}}$

$\Rightarrow \alpha = \alpha \circ \mathrm{id} \in R_X$

Precomposition

($\supseteq$) $\alpha \in R_X \Rightarrow \forall \varphi \in R_{\mathbb{R}} = \mathrm{Meas}(\mathbb{R}, \mathbb{R}).\ \alpha \circ \varphi \in R_X \Rightarrow \alpha : \mathbb{R} \to X$

$\Rightarrow \alpha \in \lfloor X^{\mathbb{R}} \rfloor$

Subspaces

For $X \in \mathbf{Qbs}$, $A \subseteq \lfloor X \rfloor$ set:

$$R_A := \{ \alpha : \mathbb{R} \to A \mid \alpha \in R_X \}$$

Then $A = (A, R_A)$ is the subspace qbs

We write $A \hookrightarrow X$

Borel subspaces ensemble

The σ-algebra $B_X := \{A \subseteq \lfloor X \rfloor \mid \forall \alpha \in R_X . \alpha^{-1}[A] \in B_{\mathbb{R}}\}$

internalises as $B_X = 2^X$, the qbs of

Borel subsets.

($\lfloor B_{(B_{\mathbb{R}})} \rfloor$ are the Borel-on-Borel sets from descriptive set theory.

cf. [Sabok et al.'21])

Standard Borel Spaces

Def: A qbs S is standard Borel when

$$S \cong A \quad \text{for some } A \in \mathcal{B}_{\mathbb{R}}$$

$$\begin{tikzcd}
\mathbf{Meas} \arrow[rr, bend left] \arrow[rr, phantom, "\top"] & & \mathbf{Qbs} \arrow[ll, bend left] \\
\mathbf{Sbs} \arrow[u, hook] \arrow[rr, bend left] \arrow[rr, phantom, "\cong"] & & \mathbf{Sbs} \arrow[u, hook] \arrow[ll, bend left]
\end{tikzcd}$$

Slogan: Qbs conservative extension of Sbs

Example $C_0 := \{ f : \mathbb{R} \to \mathbb{R} \mid f \text{ continuous} \} \hookrightarrow \mathbb{R}^{\mathbb{R}}$

C_0 is sbs. (Well-known!)

Proof: $C_0' \in B_{\mathbb{R}^{\mathbb{Q}}}$ sbs!

$$C_0' := \left\{ g \in \mathbb{R}^{\mathbb{Q}} \;\middle|\; \begin{array}{l} \forall a,b \in \mathbb{Q}, \varepsilon \in \mathbb{Q}^{+} \\ \exists \delta \in \mathbb{Q}^{+}\ \forall p,q \in \mathbb{Q}^{+} \cap [a,b]. \\ |p-q| < \delta \Rightarrow |g\,p - g\,q| < \varepsilon \end{array} \right\}$$

on closed intervals (= compact intervals) Continuity $\Updownarrow$ uniform continuity

Borel measurable by type checks

then $C_0 \cong C_0' \in B_{\mathbb{R}^{\mathbb{Q}}}$:

$$\begin{array}{ll} C_0 \to C_0' & C_0' \to C_0 \\ \varphi \mapsto \varphi|_{\mathbb{Q}} & \varphi \mapsto \lambda r. \lim\limits_{n\to\infty} g(\text{approx}_{(\frac{1}{m})_{m\in\mathbb{N}}} r)_n \end{array}$$

Example (ctd)

C_0 is sbs, and $\mathrm{eval}: C_0 \times \mathbb{R} \to \mathbb{R}$

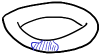
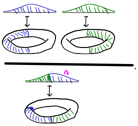
is a measurable.

Avoids:

- constructing complete separable metrics
- proving that evaluation is measurable w.r.t. metric σ-algebra.

Agenda

Slogan: Measurable by Type

- Borel sets 
- Qbs:
  - def., constructions, 
  - partiality, refinement
- Measures & integration
- Random variable spaces
- Conditional expectation

Partiality cf. [Vákár et al. '19]

A Borel embedding $e : X \hookrightarrow Y$

- injective function $e : \lfloor X \rfloor \rightarrowtail \lfloor Y \rfloor$
- its image is Borel: $e[\lfloor X \rfloor] \in \mathcal{B}_Y$
- e is Strong: $\alpha \in R_X \iff e \circ \alpha \in R_Y$

Examples

- $\mathbb{1} \hookrightarrow \mathbb{2}$
- S is sbs $\Leftrightarrow \exists\, S \hookrightarrow \mathbb{R}$

Non-examples $\rightsquigarrow$ [Sabok et al.'21]

- $\{ A \in B_{\mathbb{R}} \mid A \neq \emptyset \} \hookrightarrow B_{\mathbb{R}}$

- $\{ (A_1, A_2) \in B_{\mathbb{R}}^2 \mid A \subseteq B \} \hookrightarrow B_{\mathbb{R}}^2$

- $\{ A \in B_{\mathbb{R}} \mid A \text{ open} \} \hookrightarrow B_{\mathbb{R}}$

