Foundations for type-driven probabilistic modelling

Ohad Kammar
University of Edinburgh

Logic Summer School
Australian National University
4–16 December, 2023
Canberra, ACT, Australia

supported by: The Royal Society, The Alan Turing Institute, Facebook Research, NCSC

Computational golden era of:

logic & type rich computation

Statistical computation

Computational golden era of:

logic & type rich computation

expressive type systems:

Haskell, OCaml, Idris

mechanised mathematics:

Agda, Coq, Isabelle/HoL, Lean

Verification:

SMT-powered, realistic systems

Statistical computation

generative modelling

+

efficient inference:

Monte-Carlo simulation

or gradient-based optimisation

"AI"

Computational golden era of:

logic & type rich computation

Statistical computation

Clear connection to Foundations:

- Robert's
- John's
- Michael's
- Dominik's

courses

- this course

Why foundations?

countably supported measures

Discrete Probability

Continuous Probability

Lebesgue measure on $\mathbb{R}^n$

Well-behaved Probability

s-finite measures over standard Borel spaces

Measurable spaces

standard, established

Poor type structure

Quasi-Borel spaces

new, experimental

rich type structure

(this course)

Takeaway:

use types to develop smoothly

Plan:

1) type-driven probability: discrete case (Mon + Tue (?))

2) Borel sets & measurable spaces (Tue)

3) Quasi Borel spaces, simple type structure (Wed)

4) Dependent type structure & standard Borel spaces (Thu)

5) Integration & random variables (Fri)

Please ask questions!

Course web page

Scribble

Advertisement

Interested in the computational / mathematical / logical foundations of Computer science?

Hedging your bets on a funded PhD offer from ANU?

Check out our PhD Programs:

THE UNIVERSITY OF EDINBURGH

lfcs | Laboratory for Foundations of Computer Science

LFCS PhD

Dependable AI for Robotics CDT

Langauge of distribution & Probability

X type (= space) of values / outcomes

DX type of distributions / measures over X

$PX \subseteq DX$ Sub type of probability measures (total measure 1)

BX type of measurable events – subsets of X we wish to measure

W type of weights : $[0, \infty]$

type judgement

$$\mu : DX, E : BX \vdash \mathcal{C}e_{\mu}[E] : W$$

↳ measure μ assigns to E

Axioms for measures

Empty event: $\emptyset : \mathbf{B}X$

Its measure is $0 : \mathbf{W}$:

$$\mu : \mathbf{D}X \vdash \underset{\mu}{\mathbb{C}}e[\emptyset] = 0 \quad : \mathbf{W}$$

Axioms for measures

BX is a Boolean sub-algebra:

$$E : BX \vdash E^c : BX$$

$$E, F : BX \vdash E \cup F, E \cap F : BX$$

$E, C : BX, \mu : DX \vdash$ (disjoint additivity)

$$\underset{\mu}{\mathcal{C}e}[E] = \underset{\mu}{\mathcal{C}e}[E \cap C] + \underset{\mu}{\mathcal{C}e}[E \cap C^c] : W$$

Axioms for measures

$\omega := (\mathbb{N}, \leq)$ $\quad (\mathbb{B}, \subseteq)$ $\quad (\mathbb{W}, \leq)$ posets

$$(BX, \subseteq)^{\omega} := \{ (E_n)_{n\in\mathbb{N}} \in (BX)^{\mathbb{N}} \mid E_0 \subseteq E_1 \subseteq E_2 \subseteq \cdots \}$$

$(BX, \subseteq)$ and $(\mathbb{W}, \leq)$ are ω-chain-closed:

$$E_{-} : (BX, \subseteq)^{\omega} \vdash \bigcup_n E_n : BX \qquad a_{-} : (\mathbb{W}, \leq)^{\omega} \vdash \sup_n a_n : \mathbb{W}$$

$E_{-} : (BX, \subseteq)^{\omega}, \mu : DX \vdash$ (Scott continuity)

$$\mathrm{Ce}_{\mu}\left[\bigcup_n E_n\right] = \sup_n \mathrm{Ce}_{\mu}\left[E_n\right] \quad : \mathbb{W}$$

Axiom for Probability

$$\mathrm{Cost} : \mathrm{P}X \overset{\subseteq}{\longleftrightarrow} \mathrm{D}X \qquad 1 : \mathbb{W}$$

$$\mu : \mathrm{P}X \vdash \underset{\mathrm{Cost}\ \mu}{\mathrm{Ce}}[X] = 1 \ : \mathbb{W}$$

Avoid costing:

$$E : \mathrm{B}X, \mu : \mathrm{P}X \vdash \underset{\mu}{\mathrm{Pr}}[E] := \underset{\mathrm{Cost}\ \mu}{\mathrm{Ce}}[E] \ : \ [0,1] \subseteq \mathbb{W}$$

Axioms for measures

Integration:

$$\mu : DX,\ \varphi : W^X \vdash \int \mu\, \varphi \quad : W \qquad \text{(Lebesgue integral)}$$

Again, avoid casting:

$$\mu : PX,\ \varphi : W^X \vdash \mathop{\mathbb{E}}_{\mu}[\varphi] := \int (\mathrm{cast}\ \mu)\, \varphi \quad : W$$

(Expectation)

More structure & notation later (...technical...)

Have: language + axioms

Want: model

today: discrete measures

rest of course: discrete + continuous

Discrete model

type X : set

$$DX := \{\mu : X \to \mathbb{W} \mid \mu \text{ is countably supported}\}$$

(next slide)

Support

Powerset

$$\mu : \mathbb{W}^X, S : \mathcal{P}X \vdash S \text{ supports } \mu :=$$

$$\forall x : X.\ \mu x > 0 \Rightarrow x \in S \quad : \mathrm{Prop}$$

$$\mu : \mathbb{W}^X \vdash \mathrm{supp}\,\mu := \{x \in X \mid \mu x > 0\} \quad : \mathcal{P}X$$

$\mathrm{supp}\,\mu$ is the smallest set supporting μ

Discrete model

type X : set

$$DX := \{\mu : X \to \mathbb{W} \mid \mu \text{ is countably supported}\}$$

$$:= \{\mu : X \to \mathbb{W} \mid \operatorname{supp} \mu \text{ is countable}\}$$

Ex. measures

- X ctbl. Counting measure $\#_X : DX$

  $\#_X := \lambda x : X . 1$ (NB: $\mathrm{supp}\,\#_X = X$ $\checkmark$ ctbl)

- Dirac measure:

  $$x : X \vdash \delta_x := \lambda x'. \begin{cases} x = x' : & 1 \\ \text{o.w.} : & 0 \end{cases} : DX$$

  NB: $\mathrm{supp}\,\delta_x = \{x\}$ $\checkmark$ ctbl

- Zero measure $\underline{0} := \lambda x . 0 : DX$

  NB: $\mathrm{supp}\,\underline{0} = \emptyset$ $\checkmark$ ctbl

Discrete model

type X : set

$$DX := \{\mu : X \to W \mid \mu \text{ is countably supported}\}$$

$$\mu : DX, E : BX \vdash \underset{\mu}{Ce}[E] := \sum_{x \in E} \mu x$$

$$:= \sum_{x \in E \cap \operatorname{Supp} \mu} \mu x$$

Lemma: $\mu : DX$, $S \in \mathcal{P}_{ctbl} X$, S supports μ, $E : BX \vdash$

$$\underset{\mu}{Ce}[E] = \sum_{x \in E \cap S} \mu x$$

Ex:

- $E : \mathbb{B}X \vdash \underset{\#_X}{\mathcal{C}_e}[E] = |E| := \begin{cases} E \text{ has } n \text{ elements: } & n \\ E \text{ infinite: } & \infty \end{cases}$

- $E : \mathbb{B}X, x : X \vdash \underset{\delta_x}{\mathcal{C}_e}[E] = \begin{cases} x \in E : & 1 \\ x \notin E : & 0 \end{cases} =: [x \in E] : W$

NB: $E : \mathbb{B}X \vdash [- \in E] : X \to W$

indicator function

- $E : \mathbb{B}X \vdash \underset{\underline{0}}{\mathcal{C}_e}[E] = 0$

Validate axioms

$$\mu : DX \vdash \underset{\mu}{\mathrm{Ce}}[\emptyset] = 0 \quad : W$$

$$E, C : BX, \mu : DX \vdash$$

$$\underset{\mu}{\mathrm{Ce}}[E] = \underset{\mu}{\mathrm{Ce}}[E \cap C] + \underset{\mu}{\mathrm{Ce}}[E \cap C^c] \quad : W$$

$$E_- : (BX, \subseteq)^\omega, \mu : DX \vdash$$

$$\underset{\mu}{\mathrm{Ce}}\left[\bigcup_n E_n\right] = \sup_n \underset{\mu}{\mathrm{Ce}}[E_n] \quad : W$$

Kernels κ from Γ to X:

$$\kappa : (DX)^{\Gamma}$$

kernels are "open/parameterised" measures

Ex: Dirac kernel: $\delta_{\_} : (DX)^{X}$

Kock Integral

$$\mu : D\Gamma, \kappa : DX^{\Gamma} \vdash \oint \mu \kappa : DX$$

In discrete model:

$$\oint \mu \kappa := \lambda x{:}X. \sum_{\gamma \in \Gamma} \mu\gamma \cdot \overbrace{k(\gamma; x)}^{:= \kappa\gamma x}$$

(Weak) disintegration problem:

Input: $\mu : D\Gamma$ $\quad \nu : DX$

Output: a kernel $k : (DX)^{\Gamma}$ s.t.

$$\oint \mu k = \nu$$

Call such k $\underline{\underline{a}}$ (weak) disintegration of ν w.r.t. μ.

(non-standard terminology)

Ex disintegration:

$\underline{n} := \{0, 1, 2, \ldots, n-1\}$

disintegrate $\#_{\underline{2}^{\underline{n+1}}}$ w.r.t. $\#_{\underline{2}}$

$k : (D(\underline{2}^{\underline{n+1}}))^{\underline{2}}$

$k(x; f) := \begin{cases} f(n) = x : & 1 \\ \text{o.w.} : & 0 \end{cases}$

$(\oint \#_{\underline{2}} k) f = \sum_{x \in \underline{2}} \overbrace{\#_{\underline{2}} x}^{1} \cdot k(x; f)$

$\boxed{\text{NB: } \mathrm{supp}(kx) \\ \checkmark \quad |c + b|}$

$= k(0; f) + k(1; f) = k(fn; f) = 1 = \#_{\underline{2}^{\underline{n+1}}}(f)$

Probability measures

$$PX := \{\mu : DX \mid \oint_{\mu}[X] = 1\} \xhookrightarrow{\subseteq} DX$$

Lemma: $\delta_- : X \to DX$ and $\oint : D\Gamma \times (DX)^{\Gamma} \to DX$

lift along the inclusion $\mathrm{cast} : P \xhookrightarrow{\subseteq} D$:

$$\begin{tikzcd}
 & PX \arrow[d, hook, "\mathrm{cast}"] \\
X \arrow[ur, dashed, "\delta_-"] \arrow[r, "\delta_-"'] \arrow[ur, phantom, shift right=3ex, "="] & DX
\end{tikzcd}
\qquad
\begin{tikzcd}
P\Gamma \times (PX)^{\Gamma} \arrow[r, dashed, "\oint"] \arrow[d, hook, "\mathrm{cast} \times (\mathrm{cast})"'] \arrow[dr, phantom, "="] & PX \arrow[d, hook, "\mathrm{cast}"] \\
D\Gamma \times (DX)^{\Gamma} \arrow[r, "\oint"'] & DX
\end{tikzcd}$$

Prop (discrete Giry): (Michèle Giry '82)

$(D, \delta_{-}, \oint)$ is a monad i.e.