Def: A Partial map $f : X \rightharpoonup Y$ is a morphism

$$f : X \to Y \amalg \{\bot\}$$

Its domain of definition $\mathrm{Dom}\, f := \{x \mid fx \neq \bot\}$

$$\begin{array}{c} \mathrm{Dom}\, f \\ \downarrow \\ X \end{array}$$

Partial hom-sets are ordered:

for $f, g : X \rightharpoonup Y$ $\quad f \leq g$ when $\forall x. \ fx \neq \bot \Rightarrow gx = fx.$

[Cockett-Lack '06]

A model of restriction categories / axiomatic domain theory

[Fiore_Plotkin '94] Borel embeddings are the admissible monos

Space refinement

Let $P : X \to (\lfloor 2 \rfloor_{\text{qbs}})^Y$ X-parametrised Property

indiscrete 2-elmt qbs

$$\prod_{x \in X} P_x \hookrightarrow Y^X \qquad \coprod_{x \in X} P_x \hookrightarrow X \times Y$$

$$\prod_{x \in X} P_x := \{ f \in Y^X \mid \forall x \in X . f x \in P_x \} \qquad \coprod_{x \in X} P_x := \{ (x, y) \mid y \in P_x \}$$

When P factors as $P : X \xrightarrow{Q} 2^Y \xrightarrow{\lambda U. U} \lfloor 2 \rfloor_{\text{qbs}}^Y$,

Write $\prod_{x \in X} Q_x$ $\coprod_{x \in X} Q_x$ for the same spaces

Example

$(\Omega \in \mathbf{Qbs})$

Converging $\hookrightarrow \left([-\infty,\infty]^{\Omega}\right)^{\mathbb{N}} \cong \left([-\infty,\infty]^{\mathbb{N}}\right)^{\Omega}$

$$\text{Converging} \cong \prod_{\omega \in \Omega} \left\{ \vec{f} \;\middle|\; \exists \lim_{n\to\infty} f_n \omega \right\}$$

Refinement not dependent types

$\prod_{X} P$ require all $P_x \hookrightarrow Y^X$ $\rightsquigarrow$ independently of x

Qbs can interpret dependent types, but such ensemble spaces require a to-be-determined universe.

An introduction to statistical modelling semantics with higher-order measure theory

Ohad Kammar
University of Edinburgh

Day 2

Logic of Probabilistic Programming
Logique de la programmation probabiliste
31 January–4 February, 2022
Logic and Interactions — Logique et interactions
CIRM Thematic Month

supported by: THE ROYAL SOCIETY The Alan Turing Institute Facebook Research NCSC

http://www.denotational.co.uk

Slides

Talks (abstracts)

- An introduction to statistical modelling semantics with higher-order measure theory, Logic of Probabilistic Programming, (exercises coming soon), 03 February, 2022.

Hopefully!

Spaces

~~Measurable~~

$\uparrow$

Continuous

$\uparrow$

Discrete countable

$\uparrow$

Discrete finite

} standard Borel

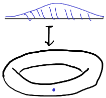
Quasi-Borel

$$X = (\lfloor X \rfloor, R_X)$$

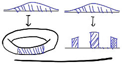
$\lfloor X \rfloor$: Set of elements

R_X: Subset of random elements $\alpha : \mathbb{R} \to \lfloor X \rfloor$

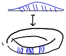
- Constants:

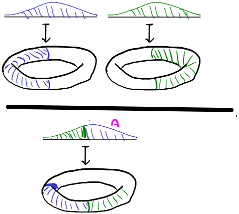


- Precomposition:



- recombination



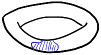
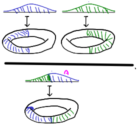


Qbs:

- (in)discrete "X", $\lfloor X \rfloor$
- Product $X \times Y$, coproducts $X + Y$, function Y^X
- Subspaces $A \hookrightarrow X \hookleftarrow V$
- refinement $\to \Pi, \coprod$
- internalisation R_X, B_X
- Partiality $\leadsto X \rightharpoonup Y$

Agenda

Slogan: Measurable by Type

- Borel sets 
- Qbs:
  def., constructions, 
  partiality, refinement $\to$, $\Rightarrow$, $\coprod$, $\prod$
- Measures & integration
- Random variable spaces
- Conditional expectation

Def: A measure μ over $\mathbb{R}$ is a function

$$\mu : B_{\mathbb{R}} \to W := [0, \infty]$$

s.t. - $\mu \emptyset = 0$

$$\frac{-\quad \vec{A} \in B_{\mathbb{R}}^{\mathbb{N}} \qquad \begin{matrix} A_n \cap A_m = \emptyset \\ (n \neq m) \end{matrix}}{\mu\left(\bigcup_{n=0}^{\infty} A_n\right) = \sum_{n=0}^{\infty} \mu A_n}$$

For measurable spaces, replace $\mathbb{R}$ with V

We write $\lfloor G V \rfloor$ for the set of measures on V

For qbs X, take $\lfloor G \ulcorner X^{\text{Meas}} \urcorner \rfloor$

The unrestricted Giry space

Equip $\lfloor GV \rfloor$ with

$$R_{GV} := \{ \alpha : \mathbb{R} \to GV \mid \forall A \in B_V . \lambda r . \alpha(r, A) : \mathbb{R} \to \mathbb{W} \}$$

α is a kernel.