$\mu : \Gamma, k : (DX)^{\Gamma} \vdash \oint \delta_{\mu} k = k\mu$

$\mu : DX \vdash \oint \mu(dx) \delta_{x} = \mu$: DX

$\mu : D\Gamma, k : (DX)^{\Gamma}, t : (DY)^{X} \vdash$

$$\oint \mu(d\gamma) \left(\oint (k\gamma) t \right) = \oint \left(\oint \mu k \right)(dx)\, t(x)$$

Corollary: $(P, \delta_{-}, \oint)$ is a monad.

Language of distribution & Probability

X type (=space) of values / outcomes

DX type of distributions / measures over X

$PX \subseteq DX$ Sub type of Probability measures (total measure 1)

BX type of measurable events - subsets of X we wish to measure

W type of weights : $[0, \infty]$

type judgment

$$\mu : DX, E : BX \vdash \mathbb{C}\mathrm{e}_\mu[E] : W$$

↳ measure μ assigns to E

Plan:

1) type-driven probability: discrete case (Mon + Tue) ✓
2) Borel sets & measurable spaces (Tue)
3) Quasi Borel spaces, simple type structure (Wed)
4) Dependent type structure & standard Borel spaces (Thu)
5) Integration & random variables (Fri)

Please ask questions!

Course web page

Scribble

Foundations for type-driven probabilistic modelling

Ohad Kammar
University of Edinburgh

Logic Summer School
Australian National University
4–16 December, 2023
Canberra, ACT, Australia

THE UNIVERSITY of EDINBURGH informatics lfcs Laboratory for Foundations of Computer Science BayesCentre

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

Recap

Languge of distribution & Probability

X type (= space) of values / outcomes

DX type of distributions / measures over X

$PX \subseteq DX$ Sub type of Probability measures (total measue 1)

BX type of measurable events - Subsets of X we wish to measure

W type of weights : $[0, \infty]$

type judgment

$$\mu : DX, E : BX \vdash \mathcal{C}e_{\mu}[E] : W$$

measure μ assigns to E

Axioms for measures/distributions

Recap

$$\mu : DX \vdash \underset{\mu}{Ce}[\emptyset] = 0 \quad : W$$

$$E, C : BX, \mu : DX \vdash$$

$$\underset{\mu}{Ce}[E] = \underset{\mu}{Ce}[E \cap C] + \underset{\mu}{Ce}[E \cap C^c] \quad : W$$

$$E_{\_} : (BX, \subseteq)^{\omega}, \mu : DX \vdash$$

$$\underset{\mu}{Ce}[\bigcup_n E_n] = \sup_n \underset{\mu}{Ce}[E_n] \quad : W$$

Kernels & their Kock integral

Recap

Kernel from Γ to X: $k : (DX)^{\Gamma}$ or $k : \Gamma \longrightarrow DX$

Dirac kernel: $\delta_{\_} : X \longrightarrow DX$

Kock integral: $\mu : D\Gamma,\ k : (DX)^{\Gamma} \vdash \oint \mu k \quad : DX$

or $\oint \mu(dx)\, k(x)$ (dx binding occurrence in $k(x)$)

Giry monads: $(D, \delta_{\_}, \oint)$ & $(P, \delta_{\_}, \oint)$.

Discrete model

Recap

type : Set $\qquad W := [0, \infty] \qquad \mathcal{B}X := \mathcal{P}X$

$$DX := \{ \mu : X \to W \mid \mathrm{supp}\,\mu \text{ countable} \}$$

$$\mathcal{P}X := \{ \mu \in DX \mid \underset{\mu}{\mathrm{Ce}}[X] = 1 \}$$

$$\underset{\mu}{\mathrm{Ce}}[E] := \sum_{x \in E} \mu x \qquad \delta_x := \lambda x'. \begin{cases} x = x' : & 0 \\ x \neq x' : & 1 \end{cases}$$

$$\oint \mu k := \lambda x. \sum_{r \in \Gamma} \mu r \cdot k(r; x)$$

Ex distributions

Recap

Counting measure (X ctbl): $\#_X := \lambda x. 1$

Dirac measure δ_x (prev slide)

Zero measure $\underline{0} := \lambda x. 0$

Plan:

1) type-driven probability: discrete case (Mon + Tue)

2) Borel sets & measurable spaces (Tue)

3) Quasi Borel spaces, simple type structure (Wed)

4) Dependent type structure & standard Borel spaces (Thu)

5) Integration & random variables (Fri)

Please ask questions!

Course web page

Scribble

Product measures

$$\mu : DX, \nu : DY \vdash \mu \otimes \nu := \oint \mu(dx) \oint \nu(dy)\, \delta_{\langle x,y \rangle} : D(X \times Y)$$

$$= \lambda(x,y).\ \mu x \cdot \nu y$$

($\otimes$ lifts along $P \hookrightarrow D$)

discrete model

Ex: $\#_{X \times Y} = \#_X \otimes \#_Y$

build measures compositionally

Indeed:

$$(\# \otimes \#)(x,y) = \# x \cdot \# y = 1 \cdot 1 = 1 = \#(x,y)$$

Notation: $\lambda : D(X \times Y),\ k : (DZ)^{X \times Y} \vdash \oint\!\!\oint \lambda(dx, dy)\, k(x, y)$

$:= \oint \lambda k$

Fubini-Tonelli Thm:

Integrate in any order:

$\mu : DX,\ \nu : DY,\ k : (DZ)^{X \times Y} \vdash$

$$\oint \mu(dx) \oint \nu(dy)\, k(x, y) = \oint\!\!\oint (\mu \otimes \nu)(dx, dy)$$

$$= \oint \nu(dy) \oint \mu(dx)\, k(x, y)$$

Pushing a measure forward

$\mu : D\Omega,\ \alpha : X^\Omega \vdash \mu_f := \oint \mu(d\omega)\,\delta_{\alpha\omega} : DX$

$$= \lambda x. \sum_{\substack{\omega \in \Omega \\ \alpha\omega = x}} \mu\,\omega$$

$\alpha : X^\Omega$: random element (w.r.t. μ)

$\mu_\alpha : DX$: the law of α

Ex: We can represent configurations of 2 dice using $\underline{6} \times \underline{6}$

Letting $(+) : \underline{6}^2 \longrightarrow \mathbb{N}^2 \xrightarrow{(+)} \mathbb{N}$

we have that the law of $(+)$:

$(\#_{\underline{6}} \otimes \#_{\underline{6}})_{(+)} : D\mathbb{N}$ is the number of rolls whose sum is given

build measures compositionally

Scaling a measure

$$(\cdot) : W \times DX \longrightarrow DX$$

$$a \cdot \mu := \lambda x.\ a \cdot \mu x$$

NB: $\mathrm{supp}(a \cdot \mu) = \begin{cases} a = 0 : & \phi \\ a \neq 0 : & \mathrm{supp}\,\mu \end{cases}$

✓ c+bl

$(\cdot) : W \times DX \to DX$ is an action of monoid $(W, (\cdot), 1)$ on DX:

$\mu : DX \vdash$

$$1 \cdot \mu = \mu$$

$a, b : W, \mu : DX \vdash$

$$a \cdot (b \cdot \mu) = (a \cdot b) \cdot \mu$$

Normalisation

unit type $\rightarrow$ $\mathbb{1} := \{\langle\rangle\}$ $\leftarrow$ unit value

$$\mu : DX,\quad Ce_{\mu}[X] \neq 0, \infty \vdash$$

$$\|\mu\| := \left(\frac{1}{Ce_{\mu}[X]}\right) \cdot \mu \quad : PX$$

Ex:

$$\emptyset \neq A \subseteq_{fin} X \quad : \mathcal{U}_{A \subseteq X} := \|(\#_A)_{A \subseteq X}\| \quad : PX$$

$$\mathbb{1} \xrightarrow{\#_A} DA \xrightarrow{(-)_{A \subseteq X}} DX \xrightarrow{\|-\|} PX$$

I.e. $\mathcal{U}_{A \subseteq X} := \lambda x. \begin{cases} x \in A : & \frac{1}{|A|} \\ x \notin A : & 0 \end{cases}$ so $\mathcal{U}_{\{x\} \subseteq X} = \delta_x$

Standard vocabulary

Joint distributions: $\mu : D(X_1 \times X_2)$

Marginal distribution:

$$X_1 \xleftarrow{\pi_1} X_1 \times X_2 \xrightarrow{\pi_2} X_2$$

law of projection

$$\mu_{\pi_i} : DX_i$$

marginalisation: $\mu_{\pi_1} = \iint \mu(dx, dy)\, \delta_x$

integrate out y

Exercise: $\mu : PX,\ \nu : DX \vdash (\mu \otimes \nu)_{\pi_2} = \nu$

independence

Pairing r.e.s:

$\alpha : X^{\Omega}, \beta : Y^{\Omega} \vdash$

$$\begin{tikzcd} & \Omega \arrow[dl, "\alpha"'] \arrow[d, dashed, "\langle \alpha, \beta \rangle"] \arrow[dr, "\beta"] & \\ X \arrow[r, phantom, "="] & X \times Y \arrow[l, "\pi_1"] \arrow[r, "\pi_2"'] \arrow[r, phantom, shift left=3, "="] & Y \end{tikzcd}$$

$$\langle \alpha, \beta \rangle := \lambda \omega . \langle \alpha \omega, \beta \omega \rangle : (X \times Y)^{\Omega}$$

$$\lambda : D\Omega, \alpha : X^{\Omega}, \beta : Y^{\Omega} \vdash \alpha \perp_{\lambda} \beta := \lambda_{\langle \alpha, \beta \rangle} = \lambda_{\alpha} \otimes \lambda_{\beta} : \mathrm{Prop}$$

α, β independent w.r.t. λ

$\underline{\text{Ex}}$ (Durrett) represent outcomes of 3 coin tosses:

$$C := \{T, H\} \qquad \Omega := C \times C \times C \qquad \lambda : U_C \otimes U_C \otimes U_C : P\Omega$$

$\pi_i : \Omega \longrightarrow C$ outcome of i^{th} toss

$$\mathrm{Same}_{ij} : \Omega \xrightarrow{\langle \pi_i, \pi_j \rangle} C \times C \xrightarrow{(\overset{?}{=})} \mathbb{B}$$

where: $(\overset{?}{=}) : C^2 \longrightarrow \mathbb{B} := \{\mathrm{True}, \mathrm{False}\}$

$$x \overset{?}{=} y := \begin{cases} x = y : & \mathrm{True} \\ x \neq y : & \mathrm{False} \end{cases}$$

Ex (Durrett) represent outcomes of 3 coin tosses:

$C := \{T, H\}$ $\Omega := C \times C \times C$ $\lambda : U_C \otimes U_C \otimes U_C : P\Omega$

$\pi_i : \Omega \longrightarrow C$ outcome of i^{th} toss

$\mathrm{Same}_{ij} : \Omega \xrightarrow{\langle \pi_i, \pi_j \rangle} C \times C \xrightarrow{(\overset{?}{=})} \mathbb{B}$

$$\lambda_{\mathrm{Same}_{12}}^{T} \overset{\text{marginalisation}}{=} (U_C \otimes U_C)^{T}_{(\overset{?}{=})} = \underbrace{\frac{1}{4}}_{U_C(H) \cdot U_C(H)} + \overset{U_C(T) \cdot U_C(T)}{\frac{1}{4}} = \frac{1}{2}$$

So $\lambda_{\mathrm{Same}_{12}}^{F} = \frac{1}{2}$ too

Ex (Durrett) represent outcomes of 3 coin tosses:

$C := \{T, H\}$ $\quad \Omega := C \times C \times C$ $\quad \lambda : U_C \otimes U_C \otimes U_C : P\Omega$