Farewell Meas

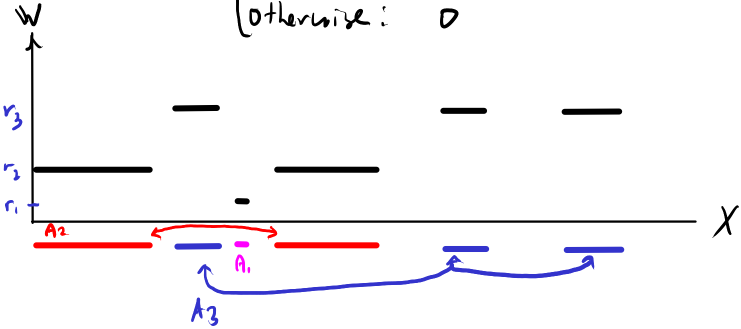
Now on:

1. All spaces are quasi-Borel
2. "measurable function" means qbs morphism!

Def: Simple function $\varphi : X \to W$ when

$\exists\, n \in \mathbb{N},\ \vec{A} \in B_X^n,\ A_i \cap A_j = \emptyset\ (i \neq j),\ \vec{r} \in W$ s.t.

$$\varphi x = \begin{cases} \vdots & \\ x \in A_i : & r_i \\ \vdots & \\ \text{otherwise} : & 0 \end{cases}$$



Encode into a space:

$$\mathrm{SimpleCode} := \coprod_{n \in \mathbb{N}} B_X^n \times W^n$$

$$\mathrm{Simple} := \{ f \in W^X \mid f \text{ simple} \} \hookrightarrow W^X$$

and define an interpretation:

$$[\![-]\!] : \mathrm{SimpleCode} \twoheadrightarrow \mathrm{Simple}$$

$$[\![(n, \vec{A}, \vec{r})]\!] := \sum_{i=1}^{n} r_i \cdot [- \in A_i]$$

↳ characteristic function for A_i

Lemma: $f : X \to W$ is measurable

iff $f = \lim_{n \to \infty} f_n$ for some monotone sequence

$\vec{f} \in \mathrm{Simple}$.

remember! QBS morphism!

Moreover, we have measurable such choice:

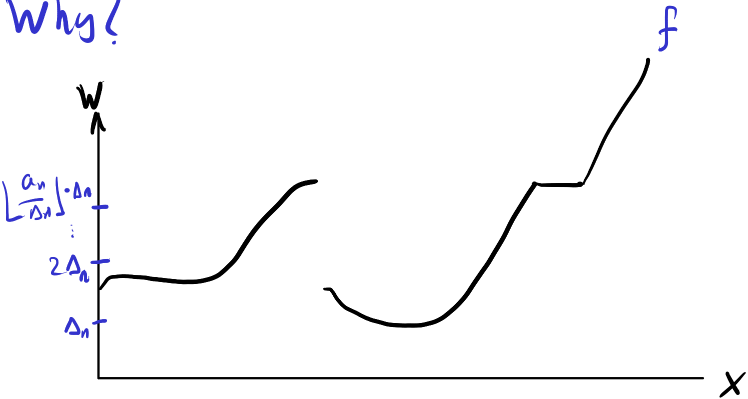
Simple Approx :

$$\{\vec{\Delta} \in \mathbb{R}^+ \mid \Delta_n \to 0\} \times \{\vec{a} \in W^{\mathbb{N}} \mid \begin{array}{l}\vec{a} \text{ monotone} \\ a_n \to \infty\end{array}\} \times W^X \longrightarrow \mathrm{SimpleCode}$$

rate of convergence

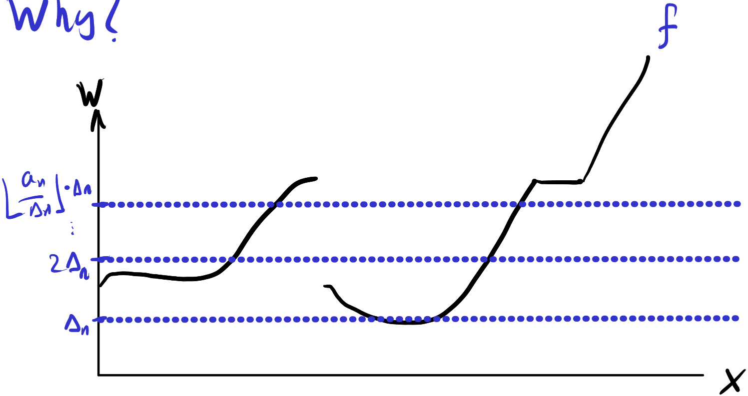
range of approximation

Why?