$\pi_i : \Omega \longrightarrow C$ outcome of i^{th} toss

$\mathrm{Same}_{ij} : \Omega \xrightarrow{\langle \pi_i, \pi_j \rangle} C \times C \xrightarrow{(\overset{?}{=})} \mathbb{B}$

$i \neq j$: $\lambda_{\mathrm{Same}_{ij}} = U_{\mathbb{B}}$

$\lambda_{\langle \mathrm{Same}_{12}, \mathrm{Same}_{23} \rangle} :$ $(T,T) \mapsto \underbrace{\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}}_{\lambda(H,H,H)} + \underbrace{\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}}_{\lambda(T,T,T)} = \frac{1}{4}$

$(T, F) \mapsto \underbrace{\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}}_{\lambda(H,H,T)} + \underbrace{\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}}_{\lambda(T,T,H)} = \frac{1}{4}$

<u>Ex</u> (Durrett) represent outcomes of 3 coin tosses:

$C := \{T, H\}$ $\quad \Omega := C \times C \times C$ $\quad \lambda : U_C \otimes U_C \otimes U_C : P\Omega$

$\pi_i : \Omega \longrightarrow C$ outcome of i^{th} toss

$\mathrm{Same}_{ij} : \Omega \xrightarrow{\langle \pi_i, \pi_j \rangle} C \times C \xrightarrow{(\overset{?}{=})} \mathbb{B}$

<u>$i \neq j$:</u> $\lambda_{\mathrm{Same}_{ij}} = U_{\mathbb{B}}$

$$\lambda_{\langle \mathrm{Same}_{12}, \mathrm{Same}_{23} \rangle} = U_{\mathbb{B} \times \mathbb{B}} = U_{\mathbb{B}} \otimes U_{\mathbb{B}} = \lambda_{\mathrm{Same}_{12}} \otimes \lambda_{\mathrm{Same}_{13}}$$

So $\mathrm{Same}_{12} \perp_{\lambda} \mathrm{Same}_{13}$

independence

$$\begin{tikzcd} & \Omega \arrow[dl, "\alpha"'] \arrow[d, dashed, "\langle\alpha,\beta\rangle"] \arrow[dr, "\beta"] & \\ X \arrow[r, phantom, "="] & X \times Y \arrow[l, "\pi_1"] \arrow[r, "\pi_2"'] & Y \end{tikzcd}$$

Pairing r.e.s:

$\alpha : X^\Omega, \beta : Y^\Omega \vdash$

$\langle\alpha,\beta\rangle := \lambda\omega. \langle \alpha\omega, \beta\omega\rangle : (X \times Y)^\Omega$

$\lambda : D\Omega, \alpha : X^\Omega, \beta : Y^\Omega \vdash \alpha \perp_\lambda \beta := \lambda_{\langle\alpha,\beta\rangle} = \lambda_\alpha \otimes \lambda_\beta : \mathrm{Prop}$

α, β independent w.r.t. λ

I-ary version:

$\lambda : D\Omega, \alpha_- : \prod_{i \in I} X_i^\Omega \vdash \perp_{\lambda}{}_{i \in I} \alpha_i :=$

$\forall J \subseteq_{fin} I.\ \lambda_{\langle\alpha_j\rangle_{j \in J}} = \bigotimes_{j \in J} \lambda_{\alpha_j} : \mathrm{Prop}$

α_- independent w.r.t. λ

Ex (Durrett) represent outcomes of 3 coin tosses:

$C := \{T, H\}$ $\quad \Omega := C \times C \times C$ $\quad \lambda : U_C \otimes U_C \otimes U_C : P\Omega$

$\pi_i : \Omega \longrightarrow C$ outcome of i^{th} toss

$\mathrm{Same}_{ij} : \Omega \xrightarrow{\langle \pi_i, \pi_j \rangle} C \times C \xrightarrow{(\overset{?}{=})} \mathbb{B}$

$i \neq j$: $\quad \lambda_{\mathrm{Same}_{ij}} = U_{\mathbb{B}}$

$\underset{\substack{* \; * \\ k}}{i \neq j}$: $\quad \mathrm{Same}_{ij} \perp_\lambda \mathrm{Same}_{jk}$ $\qquad \not\perp_\lambda \{\mathrm{Same}_{12}, \mathrm{Same}_{23}, \mathrm{Same}_{13}\}$

Intuition: $\mathrm{Same}_{13} = \mathrm{IFF}(\mathrm{Same}_{12}, \mathrm{Same}_{23})$

Calc:

$$\lambda_{\langle \mathrm{Same}_{12}, \mathrm{Same}_{23}, \mathrm{Same}_{13} \rangle}(T,T,T) = \underbrace{\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}}_{\lambda(H,H,H)} + \underbrace{\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}}_{\lambda(T,T,T)} = \frac{1}{4} \neq \frac{1}{2^3} = \lambda_{\mathrm{Same}_{12}} \otimes \lambda_{\mathrm{Same}_{23}} \otimes \lambda_{\mathrm{Same}_{13}}$$

Vocabulary

(Discrete) Measure space $(X, \mu : DX)$

measure preserving $f : (X, \mu) \to (Y, \nu)$

function $f : X \to Y$ s.t. $\mu_f = \nu$

$\mu : DX, f : X \to Y \vdash \mu$ invariant under $f :=$

$$f : (X, \mu) \to (X, \mu)$$

Ex:

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

$\mathrm{Swap} : (X \times Y, \mu \otimes \nu) \longrightarrow (Y \times X, \nu \otimes \mu)$ so

$\mu : DX \vdash \mu \otimes \mu$ invariant under swap

Weights as measures

NB: unit type $\to$ $\mathbb{1} := \{\langle\rangle\}$ $\leftarrow$ unit value

Observation:

$$\begin{tikzcd} D\mathbb{1} \arrow[rr, bend left, "\lambda\mu.\, ce_{\mu}[\mathbb{1}]"] & & W \arrow[ll, bend left, "(-\odot\delta_{\langle\rangle})"] \end{tikzcd}$$

mutually inverse

Proof:

$$\mu \longmapsto \mu(\langle\rangle) \qquad\qquad \lambda\langle\rangle.a \longleftarrow\!\!\mid a$$

$$\mu = \lambda\langle\rangle.\mu\langle\rangle \qquad \wedge \qquad a = (\lambda\langle\rangle.a)\langle\rangle$$

$\blacksquare$

Integration

$$\mu : DX,\ \varphi : W^X \vdash \int \mu\, \varphi \quad : W \qquad \text{(Lebesgue integral)}$$

$$:= \sum_{x \in X} \mu x \cdot \varphi x$$

Can derive it:

$$\begin{tikzcd}
DX \times W^X \arrow[rr, "DX \times (\cong \circ -)"] \arrow[d, "\int"'] \arrow[drr, phantom, ":="] & & DX \times (D\mathbb{1})^X \arrow[d, "\oint"] \\
W & & D\mathbb{1} \arrow[ll, "\cong"]
\end{tikzcd}$$

Additivity:

$$I \text{ ctbl}, \ \mu_{-} : (DX)^I \vdash \sum_{i \in I} \mu_i \ : DX$$

$$:= \lambda x. \sum_{i \in I} \mu_i x$$

NB:

$\operatorname{supp} \sum_i \mu_i \subseteq \bigcup_i \operatorname{supp} \mu_i$

✓ctbl

Ex: Bernoulli distribution

$$p : [0,1] \vdash B(p) := p \cdot \delta_{\mathrm{True}} + (1-p) \cdot \delta_{\mathrm{False}} \ : P\mathbb{B}$$

i.e. B_p : $\mathrm{True} \mapsto p$

$\mathrm{False} \mapsto 1-p$

Thm (affine-linearity):

$\oint$ is affine-linear in each argument:

$$\frac{I \text{ ctbl} \quad \mu_{-} : (D\Gamma)^{I},\ k : (DX)^{\Gamma}_{,} \quad a_{-} : \mathbb{W}^{I}}{\vdash \oint \left(\sum_{i \in I} a_i \cdot \mu_i\right) k = \sum_{i \in I} a_i \cdot \oint \mu_i \, k}$$

$$I \text{ ctbl},\ \mu : D\Gamma,\ a_{-} : \mathbb{W}^{I},\ k_{-} : DX^{I} \vdash$$

$$\int \mu(dx) \left(\sum_{i \in I} a_i \cdot k_i(x) \right) = \sum_{i \in I} a_i \cdot \oint \mu \, k_i$$

<u>Prop:</u> $\mathbb{W} \cong D\mathbb{1}$ is a σ-semi-ring isomorphism:

$$(\mathbb{W}, \Sigma, (\cdot), 1) \cong (D\mathbb{1}, \Sigma, (\cdot), \delta_{\langle\rangle})$$

and $(\cdot) : \mathbb{W} \times DX \to DX$ makes DX into a <u>module</u>:

$$\left(\sum_{i \in I} a_i\right) \cdot \mu = \sum_{i \in I} (a_i \cdot \mu) \qquad a \cdot \sum_{i \in I} \mu_i = \sum_{i \in I} a \cdot \mu_i$$

<u>Corollary:</u> $\int$ is affine-linear in each argument.

Random variable:

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

A random element $\alpha : \overline{\mathbb{R}}^{\Omega}$ (wrt some $\mu : D\Omega$)

Can add, multiply r.v.'s.

To integrate r.v.'s:

$$(-)^{+} : \overline{\mathbb{R}}^{\Omega} \longrightarrow W^{\Omega}$$

$$\alpha^{+} := \lambda w. \begin{cases} \alpha \cdot w \geq 0 : & \alpha w \\ 0.w : & 0 \end{cases} = [\alpha - \geq 0] \cdot |\alpha|$$

$$\alpha^{-} := \lambda w. \begin{cases} \alpha \cdot w \leq 0 : & |\alpha w| \\ 0.w. : & 0 \end{cases} = [\alpha - \leq 0] \cdot |\alpha|$$

So $\alpha = \alpha^{+} - \alpha^{-}$

$$\mu : D\Omega, \alpha : \overline{\mathbb{R}}^{\Omega}, \int \mu \alpha^{+} < \infty \text{ or } \int \mu \alpha^{-} < \infty \vdash$$

$$\int \mu\, \alpha := \int \mu \alpha^{+} - \int \mu \alpha^{-} \quad : \overline{\mathbb{R}}$$

<u>Ex.</u> The (discrete) Lebesgue p-space:

$$p : [1, \infty), \mu : P\Omega \vdash \mathcal{L}_p(\Omega, \mu) :=$$

$$\left\{ \alpha : \overline{\mathbb{R}}^{\Omega} \;\middle|\; \mathop{\mathbb{E}}_{\mu}[|\alpha|^p] < \infty \right\}$$

$\mathcal{L}_p(\Omega, \mu)$ has a norm $\|\alpha\| := \sqrt[p]{\mathop{\mathbb{E}}_{\mu}[|\alpha|^p]}$ almost Banach

$\mathcal{L}_2(\Omega, \mu)$ has an inner product $\langle \alpha, \beta \rangle := \mathop{\mathbb{E}}_{\mu}[\alpha \cdot \beta]$ almost Hilbert

Density

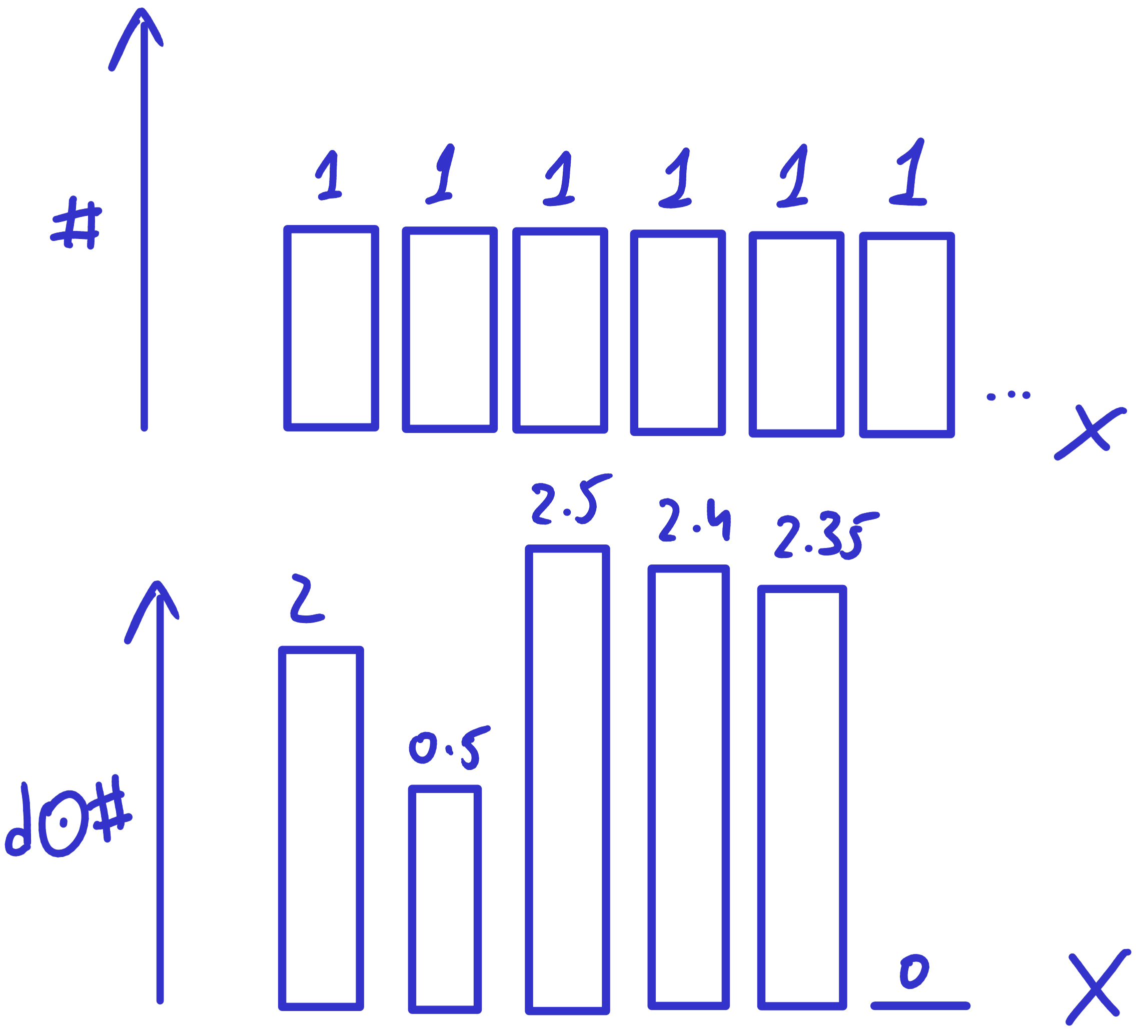
a density over X : $d : X \to W$

$$d : W^X,\ \mu : DX \vdash d \odot \mu \quad : DX$$

$$:= \oint \mu(dx)(d\,x \cdot \delta_x)$$

Warning The types of measures & densities in the discrete model are close, but still different. They coincide on countable sets, so people often confuse them. Types help us keep them separate.

Intuition:



Almost certain properties

$$E : \mathcal{B}X, \mu : DX \vdash \mu(dx)\text{-almost certainly } x \in E \quad : \mathrm{Prop}$$

$$:= [- \in E] \odot \mu = \mu$$

$$\text{NB: } [x \in E] = \begin{cases} x \in E : & 1 \\ x \notin E : & 0 \end{cases} \quad : W$$

When $\mu : PX$ we say instead

$$\mu(dx)\text{-almost surely } x \in E$$

Exercise Look up the def. of a normed space and modify the definition so that $\mathcal{L}_p(\Omega, \mu)$ is a normed space up-to almost sure equality.

Absolute continuity

d is a density of μ w.r.t. ν or

d is a Radon-Nikodym derivative w.r.t. ν

$$\mu, \nu : DX,\ d : W^X \vdash d = \frac{d\mu}{d\nu} \quad : \mathrm{Prop}$$

$$:= \mu = d \odot \nu$$

$$\mu, \nu : DX \vdash \mu \ll \nu := \mu \text{ is absolutely continuous w.r.t. } \nu : \mathrm{Prop}$$

$$:= \exists d : W^X.\ d = \frac{d\mu}{d\nu}.$$

$$=: \mu \text{ has a density w.r.t. } \nu$$

Lemma: $\mu, \nu : DX$, $\mu \ll \nu$, $k : (DY)^X$

$$\oint \nu(dx)\, \frac{d\mu}{d\nu}(x) \cdot k\,x = \oint \mu(dx)\, k\,x$$

Ex: $U_{A \subseteq X} \ll (\#_A)_{\mathrm{Cnst} : A \subseteq X}$

$$\frac{dU_{A \subseteq X}}{d(\#_A)_{\mathrm{Cnst}}} = \lambda x. \begin{cases} x \in A : & \frac{1}{|A|} \\ \text{o.w.} : & 0 \end{cases}$$

but also: $$\frac{dU_{A \subseteq X}}{d(\#_A)_{\mathrm{Cnst}}} = \lambda x. \frac{1}{|A|}$$

Radon-Nikodym Thm: (discrete version)

$\mu, \nu : PX \vdash \mu \ll \nu$ iff $\forall x.\ \nu x = 0 \Rightarrow \mu x = 0$

i.e. $\mathrm{Supp}\,\mu \subseteq \mathrm{Supp}\,\nu$

In that case, if $d_1, d_2 = \frac{d\mu}{d\nu}$ then

$\nu(dx)$-a.s. $d_1 x = d_2 x$

<u>Ex</u>: for ctbl X, $\forall \mu : DX.\ \mu \ll \#_X$. Proof: vacuously, as $\#_X x \neq 0$.

Then $\lambda x.\,\mu x = \frac{d\mu}{d\#}$.

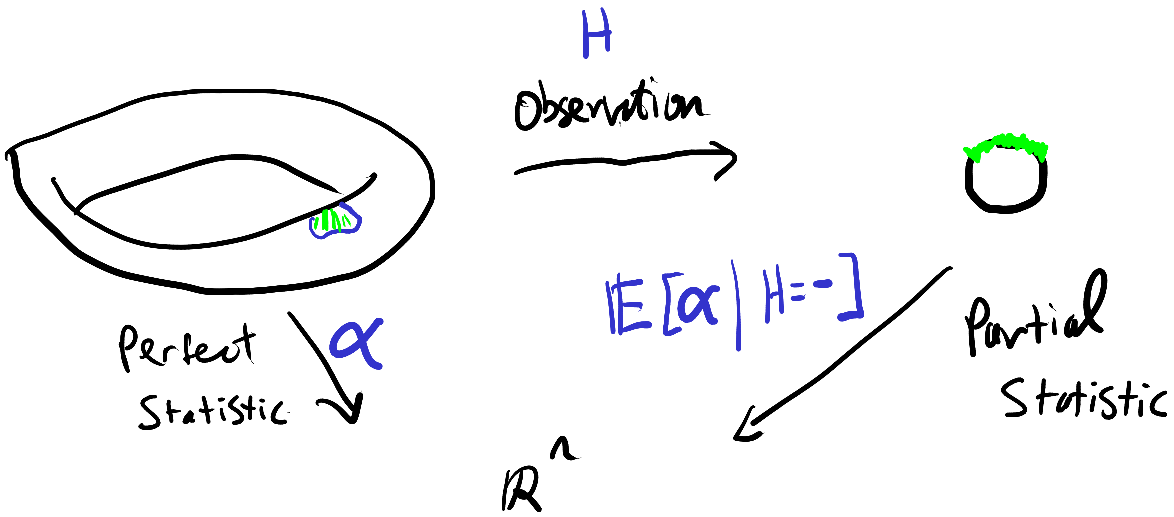
Conditional expectation

β is a conditional expectation of α w.r.t. μ along H

$$\mu : D\Omega,\ H : X^{\Omega},\ \alpha : \mathcal{L}_1(\Omega, \mu),\ \beta : \mathcal{L}_1(X, \mu_H)$$

$$\vdash \beta = \mathbb{E}[\alpha \mid H = -] \qquad : \mathrm{Prop}$$

$$:= \forall \varphi : \mathcal{L}_1(Y, \mu_H^M).\ \int \mu_H(dx)\, \beta(x) \cdot \varphi(x) = \int \mu(d\omega)\, \alpha(\omega) \cdot \varphi(H\omega)$$



Thm (Kolmogorov): (discrete version)

There is a function

$$\mathbb{E}[-|-=-] \in \prod_{\mu : P\Omega} \prod_{H : X^{\Omega}} \mathcal{L}_1(\Omega, \mu) \to \mathcal{L}_1(X, \mu_H)$$

s.t. $\mathbb{E}_{\mu}[\alpha | H = -]$ is a conditional expectation of α w.r.t. μ along H.

Conditional Probability (discrete version):

$$H : X^{\Omega},\ \mu : PX \vdash \mathbb{P}_{\mu}[- \mid H = -] : (P\Omega)^{X}$$

$$:= \lambda x_0 : X.\ \lambda \omega_0 : \Omega.\ \mathop{\mathbb{E}}_{\omega \sim \mu}\big[[\omega_0 = \omega] \,\big|\, H\omega = x_0\big]$$

Bayes's Thm (discrete version, adapted from Williams):

Let $\lambda : P(X \times \Theta)$ joint probability distribution.

Assume $\mu : DX$, $\nu : D\Theta$ s.t. $\lambda \ll \mu \otimes \nu$.

with $d_{X,H} = \dfrac{d\lambda}{d(\mu \otimes \nu)}$.

$$\begin{array}{ccc} & X \times \Theta & \\ {}^{\alpha := \pi_1} \swarrow & & \searrow {}^{H := \pi_2} \\ X & & \Theta \end{array}$$

observation space.

Obs 1: $d_X : W^X$

$d_X := \lambda x. \int \nu(d\theta)\, d_{X,H}(x, \theta)$

then $d_X = \dfrac{d\lambda_\alpha}{d\mu}$

& similarly $(d_H : W^\Theta) := \lambda\theta. \int \mu(dx)\, d_{X,H}(x, \theta) = \dfrac{d\lambda_H}{d\nu}$

Bayes's Thm (discrete version, adapted from Williams):

$$\begin{tikzcd} & X \times \Theta \arrow[dl, "\alpha := \pi_1"'] \arrow[dr, "H := \pi_2"] & \\ X & & \Theta \end{tikzcd}$$

observation space.

Let $\lambda : P(X \times \Theta)$ joint probability distribution.

Assume $\mu : DX$, $\nu : D\Theta$ s.t. $\lambda \ll \mu \otimes \nu$.

with $d_{X,H} = \dfrac{d\lambda}{d(\mu \otimes \nu)}$. $d_X = \dfrac{d\lambda_\alpha}{d\mu}$ $\quad d_\Theta = \dfrac{d\lambda_H}{d\nu}$

Let $d_{X|H}(-|-) : X \times \Theta \to W$

$$d_{X|H}(x|\theta) := \begin{cases} d_H\theta \neq 0: & \dfrac{d_{X,H}(x,\theta)}{d_H\theta} \\ \text{o.w.}: & 0 \end{cases}$$

$\lambda_{X|H=-} : \Theta \to PX$

$\lambda_{X|H=\theta} := d_{X|H}(-|\theta) \odot \mu$

Bayes's formula:

$$\mathbb{P}\mathrm{r}_\lambda[-|H=-] = \lambda_{X|H=-}$$

Summary

$\mu \otimes \nu$ Product measures & Fubini-Tonelli

μ_H push-forward / law

$(Dx, \Sigma, (\cdot))$ module structure and affine linearity of $\oint$

$\int$ Lebesgue integration

Standard vocabulary: joint dist., marginalisation. independence, invariance

density & Radon-Nikodym derivatives (heed the warning)

almost certain properties

Conditional expectation & Probability

with Bayes's Thm.