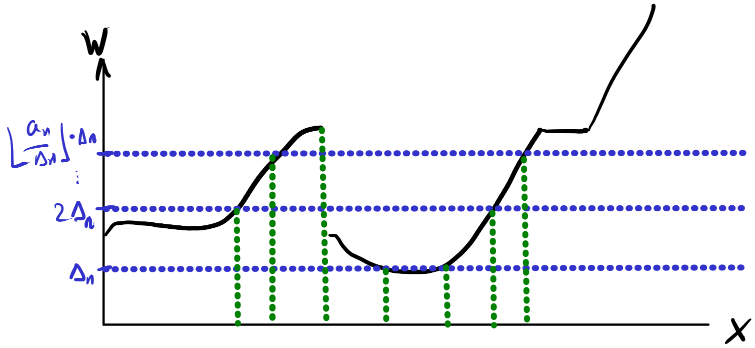
f
W
$\lfloor \frac{a_n}{\Delta_n} \rfloor \cdot \Delta_n$
$2\Delta_n$
Δ_n
X

Why?



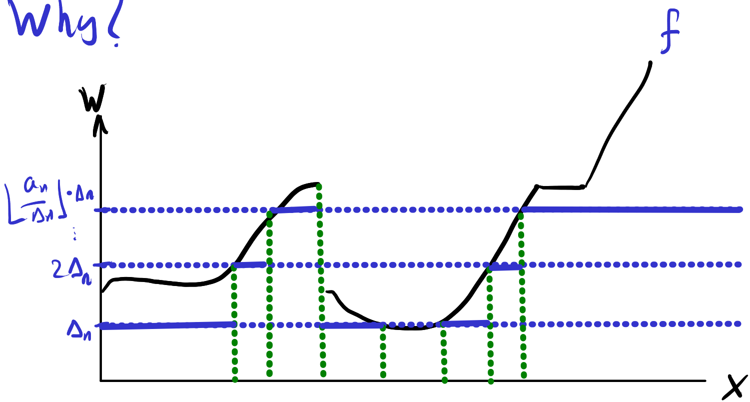
f
W
$\lfloor \frac{a_n}{\Delta_n} \rfloor \cdot \Delta_n$
$\vdots$
$2\Delta_n$
Δ_n
X

Why?



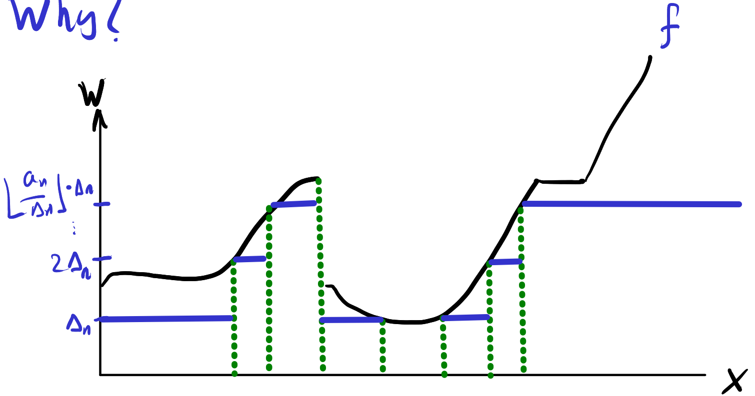
f
W
X
$\lfloor \frac{a_n}{\Delta_n} \rfloor \cdot \Delta_n$
$2\Delta_n$
Δ_n

Why?

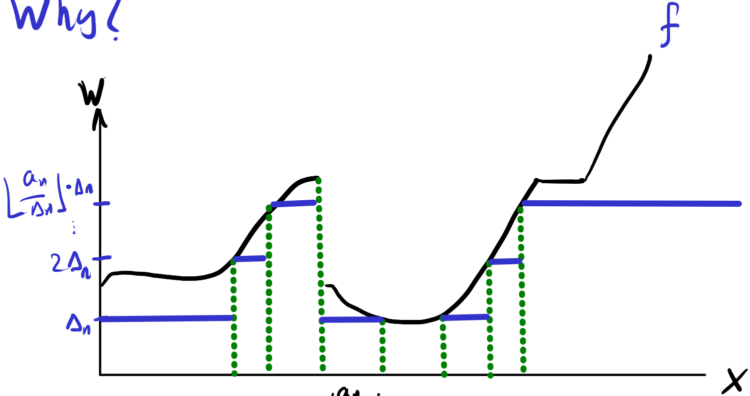


f
W
$\lfloor \frac{a_n}{\Delta_n} \rfloor \cdot \Delta_n$
$2\Delta_n$
Δ_n
X

Why?