Plan:

1) type-driven probability: discrete case (Mon + Tue ✓)

2) Borel sets & measurable spaces (~~Tue~~ Wed)

3) Quasi Borel spaces, simple type structure (Wed)

4) Dependent type structure & standard Borel spaces (Thu)

5) Integration & random variables (Fri)

Please ask questions!

Course web page

scribble

Foundations for type-driven probabilistic modelling

Ohad Kammar
University of Edinburgh

Logic Summer School
Australian National University
4–16 December, 2023
Canberra, ACT, Australia

THE UNIVERSITY of EDINBURGH informatics lfcs Laboratory for Foundations of Computer Science BayesCentre

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

Plan:

1) type-driven probability: discrete case (Mon + Tue)

2) Borel sets & measurable spaces (Wed)

3) Quasi Borel spaces (Wed) Simple type structure (Thu)

4) Dependent type structure & standard Borel spaces (Thu)

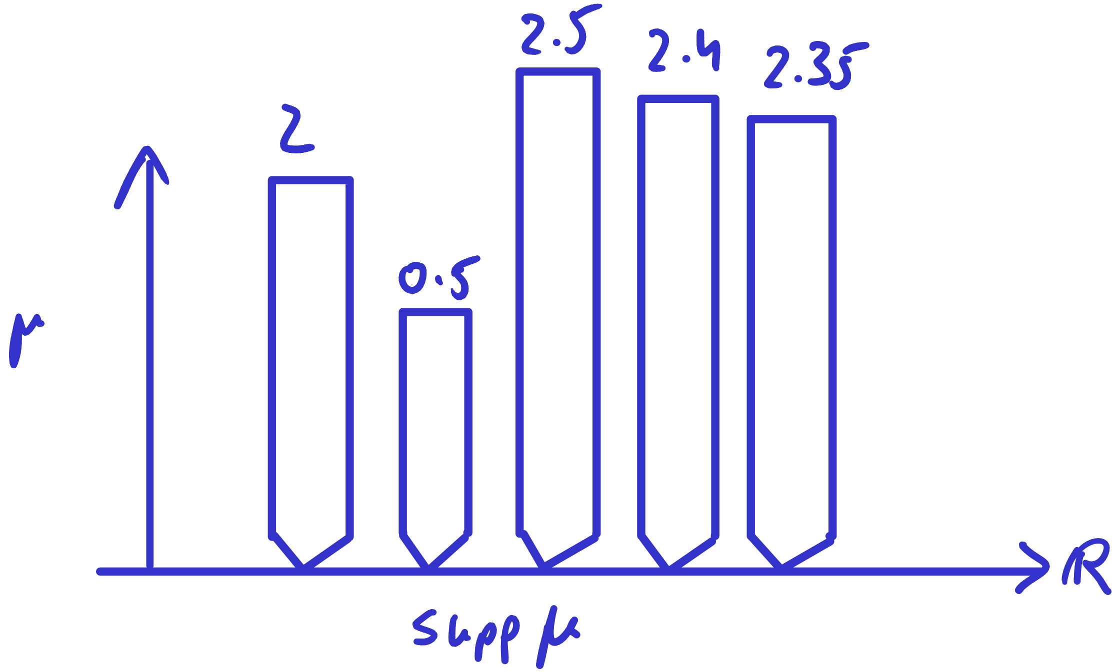
5) Integration & random variables (Fri)

Please ask questions!

Course web page

scribble

discrete model measure only histograms:



Want:

- lengths
- areas
- volumes.

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

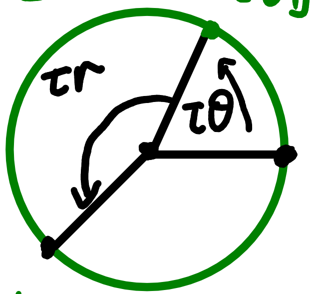
Thm: no $\lambda : \mathcal{P}\mathbb{R} \to [0,\infty]$:

$$\lambda(a,b) = b - a \qquad \lambda(r + A) = \lambda A \qquad \lambda\left(\biguplus_{n=0}^{\infty} A_n\right) = \sum_{n=0}^{\infty} \lambda A_n$$

(generalises length) (translation invariant) σ-additive

Direct proof in standard analysis courses. Idea behind typical proof is:

S^1 $[\![\theta]\!] := (\cos\theta, \sin\theta)$



$[\![0]\!] := (1,0)$

Thm: no $\lambda : \mathcal{P}S^1 \to [0,\infty]$
s.t.

$r : \mathbb{R} \vdash \mathrm{rotate}_r[\![\theta]\!] := [\![\theta + \tau r]\!]$

a) Satisfy measure axioms for $\mathcal{B}S^1 := \mathcal{P}S^1$

b) invariant under rotations: $E : \mathcal{B}S^1 \vdash$

$$\lambda\, \mathrm{rotate}[E] = \lambda E$$

c) $\lambda S^1 = \tau \ (:= 2\pi)$

reduce (S^1, λ^{S^1}) to $(\mathbb{R}, \lambda^{\mathbb{R}})$ via restriction & push forward

$$\lambda^{\mathbb{R}}\Big|_{\mathcal{P}[0,1)} := \lambda E \subseteq [0,1] .\ \lambda E : \mathcal{P}[0,1) \to \mathbb{W}$$

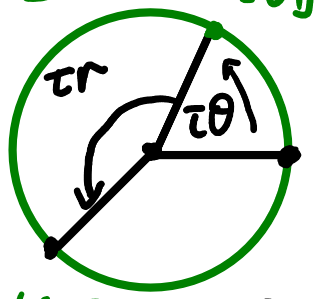
$$\lambda^{S^1} := \lambda E \subseteq S^1 .\ \lambda^{\mathbb{R}}_{\mathcal{P}[0,1)}\left(\llbracket - \rrbracket^{-1}[E]\right) : \mathcal{P}S^1 \xrightarrow{\llbracket - \rrbracket^{-1}} \mathcal{P}[0,1) \xrightarrow{\lambda^{\mathbb{R}}_{[0,1)}} \mathbb{W}$$

noting

rotations in $S^1 \longleftrightarrow$ translations in $\mathbb{R}$

Since $\nexists\, \lambda^{S^1}$, we have $\nexists\, \lambda^{\mathbb{R}}$ either.

S^1 $\quad [\![\theta]\!] := (\cos\theta, \sin\theta)$



$[\![0]\!] := (1,0)$

$r : \mathbb{R} \vdash \mathrm{rotate}_r [\![\theta]\!] := [\![\theta + \tau r]\!]$

Thm: no $\lambda : \mathcal{P}S^1 \longrightarrow [0,\infty]$
s.t.

a) satisfy measure axioms for $\mathcal{B}S^1 := \mathcal{P}S^1$

b) invariant under rotations: $E : \mathcal{B}S^1 \vdash \lambda\, \mathrm{rotate}[E] = \lambda E$

c) $\lambda S^1 = \tau \; (:= 2\pi)$

Proof: $a + b \Rightarrow \neg c$:

1) Using axiom of choice (AoC):

$$S^1 = \biguplus_{i=0}^{\infty} E_i \qquad E_i = \mathrm{rotate}_{r_i}[E_0]$$

2) $$\lambda S^1 = \sum_{i=0}^{\infty} \lambda E_i = \sum_i \lambda\, \mathrm{rotate}_{r_i} E_0 = \sum_{i=0}^{\infty} \lambda E_0 = \begin{cases} \lambda E_0 = 0 : 0 \\ \lambda E_0 > 0 : \infty \end{cases} \neq \tau$$

Constructing E_i:

$$x, y : S' \vdash x \sim y := \exists q \in \mathbb{Q}.\ \mathrm{rotate}_q\, x = y \quad : \mathrm{Prop}$$

$$\equiv \exists q \in [0,1) \cap \mathbb{Q}.\ \mathrm{rotate}\, x = y$$

$\sim$-equivalence classes:

$$x : S' \vdash [x]_\sim := \{ y \in S' \mid x \sim y \} \quad : \mathcal{P} S'$$

$$C := \{ [x]_\sim \in \mathcal{P} S' \mid x \in S' \}$$

$\forall e \in C,\ e \neq \emptyset$, so by AoC: $\exists \xi_- : C \longrightarrow S'.\ \xi_e \in e$.

NB: ξ injective

Take $C_0 := \{\xi_e \in S' \mid e \in C\} \in \mathcal{P}S'$

Note: $x \sim y,\ x, y \in C_0 \vdash x = y$.

$q : \mathbb{Q} \vdash C_q := \mathrm{rotate}_q[C_0] \in \mathcal{P}S'$

NB: $x, y : S' \vdash$
$x \sim y : \mathrm{Prop}$

$C = \sim$-equiv.

$\xi_{-} : C \rightarrowtail S'$

$e : C \vdash \xi_e \in e$

Let $(r_i)_{i=0}^{\infty}$ enumerate $\mathbb{Q} \cap [0,1)$ s.t. $r_0 = 0$

Take $E_i := C_{r_i}$

By fiat: $E_i = C_{r_i} = \mathrm{rotate}_{r_i}[C_0] = \mathrm{rotate}_{r_i}[E_0]$

RTP: $S' = \bigoplus_{i=0}^{\infty} E_i$

$E_i \cap E_j = \emptyset, \quad i \neq j:$

$x \in E_1 \cap E_2 \Rightarrow \exists\, y_i \in C_0 .\ x = \mathrm{rotate}_{r_i}\, y_i$

$\Rightarrow y_1 \sim x \sim y_2 \Rightarrow y_1 = y_2 =: y$

$\Rightarrow \mathrm{rotate}_{r_2 - r_1}\, y = y,\ |r_2 - r_1| < 1$

$\Rightarrow r_1 = r_2$

$\underline{S' = \bigcup_{i=0}^{\infty} E_i}$: $x \in S'$. Letting $e := \xi_{[x]_\sim} : \rho S'$

$\xi_e, x \in e \Rightarrow \xi_e \sim x$

$\Rightarrow \exists q \in \mathbb{Q} \cap [0,1).\ \mathrm{rotate}_q \xi_e = x.$

As $\xi_e \in C_0$: $x \in C_q$. Find i s.t. $r_i = q$

and $x \in C_{r_i} = E_i$. $\blacksquare$

Takeaway: taking $\mathcal{B}\mathbb{R} := \mathcal{P}\mathbb{R}$

excludes measures such as:

length, area, volume

Workaround: only measure well-behaved subsets

Def: The Borel subsets $B_{\mathbb{R}} \subseteq P\mathbb{R}$:

- open intervals $(a,b) \in B_{\mathbb{R}}$

Closure under σ-algebra operations:

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

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

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

Non-examples?

More complicated: analytic, Lebesgue

Def: Measurable space $V = (\lfloor V \rfloor, B_V)$

Set (Carrier)

family of subsets $B_V \subseteq P\lfloor V \rfloor$

closed under σ-algebra operations:

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

empty set

complements

countable unions

Idea: Structure all spaces after the worst-case scenario

Examples

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

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

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

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

- Sub spaces: $A \in P \llcorner V \lrcorner$ $\quad A := (A, [B_V] \cap A)$

- Products: $A \times B := (\llcorner A \lrcorner \times \llcorner B \lrcorner, \sigma([B_A] \times [B_B]))$

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:

$id : V \longrightarrow V$

Composition:

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

Meas Category

Products, coproducts/disjoint union, subspaces

Categorical limits, colimits, but:

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

membership predicate

$$(\ni) : (B_{\mathbb{R}}, B_{B_{\mathbb{R}}}) \times \mathbb{R} \longrightarrow \mathrm{Bool}$$

$$(U, r) \longmapsto [r \in U]$$

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

$$(f, r) \mapsto f(r)$$

Questions? skip proof?