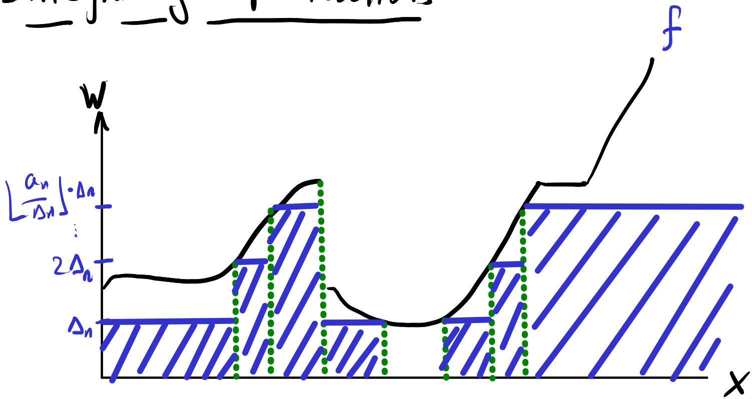

W
f
X
$\lfloor \frac{a_n}{\Delta_n} \rfloor \cdot \Delta_n$
$2\Delta_n$
Δ_n

Why?



$$[\![\mathrm{SimpleApprox}_{\vec{\Delta},\vec{a}}\, f]\!] := \sum_{i=1}^{\lfloor \frac{a_n}{\Delta_n} \rfloor} i\cdot\Delta_n \left[i\cdot\Delta_n \le f < (i+1)\Delta_n \right] + \left\lfloor \frac{a_n}{\Delta_n} \right\rfloor \Delta_n \cdot \left[f \geq \left\lfloor \frac{a_n}{\Delta_n} \right\rfloor \cdot \Delta_n \right] \in \mathrm{Simple}$$

Integrating Simple Functions



$$\int : GX \times \text{SimpleCode} \to W$$

$$\int \mu\,(n, \vec{A}, \vec{r}) := \sum_{I \subseteq \{1,\ldots,n\}} \Big(\sum_{i \in I} r_i\Big) \cdot \mu\Big(\bigcap_{i \in I} A_i \setminus \bigcup_{i \notin I} A_i\Big)$$

Integration

$\rightsquigarrow$ Properly higher-order operation

$$\int : GX \times \mathbb{W}^X \longrightarrow \mathbb{W}$$

$$\int \mu f := \sup\left\{ \int \mu \varphi \;\middle|\; \varphi \in \mathrm{Simple},\ \varphi \leq f \right\}$$

we also write $\int \mu(dx)t$ for $\int \mu(\lambda x; t)$

$$= \lim_{n\to\infty} \int \mu\left(\mathrm{SimpleApprox}_{\vec{\Delta},\vec{a}}\, f\right)_n$$

$\rightsquigarrow$ measurable by type

for $\frac{a_n}{\Delta_n} \to 0$, eg. $\Delta_n = \frac{1}{2^n}$ $a_n = n$.

resolution

The unrestricted Giry Strong Monad

Unlike the unrestricted Giry on Meas.

Dirac:

$$\delta : X \to GX$$
$$x \mapsto \lambda A. \begin{cases} x \in A : 1 \\ x \notin A : 0 \end{cases}$$

Kleisli extension/Kock integral:

$$\oint : GX \times GY^X \longrightarrow GY$$

$$\oint \mu f := \lambda A. \int \mu(\mathrm{d}x)\, f\, x\, A$$

but: non-commutative

(Fubini fails, just like in Meas)

Randomisable measures monad

$$\lambda A. \int_{\mathrm{Dom}\,\alpha} \lambda\ \alpha^{-1}[A]$$

$$D \rightarrowtail G$$

$$\lfloor DX \rfloor := \{ \alpha_* \lambda \mid \alpha : \mathbb{R} \to X \}$$

Lebesgue measure

$$M_{DX} := \{ \lambda x. (\alpha x)_* \lambda \mid \alpha : \mathbb{R} \times \mathbb{R} \to X \}$$

D is commutative (Fubini's Theorem)

$\mu \in DX$, $\nu \in DY$:

$$\oint \mu(dx) \oint \nu(dy)\, \delta_{(x,y)} = \oint \nu(dy) \oint \mu(dx)\, \delta_{(x,y)} =: \mu \otimes \nu$$

Models Kock's Synthetic measure theory [Kock'12, Ścibior et al.'18]

Distribution Submonads

A measure space

$\Omega = (\Omega, \mu)$

is a qbs Ω with

$\mu \in DX$.

Similarly: finite measure space

- (sub) probability space.

$$
\begin{array}{c}
PX := \{\mu \in DX \mid \mu X = 1\} \\
\downarrow \\
P_{\leq 1} X := \{\mu \in DX \mid \mu X \leq 1\} \\
\downarrow \\
P_{<\infty} X := \{\mu \in DX \mid \mu X < \infty\} \\
\downarrow \\
DX
\end{array}
$$

Thm: For sbs S, PS, $D_{\le 1}S$, $D_{<\infty}S \in Sbs$

and agree with their counterparts on Meas.

$|DS| = \{\mu \mid \mu \text{ s-finite}\}$

See [Staton'16]

$R_{DS} = \{k : \mathbb{R} \to GD \mid k \text{ s-finite kernel}\}$

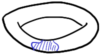
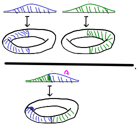
Open: Is there a conterpart to D in Meas?

More modestly, is $DS \in Sbs$?

(Hypothesis: no)

Agenda

Slogan: Measurable by Type

- Borel sets 
- Qbs:
  def., constructions, 
  partiality, refinement $\rightharpoonup, \Rightarrow, \coprod, \prod$
- Measures & integration $\delta, \int, \fint, D, \otimes, P$
- Random variable spaces
- Conditional expectation

Random variable: $\xi : \Omega \to \Theta \hookrightarrow \overline{\mathbb{R}}$

Θ:

$$\begin{tikzcd}
 & & \overset{[0,\infty]}{W} \arrow[dr] & \\
\overset{[0,1]}{\mathbb{I}} \arrow[r] & \overset{[0,\infty)}{\mathbb{R}_{\geq 0}} \arrow[ur] \arrow[dr] & & \underset{[-\infty,\infty]}{\overline{\mathbb{R}}} \\
\underset{(0,\infty)}{\mathbb{R}^+} \arrow[ur] & & \underset{(-\infty,\infty)}{\mathbb{R}} \arrow[ur] &
\end{tikzcd}$$

- Θ^Ω is a space
- $\mathbb{R}^\Omega$ measurable vector space: $\alpha\xi + \zeta := \lambda\omega.\ \alpha\cdot\xi\omega + \zeta\omega$
- W^Ω measurable σ-semi-module for W: $\sum_{n=0}^{\infty} \alpha_n \xi_n := \lambda\omega.\ \sum_{n=0}^{\infty} \alpha_n \cdot \xi_n$

$\Pr_- : P\Omega \times B_\Omega \to W$

$\Pr_\lambda A := \mathrm{eval}(\lambda, A) = \lambda A$

Probability Space $\Omega = (\Omega, \lambda_\Omega)$

"$P\,x$ holds $\lambda(dx)$-almost surely" $(P \hookrightarrow \Omega)$

for some $Q \hookrightarrow \Omega$, $P \supseteq Q$, $\Pr_\lambda Q^c = 0$

so $\Pr Q = 1$

Example $(\xi, \zeta \in \mathbb{H}^\Omega)$

$\xi = \zeta$ a.s. when $\Pr_{\omega \sim \lambda}[\xi\,\omega \neq \zeta\,\omega] = 0$

Integrating random variables

$(-)_+, (-)_- : \overline{\mathbb{R}}^\Omega \longrightarrow W^\Omega$ in Qbs!

$\xi_+ := \max(\xi, 0) \quad \xi_- := \max(-\xi, 0)$

So: $\xi = \xi_+ - \xi_-$

$$\int : P\Omega \times W^\Omega \longrightarrow W$$

$$\int \lambda\, \xi := \int \lambda\, \xi_+ - \int \lambda\, \xi_-$$

$\int$ respects a.s. equality:

$\xi = \zeta\ (\lambda\text{-a.s})$

$\Rightarrow \int \lambda\, \xi = \int \zeta.$

Example

$$\mathrm{ASConverge}(\overline{\mathbb{R}})^{\Omega} := \left\{ \vec{z} \in \overline{\mathbb{R}}^{\mathbb{N}\times\Omega} \;\middle|\; \Pr_{\omega \sim \lambda}\left[\lim z_n \omega \neq \bot\right] \right\}$$

$$\begin{array}{c} \cap \\ \downarrow \\ \overline{\mathbb{R}}^{\mathbb{N}\times\Omega} \end{array}$$

So:

$$\overset{as}{\lim} : \overline{\mathbb{R}}^{\mathbb{N}\times\Omega} \longrightarrow \overline{\mathbb{R}}^{\Omega} \qquad \mathrm{Dom}\, \overset{as}{\lim} := \mathrm{ASConverge}(\overline{\mathbb{R}})^{\Omega}$$

$$\overset{as}{\lim} \vec{z} := \lambda \omega.\ \limsup_{n\to\infty} f_n\, \omega$$

$\llcorner$ $\overset{as}{\lim}$ respects a.s. equality.

Thm (monotone convergence):

Let $\vec{\xi} \in \mathbb{W}^{\mathbb{N}\times\Omega}$ λ-a.s. monotone.

$$\xi = \lim^{as}_{n\to\infty} \xi_n \qquad (a.s.)$$

$$\Downarrow$$

$$\int \lambda\, \xi = \lim_{n\to\infty} \int \lambda\, \xi_n$$

Lebesgue space $\left(\begin{array}{l}\Omega \text{ prob. space,} \\ p \in [1,\infty)\end{array}\right)$

$$\mathcal{L}^p_\Omega := \left\{ \xi \in \mathbb{R}^\Omega \,\middle|\, \int \lambda |\xi|^p < \infty \right\} \hookrightarrow \mathbb{R}^\Omega$$

Ensemble

$$\mathcal{L}_\Omega := \prod_{\substack{\lambda \in P\Omega \\ p \in [1,\infty)}} \mathcal{L}^p_{(\Omega,\lambda)} \hookrightarrow B_{\mathbb{R}^\Omega}^{P\Omega \times [1,\infty)}$$

$\llcorner$ $p \leq q \Rightarrow \mathcal{L}^p_\Omega \supseteq \mathcal{L}^q_\Omega$