Proof (sketch):

Borel hierarchy:

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

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

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

even for $B_{B_{\mathbb{R}}} = P(B_{\mathbb{R}})$

$$(\ni) : (B_{\mathbb{R}}, B_{B_{\mathbb{R}}}) \times \mathbb{R} \to \mathbb{R}$$

$$(U, r) \mapsto [r \in U]$$

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

If measurable:

$$\alpha := \operatorname{rank}\left((\ni)^{-1}[\mathrm{true}]\right) < \omega_1$$

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

But:

$$\alpha < \operatorname{rank} A = \operatorname{rank}\left(\langle A, -\rangle^{-1}\left[(\ni)^{-1}[\mathrm{true}]\right]\right) \leq \operatorname{rank}\left((\ni)^{-1}[\mathrm{true}]\right) \leq \alpha$$

$\#$

More details in Ex. B

Sequential Higher-order structure:

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

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

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

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

$\lim\sup : [-\infty,\infty]^{\mathbb{N}} \to [-\infty,\infty] \qquad \lim : \mathrm{Cauchy} \to \mathbb{R}$

Compose higher-order building blocks:

lim is measurable!

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

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

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

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: measurability by type!

But

For higher-order building blocks

defer measurability proofs until

we resume 1st order fragment ⇒ non compositional

Plan:

1) type-driven probability: discrete case (Mon + Tue)

2) Borel sets & measurable spaces (Wed)

3) Quasi Borel spaces (Wed) Simple type structure (Thu)

4) Dependent type structure & standard Borel spaces (Thu)

5) Integration & random variables (Fri)

Please ask questions!

Course web page

scribble

Plan

Def: $V \in$ 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 include

- 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:

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

- Countable coproducts of Sbs:

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

Conservative extensions:

concrete spaces
we "observe"

Standard Borel spaces

↙ ↘

Meas Qbs

abstract (auxiliary?)

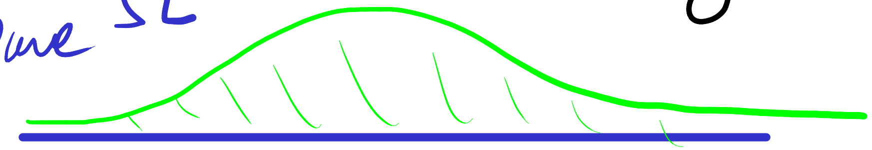
spaces

Core idea

Measure Theory

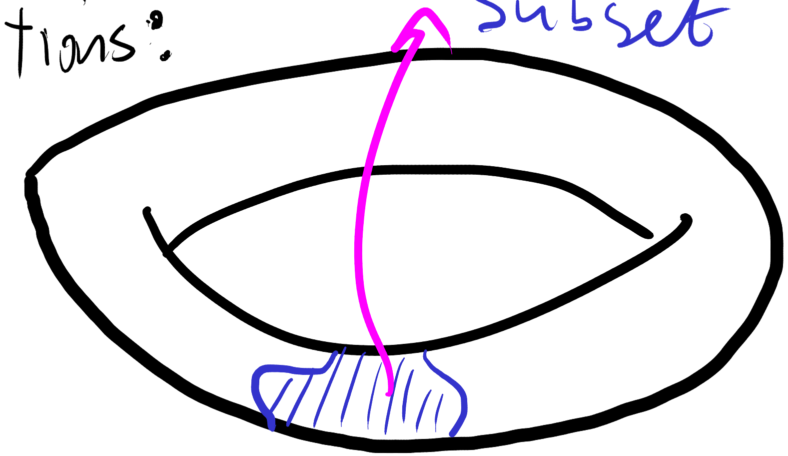
Qbs Theory

sample space Ω



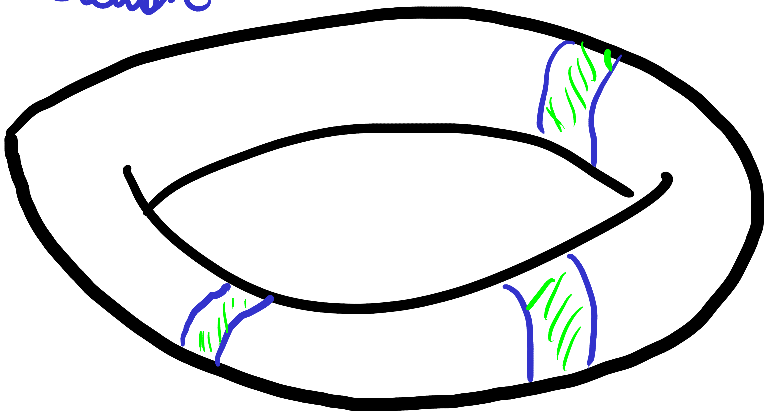
Primitive notions:

measurable subset



random element

$\downarrow \alpha$



Derived notions:

measure

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

measurable subsets

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

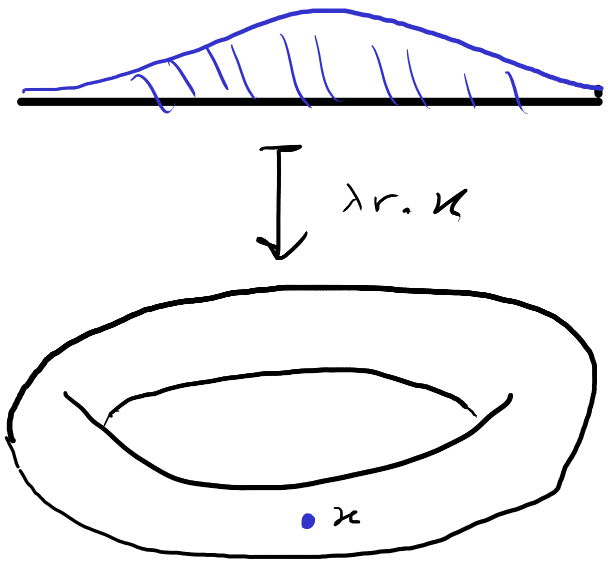
set "carrier"

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

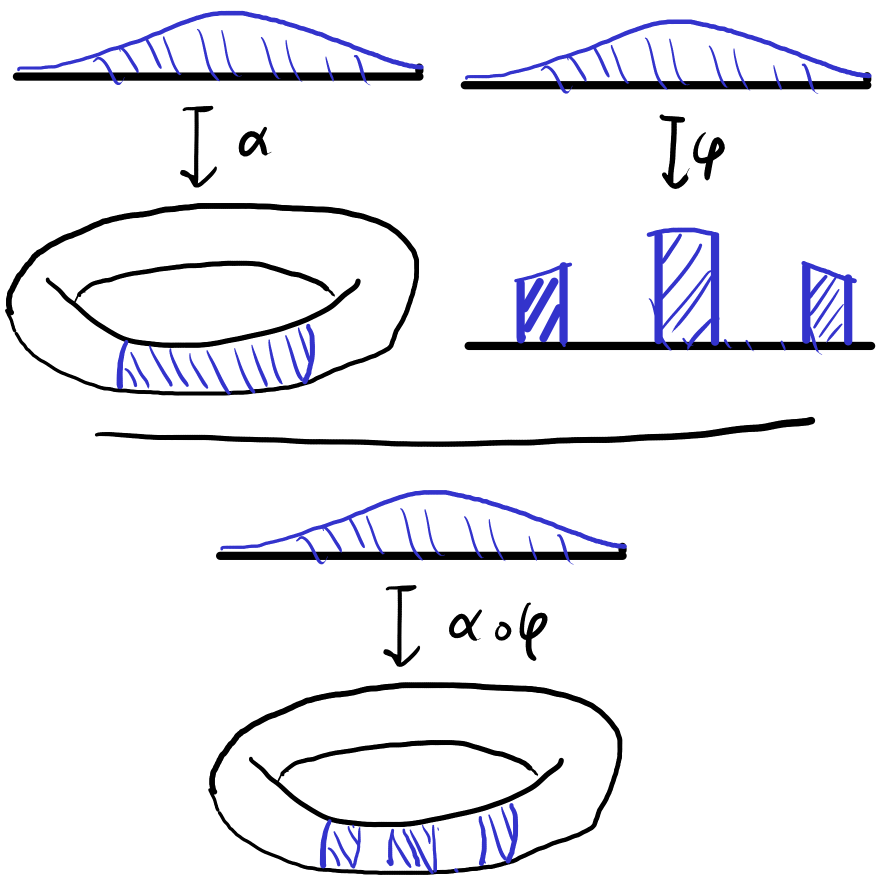
set "carrier"

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

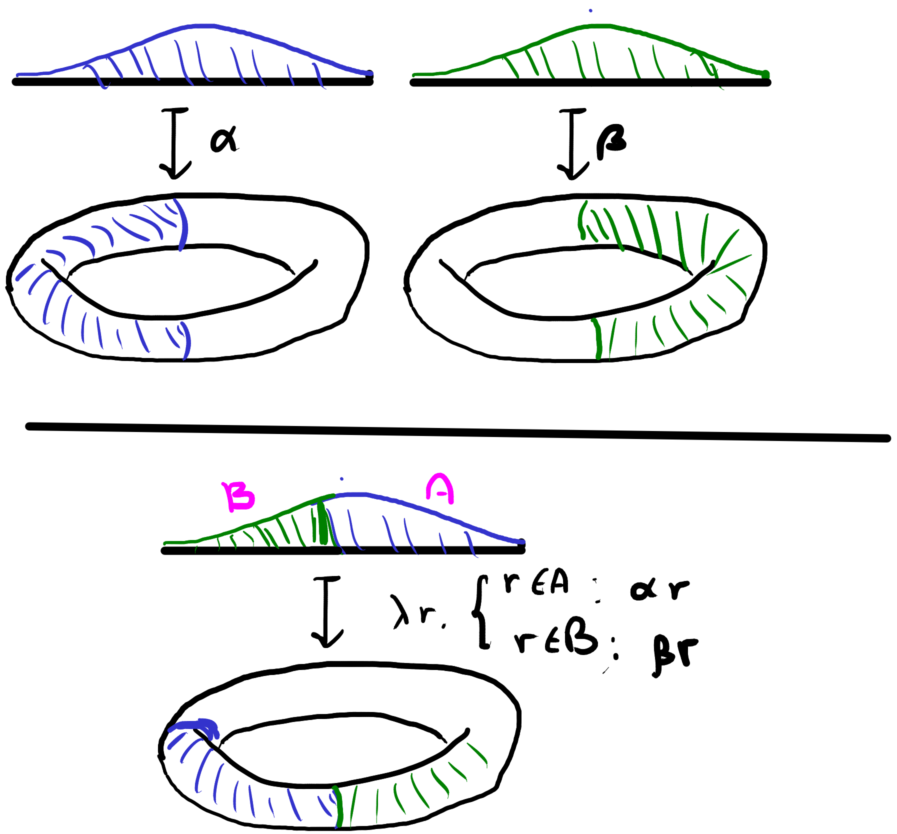
set "carrier"

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

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



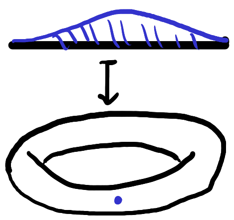
Def: Quasi-Borel space $X = (\lfloor X \rfloor, R_X)$

set "carrier"

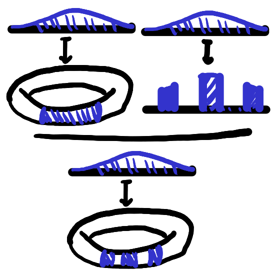
set of functions $\alpha : \mathbb{R} \to \lfloor X \rfloor$ "random elements"