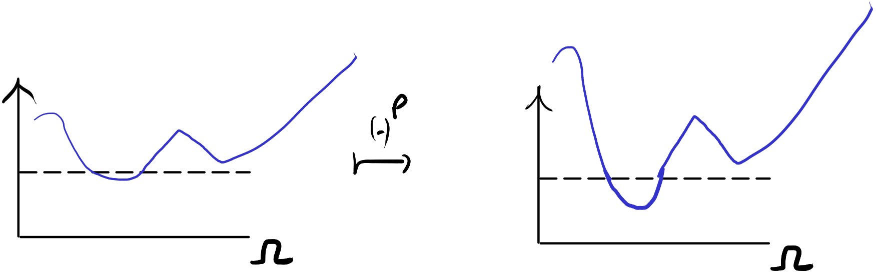
$\mathcal{L}^p$ Semi norms

$\|-\| : \coprod_{p,\lambda} \mathcal{L}^p_{(\Omega,\lambda)} \longrightarrow \mathbb{R}_{\geq 0}$ $\quad \|\xi\|_p := \sqrt[p]{\int \lambda |\xi|^p}$

$\mathcal{L}^2$ inner product

$\langle -, - \rangle : \coprod_{p,\lambda} \mathcal{L}^p_{(\Omega,\lambda)} \times \mathcal{L}^p_{(\Omega,\lambda)} \longrightarrow \mathbb{R}$

$\langle \xi, \zeta \rangle_{p,\lambda} := \int \lambda \, \xi \zeta$



Statistics

Expectation

$$\mathbb{E} : \coprod_\lambda \mathcal{L}^1 \to \mathbb{R}$$

$$\mathbb{E}_\lambda \xi := \int \lambda \xi$$

Covariance and Correlation

$$\mathrm{Cov}, \mathrm{Corr} : \coprod_\lambda \mathcal{L}^2 \to \mathbb{R}$$

$$\mathrm{Cov}(\xi, \zeta) := \langle \xi - \mathbb{E}\xi, \zeta - \mathbb{E}\zeta \rangle$$

$$\mathrm{Corr}(\xi, \zeta) := \frac{\langle \xi, \zeta \rangle}{\|\xi\|_2 \cdot \|\zeta\|_2} = \cos(\mathrm{angle}(\xi, \zeta))$$

Sequential limits

$\text{Cauchy}\,\mathcal{L}^p_\Omega \hookrightarrow (\mathcal{L}^p)^{\mathbb{N}}$

$$\text{Cauchy}\,\mathcal{L}^p_\Omega := \left\{ \vec{\Sigma} \;\middle|\; \forall \varepsilon \in \mathbb{Q}^+\, \exists k \in \mathbb{N}\, \forall m,n \geq k.\ \| \Sigma_{k+n} - \Sigma_{k+m} \|_p < \varepsilon \right\}$$

Thm: $\mathcal{L}^p_\Omega$ is Cauchy-complete

$$\lim^p : \text{Cauchy}\,\mathcal{L}^p \longrightarrow \mathcal{L}^p$$

Why?

1. Every Cauchy sequence has an a.s. converging subseq.
2. We can find it measurably

Example

Thm (dominated convergence)

For $\vec{\xi}, \zeta \in \mathcal{L}^1$ s.t. $\xi_n \leq \zeta$ a.s.:

1. $\overset{as}{\lim} \vec{\xi} \in \mathcal{L}^1$
2. $\overset{1}{\lim} \vec{\xi} = \overset{as}{\lim} \vec{\xi}$
3. $\lim_{n \to \infty} \int \lambda\, \xi_n = \int \lambda \lim_{n \to \infty} \xi_n$

Separability

Def: $\mathcal{L}^p$ separable: has countable dense subset

Fact: Separability is property of λ_{Ω}:

TFAE:

- $\exists p \geq 1.\ \mathcal{L}^p$ separable
- $\forall p \geq 1.\ \mathcal{L}^p$ separable

Measurable separability in $\mathcal{I} \hookrightarrow P\Omega \times [1,\infty)$

$$\vec{\beta} : \prod_{(\lambda,p)\in\mathcal{I}} {\mathcal{L}^{p}_{(\Omega,\lambda)}}^{\mathbb{N}} \quad \text{s.t.}$$

$$\{ \vec{\beta}^{\lambda,p}_{n} \mid n \in \mathbb{N} \} \text{ dense in } \mathcal{L}^{p}_{(\Omega,\lambda)}$$

Prop. - every sbs S measurably separable in $PS \times [1,\infty)$

- $\mathcal{I} \hookrightarrow P\Omega \times \{2\}$ measurably separable

$\Rightarrow \exists \vec{\beta} \in \prod_{\lambda\in\mathcal{I}} \mathcal{L}^{2}_{(\Omega,\lambda)}$ orthonormal system

$\langle \beta_n, \beta_m \rangle = 0$

$\|\beta_n\|_2 = 1$

(β_n) dense

Example

Let $S \hookrightarrow \mathcal{L}^2$ closed vector subspace.

orthogonal decomposition

linear in fact.

$$\langle P, P^{\perp}\rangle : \lfloor \mathcal{L}^{P} \rfloor \to \lfloor S \rfloor \times \lfloor S^{\perp} \rfloor$$

When S is separable with orthonormal system β

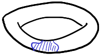
We have a measurable version of

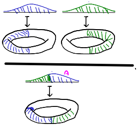
$\langle P, P^{\perp}\rangle : \mathcal{L}^2 \to S \times S^{\perp}$

$$P\xi := \sum_{n=0}^{\infty} \langle \xi, \beta_n \rangle \beta_n \qquad P^{\perp} := \mathrm{Id} - P.$$

Agenda

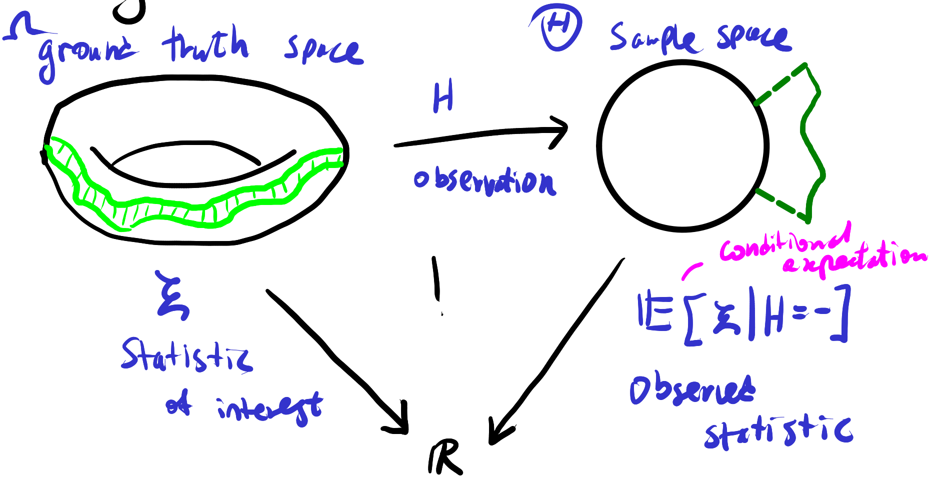
Slogan: Measurable by Type

- Borel sets 
- Qbs:

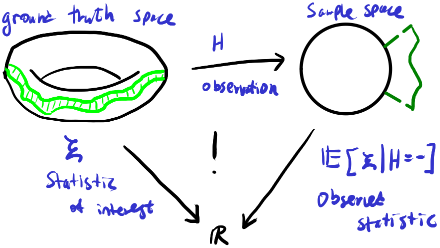
  

  def., constructions,
  partiality, refinement $\rightharpoonup, \hookrightarrow, \coprod, \prod$
- Measures & integration $\delta, \int, ⨏, D, \otimes, P$
- Random variable spaces $\mathcal{L}^p, \|-\|_p, \mathbb{E}$
- Conditional expectation

Kolmogorov's Conditional Expectation



Kolmogorov's Conditional Expectation



A conditional expectation of $\xi \in \mathcal{L}^1_\Omega$ wrt $H : \Omega \to \Theta$ is $Z \in \mathcal{L}^1_\Theta$ s.t. for all $A \in \mathcal{B}_\Theta$:

$$\int_A \mu Z = \int_{H^{-1}[A]} \lambda \xi$$

where $\mu := H_* \lambda$

Conditional expectations

1. unique a.s.

2. fundamental to modern probability, eg:

   a martingale

$$
\begin{tikzcd}
\cdots & \Theta_n \arrow[l] \arrow[dr, "\xi_n"'] & {} \arrow[d, phantom, "!"] & \Theta_{n+1} \arrow[ll, "H_n"'] \arrow[dl, "\xi_{n+1}"] & \cdots \arrow[l] \\
 & & \mathbb{R} & &
\end{tikzcd}
$$

s.t. $\xi_n = \mathbb{E}[\xi_{n+1} \mid H_n = -]$

Thm (Existence)

- $\exists$ $\mathbb{E}[- \mid H = -] : \lfloor \mathcal{L}^1_{(\Omega,\lambda)} \rfloor \longrightarrow \mathcal{L}^1_{(\Theta,\mu)}$

- When (Ω,λ) is Separable

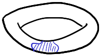
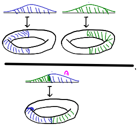
$$\mathbb{E}[- \mid H = -] : \mathcal{L}^1_{(\Omega,\lambda)} \longrightarrow \mathcal{L}^1_{(\Theta,\mu)}$$

- When Θ is $\mathbb{I}$-measurably separable

$$\mathbb{E}_{-}[- \mid - = -] : \prod_{\substack{H \in \Theta^{\Omega} \\ \lambda \in H^{-1}_{*}[\mathbb{I}]}} \mathcal{L}^1_{(\Omega,\lambda)} \longrightarrow \mathcal{L}^1_{(\Theta,\mu)}$$

Agenda

Slogan: Measurable by Type

- Borel sets 
- Qbs: 

  def., constructions, partiality, refinement $\to$, $\rightharpoonup$, $\coprod$, $\prod$
- Measures & integration δ, $\int$, $⨏$, D, $\otimes$, P
- Random variable spaces $\mathcal{L}^p$, $\|-\|_p$, $\mathbb{E}$
- Conditional expectation $\mathbb{E}[-|-=-]$