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

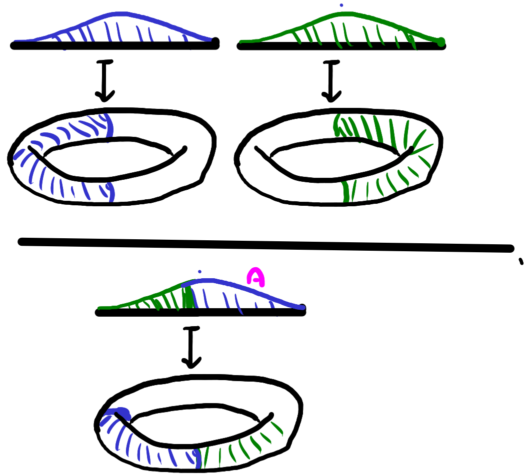
- constants:



- precomposition:



- recombination



Examples

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

qbs underlying $\mathbb{R}$

recombination of constants

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

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

discrete qbs on X

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

↳ all funtions

Indiscrete qbs on X

Qbs morphism $f : X \to Y$

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

Example

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

Category Qbs $\Leftarrow$ - identity, composition

Full model

type : Qbs $\quad$ $\mathbb{W} := [0,\infty]$ $\quad$ $\mathcal{B}X :=$ (Thur)

$DX :=$ (Fri)

$PX := \{\mu \in DX \mid \mathcal{C}e_{\mu}[X] = 1\}$ (Thu)

$\mathcal{C}e_{\mu}[E] :=$ (Fri) $\quad$ $\delta_x :=$ (Fri)

$\oint \mu k :=$ (Fri)

Plan:

1) type-driven probability: discrete case (Mon + Tue)

2) Borel sets & measurable spaces (Wed)

3) Quasi Borel spaces (Wed) Simple type structure (Thu)

4) Dependent type structure & standard Borel spaces (Thu)

5) Integration & random variables (Fri)

Please ask questions!

Course web page

Scribble

Foundations for type-driven probabilistic modelling

Ohad Kammar
University of Edinburgh

Logic Summer School
Australian National University
4–16 December, 2023
Canberra, ACT, Australia

THE UNIVERSITY of EDINBURGH informatics lfcs Laboratory for Foundations of Computer Science BayesCentre

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

Plan:

1) type-driven probability: discrete case (Mon + Tue)

2) Borel sets & measurable spaces (Wed)

3) Quasi Borel spaces (Wed) Simple type structure (Thu)

4) Dependent type structure & standard Borel spaces (Thu)

5) Integration & random variables (Fri)

Please ask questions!

Course web page

Scribble

Full model

type : Qbs $\quad W := [0,\infty] \quad BX := $ (Thur)

$DX :=$ (Fri)

$PX := \{\mu \in DX \mid \mathbb{E}_{\mu}[X] = 1\}$ (Thu)

$\mathbb{E}_{\mu}[E] :=$ (Fri) $\qquad \delta_x :=$ (Fri)

$\oint \mu k :=$ (Fri)

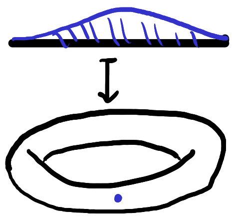
Def: Quasi-Borel space $X = (\lfloor X \rfloor, R_X)$

set "carrier"

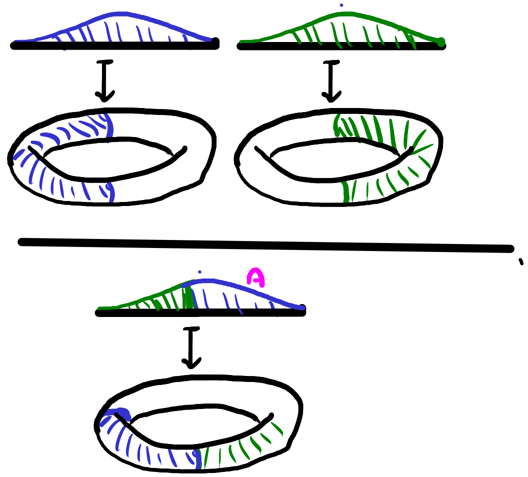
set of functions $\alpha : \mathbb{R} \to \lfloor X \rfloor$ "random elements"

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

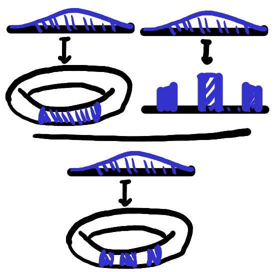
- constants:



- 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 funtions

  Indiscrete qbs on X

Validate qbs axioms for: $\mathbb{W} := ([0,\infty], \mathrm{Meas}(\mathbb{R}, \mathbb{W}))$

- Constants:

$$E : B_{\mathbb{W}},\ x : \mathbb{W} \vdash$$

$$(\lambda r:\mathbb{R}.\, x)^{-1}[E] = \begin{cases} x \in E: & \mathbb{R} \\ x \notin E: & \emptyset \end{cases} \in B_{\mathbb{R}}$$

✓

Validate qbs axioms for: $\mathbb{W} := ([0,\infty], \mathrm{Meas}(\mathbb{R}, \mathbb{W}))$

- Precomposition:

$$\alpha : \mathrm{Meas}(\mathbb{R}, \mathbb{W}),\quad \varphi : \mathrm{Meas}(\mathbb{R}, \mathbb{R}) \vdash$$

$$\mathbb{R} \xrightarrow{\varphi} \mathbb{R} \xrightarrow{\alpha} \mathbb{W} \quad \in \mathrm{Meas}(\mathbb{R}, \mathbb{W})$$

Meas is a cat.

Explicitly:

$$(\alpha \circ \varphi)^{-1}[E] \in \mathcal{B}\mathbb{R} \xleftarrow{\varphi^{-1}} \alpha^{-1}[E] \in \mathcal{B}\mathbb{R} \xleftarrow{\alpha^{-1}} E \in \mathcal{B}\mathbb{W} \quad \checkmark$$

Validate qbs axioms for: $\mathbb{W} := ([0,\infty], \mathrm{Meas}(\mathbb{R}, \mathbb{W}))$

- recombination

I ctbl, $\alpha_- : \mathrm{Meas}(\mathbb{R}, \mathbb{W})^I$, $E_- : B_{\mathbb{R}}^I$, $\mathbb{R} = \biguplus_{i \in I} E_i$, $F : B_{\mathbb{W}}$,

$$\beta := \left(\lambda r. \begin{cases} \vdots \\ r \in E_i : & \alpha_i r \\ \vdots \end{cases}\right)^{-1}[F]$$

$$= \bigcup_{i \in I} \alpha_i^{-1}[F] \cap E_i \in B_{\mathbb{R}}$$

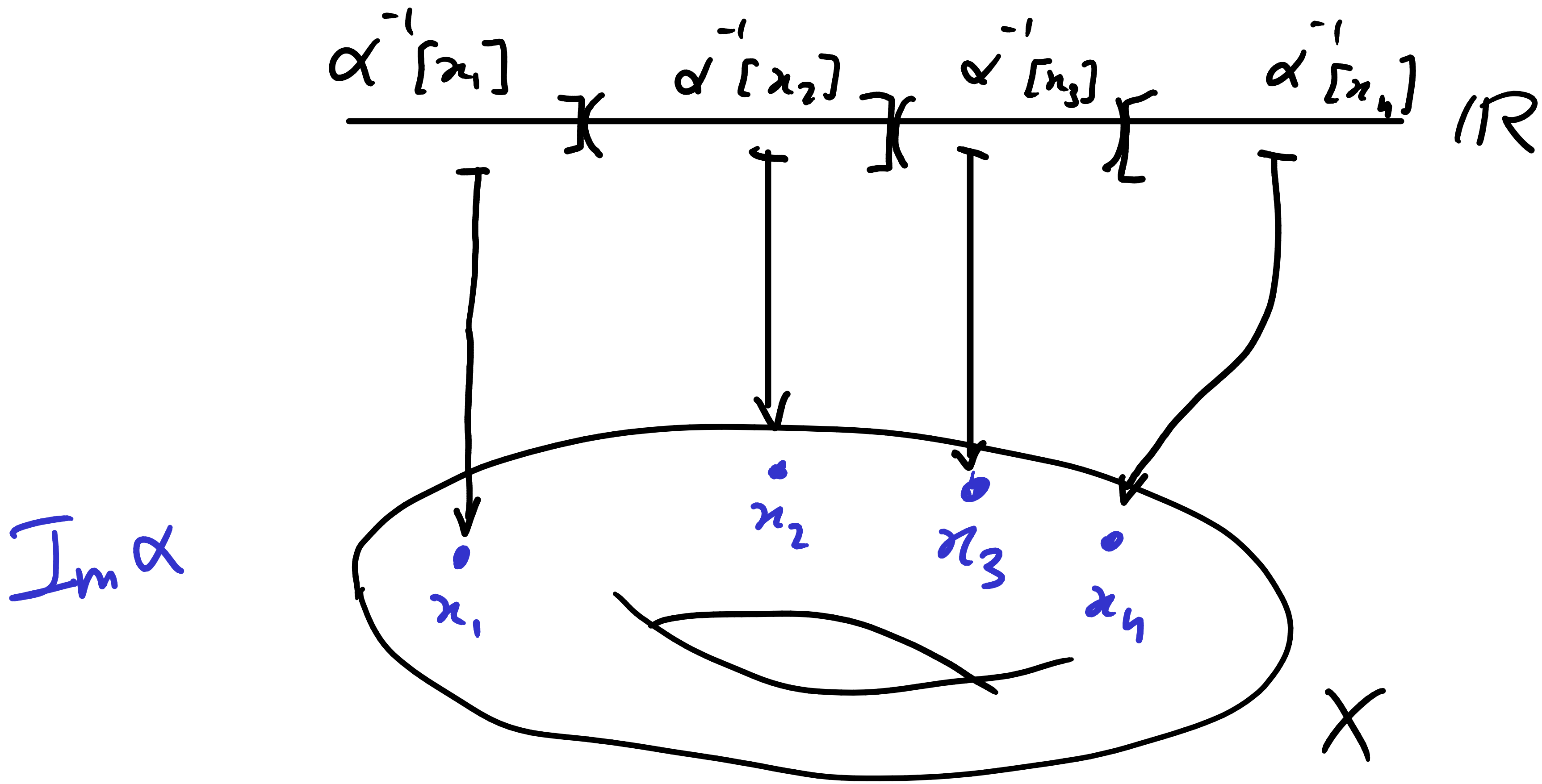
Indeed:

$r \in LHS \iff \beta r \in F \iff \exists i \in I. r \in E_i \wedge \alpha_i r \in F \iff r \in RHS$ ✓

σ-simple function

$\alpha : \mathbb{R} \longrightarrow X$ s.t. $\mathrm{Im}\,\alpha := \alpha[\mathbb{R}]$ is ctbl $\wedge$

$\forall x \in \mathrm{Im}\,\alpha.\ \alpha^{-1}[x] \in \mathcal{B}_{\mathbb{R}}$



Validate qbs axioms for: $\ulcorner X^{\mathrm{Qbs}} \urcorner := (X, \sigma\text{-Simple}(X))$

NB: f σ-Simple: $\mathrm{Im}\, f$ ctbl $\wedge$ $f^{-1}[x] \in B_{\mathbb{R}}$

- Constants

$$\mathrm{Im}(\lambda r.x) = \{x\} \text{ ctbl } \checkmark$$

$$y : X \vdash (\lambda r.x)^{-1}[y] = \begin{cases} x = y : & \mathbb{R} \\ x \neq y : & \emptyset \end{cases} \in B_{\mathbb{R}} \checkmark$$

Validate qbs axioms for: $\ulcorner X^{\mathrm{Qbs}} \urcorner := (X, \sigma\text{-Simple}(X))$

NB: f σ-Simple: $\mathrm{Im} f$ ctbl $\wedge$ $f^{-1}[x] \in B_{\mathbb{R}}$

- Precomposition:

$\alpha : \sigma\text{-simple}(X),\ \varphi : \mathrm{Meas}(\mathbb{R}, \mathbb{R}) \vdash$

$\mathrm{Im}(\alpha \circ \varphi) \subseteq \mathrm{Im}\,\alpha$ ctbl ✓

$x : X \vdash$

$(\alpha \circ \varphi)^{-1}[x] = \varphi^{-1}[\alpha^{-1}(x)] \in B_{\mathbb{R}}$ ✓

$\alpha^{-1}(x) \in B_{\mathbb{R}}$

$\varphi : \mathbb{R} \to \mathbb{R}$ measurable

Validate qbs axioms for: $\ulcorner X^{\mathrm{Qbs}} \urcorner := (X, \sigma\text{-Simple}(X))$

NB: f σ-Simple: $\operatorname{Im} f$ ctbl $\wedge$ $f^{-1}[x] \in B_{\mathbb{R}}$

- recombination:

$$\alpha_- : (\sigma\text{-Simple}(X))^I, \; E_- : B_{\mathbb{R}}^I, \; \mathbb{R} = \biguplus_{i \in I} E_i \vdash$$

$$\operatorname{Im} [E_i . \alpha_i]_{i \in I} \subseteq \bigcup_{i \in I} \operatorname{Im} \alpha_i \quad \text{ctbl} \quad \checkmark$$

$x : X \vdash$

$$[E_i . \alpha_i]_{i \in I}^{-1}(x) = \bigcup_{i \in I} \alpha_i^{-1}[x] \cap E_i \in B_{\mathbb{R}} \quad \checkmark$$

Prop: $X : \mathrm{Set},\ A : \mathrm{Qbs} \vdash$

- $\forall f : X \to \lfloor A \rfloor .\ f : \ulcorner X \urcorner^{\mathrm{Qbs}} \longrightarrow A$

- $\forall f : \lfloor A \rfloor \longrightarrow X .\ f : A \longrightarrow \llcorner X \lrcorner_{\mathrm{Qbs}}$

Prop: $X : \mathrm{Set},\ A : \mathrm{Qbs} \vdash$

$$\forall f : X \to \llcorner A \lrcorner .\ f : \ulcorner X^{\mathrm{Qbs}} \urcorner \longrightarrow A$$

Pf: $\alpha : R_{\ulcorner X^{\mathrm{Qbs}} \urcorner} \vdash$ α σ-simple $\Longrightarrow$

$$\alpha = [\alpha^{-1}[x].\lambda r.\, x]_{x \in \mathrm{Im}\,\alpha} \Longrightarrow$$

recombination

$$(f \circ \alpha) = [\alpha^{-1}[x].\lambda r.\, fx]_{x \in \mathrm{Im}\,\alpha} \in R_A \quad \checkmark$$

Borel

constant $\in B_A$

ctbl

Prop: $X : \mathrm{Set},\ A : \mathrm{Qbs} \vdash$

- $\forall f : X \longrightarrow \lfloor A \rfloor .\ \ f : \ulcorner X \urcorner^{\mathrm{Qbs}} \longrightarrow A$
- $\forall f : \lfloor A \rfloor \longrightarrow X .\ \ f : A \longrightarrow \llcorner X \lrcorner_{\mathrm{Qbs}}$

Prf: $\alpha : R_A \vdash (f \circ \alpha : \mathbb{R} \to X) \in R_{\llcorner X \lrcorner_{\mathrm{Qbs}}}$ always. ✓

$\blacksquare$

Useful adjunctions:

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

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

$$\ulcorner X \urcorner^{Meas} := \left\{ A \subseteq \lfloor X \rfloor \;\middle|\; \forall \alpha \in R_X .\ \alpha^{-1}[A] \in B_{\mathbb{R}} \right\}$$

- 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!

- $\lfloor Y^X \rfloor := \mathrm{Qbs}(X, Y)$

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

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

- $\mathrm{eval} : Y^X \times X \to Y$

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

Meas vs Qbs

By generalities:

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

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

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

No factorisation by Aumann's Theorem.

So: $\left(\ulcorner \mathbb{R}^{\mathbb{R}} \times \mathbb{R} \urcorner \not\cong \ulcorner \mathbb{R}^{\mathbb{R}} \urcorner \times \ulcorner \mathbb{R} \urcorner \right)$

Simple Type Structure

"Simple" because:

- Simply-typed λ-calculus
- types are simple: $A, B : \mathrm{Type} \vdash B^A : \mathrm{Type}$
  - no polymorphism
  - no term depending
- Contexts for terms: $\Gamma \vdash t : A$

  are simple: $\Gamma = x_1 : A_1, \ldots, x_n : A_n$

  i.e. $\mathrm{List}(\mathrm{Type})$

Simple Type Structure

"Simple" because:

- interpretation is simple:

$$[\![x_1 : A_1, \ldots, x_n : A_n]\!] := \prod_{i=1}^{n} A_i$$

$$[\![\Gamma \vdash t : A]\!] : [\![\Gamma]\!] \longrightarrow A$$

in Qbs

Simple Type Structure   Curry-Howard-Lambek

$$\frac{\Gamma \vdash t : A \qquad \Gamma \vdash s : B}{\Gamma \vdash \langle t, s \rangle : A \times B} \rightsquigarrow [\![\Gamma]\!] \xrightarrow{\lambda r.\, \langle tr, sr \rangle} A \times B$$

is measurable

measurability by type!

$$\frac{\Gamma \vdash t : A \times B \qquad \Gamma, x : A, y : B \vdash s : C}{\Gamma \vdash \mathrm{let}\ (x, y) = t\ \mathrm{in}\ s : C} \rightsquigarrow$$

$$[\![\Gamma]\!] \xrightarrow{\lambda r.\, \mathrm{let}\ (a, b) = tr\ \mathrm{in}\ s\, r[x \mapsto a, y \mapsto b]} C$$

is measurable. etc.

Random element Space

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

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

internalises as $\mathcal{B}_X = 2^X$, the qbs of

Borel subsets.

($\mathcal{B}_{\lfloor (\mathcal{B}_{\mathbb{R}}) \rfloor}$ 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

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

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

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:

$C_0' \in \mathbf{B}_{\mathbb{R}^{\mathbb{Q}}}$ 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 $C_0 \cong C_0' \in \mathbf{B}_{\mathbb{R}^{\mathbb{Q}}}$:

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

Example (ctd)

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

is measurable.

Avoids:

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

Non-examples ~~ [Sabok et al.'21]

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

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

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

Plan:

1) type-driven probability: discrete case (Mon + Tue)

2) Borel sets & measurable spaces (Wed)

3) Quasi Borel spaces (Wed) Simple type structure (Thu) ✓

4) Dependent type structure & standard Borel spaces ✓ (Thu)

5) Integration & random variables (Fri)

Please ask questions!

Course web page

scribble

Dependent Type Structure

types can contain terms :

$$X : \mathrm{Type},\ E : B_X \vdash \{x \in X \mid x \in E\} : \mathrm{Type}$$

a type referring to a term

a type, just like STLC

a term!

Dependent Type Structure

types can contain terms:

a type referring to a term

$$X : \mathrm{Type},\ E : B_X \vdash \{x \in X \mid x \in E\} : \mathrm{Type}$$

a type, just like STLC

a term!

Context formation:

$$\frac{\Gamma \vdash A : \mathrm{Type}}{\Gamma, x : A \vdash}$$

Dependent Type Structure

types denote spaces-in-context

$$
\begin{tikzcd}
[\![\Gamma \vdash A]\!] \arrow[d, "\mathrm{dep}"] \\
[\![\Gamma]\!]
\end{tikzcd}
$$

Dependent types denote spaces-in-context

$\Gamma \vdash$ context

$\Gamma \vdash A$

type in context

$[\![\Gamma \vdash A]\!]$ Space in context

$$\begin{array}{c} [\![\Gamma \vdash A]\!] \\ \big\downarrow \mathrm{dep} \\ [\![\Gamma]\!] \end{array}$$

context space

assigns environment

E.g.:

$$\begin{array}{c} A \\ \big\downarrow \\ \mathbb{1} \end{array}$$

simple types

$$[\![E : \mathcal{B}_A \vdash \{x \in A \mid x \in E\}]\!]$$

$$\begin{array}{c} \shortparallel \\ \{(E, a) \in \mathcal{B}_A \times A \mid a \in E\} \\ \big\downarrow \pi_1 \\ \mathcal{B}_A \end{array}$$

decoder

Context extension

$$\frac{\Gamma \vdash A}{\Gamma, a : A \vdash}$$

$$[\![\Gamma \vdash A]\!] \xrightarrow{\mathrm{dep}} [\![\Gamma]\!] \qquad [\![\Gamma, a : A]\!] := [\![\Gamma \vdash A]\!]$$

Substitution

$\Gamma \vdash \sigma : \Delta$ $\qquad [\![\sigma]\!] : [\![\Gamma]\!] \longrightarrow [\![\Delta]\!]$

E.g. weakening $\qquad \Gamma, a : A \vdash \mathrm{wkn} : \Gamma$

$$[\![\Gamma, a : A]\!] := [\![\Gamma \vdash A]\!] \xrightarrow[\mathrm{dep}]{\mathrm{wkn}} [\![\Gamma]\!]$$

Action of substitution on types

$$
\begin{array}{c}
[\![\Gamma \vdash A[\sigma]]\!] \\
\shortparallel
\end{array}
$$

$$
\begin{tikzcd}
\{(\gamma, a) \in [\![\Gamma]\!] \times [\![A]\!] \mid \mathrm{dep}\, a = \gamma\} \arrow[r, "\pi_2"] \arrow[d, "\mathrm{dep} := \pi_1"'] \arrow[dr, phantom, very near start, "\lrcorner"] & [\![\Delta \vdash A]\!] \arrow[d, "\mathrm{dep}"] \\
[\![\Gamma]\!] \arrow[r, "\sigma"] & \Delta
\end{tikzcd}
$$

E.g.

$$
\begin{tikzcd}
[\![x : A \vdash B[\mathrm{wkn}]]\!] := A \times B \arrow[r, "\pi_2"] \arrow[d, "\mathrm{dep} := \pi_1"'] \arrow[dr, phantom, very near start, "\lrcorner"] & B \arrow[d] \\
x : A \arrow[r, "\langle\rangle"'] & \mathbb{1}
\end{tikzcd}
$$

simple type

Terms : sections

$$\begin{tikzcd}
{[\![\Gamma]\!]} \arrow[r, "{[\![\Gamma \vdash M : A]\!]}"] \arrow[dr, equal, "="'] & {[\![\Gamma \vdash A]\!]} \arrow[d, "\mathrm{dep}"] \\
 & {[\![\Gamma]\!]}
\end{tikzcd}$$

e.g.

$$\begin{tikzcd}
\mathbb{R} \arrow[r, "{[\![x : \mathbb{R} \vdash [x, \infty) : B_{\mathbb{R}}[\mathrm{wkn}]]\!]}"] \arrow[dr, equal, "="'] & \mathbb{R} \times B_{\mathbb{R}} \arrow[d, "\pi_1"] \\
 & \mathbb{R}
\end{tikzcd}$$

E.g. Variables:

$$\begin{tikzcd}
{[\![\Gamma, a : A]\!]} \arrow[r, "{[\![\Gamma, a : A \vdash a : A]\!]}", "{\langle \mathrm{id}, \mathrm{dep}_{\Gamma \vdash A} \rangle}"'] \arrow[dr, equal, "="'] & {[\![\Gamma, a : A \vdash A[\mathrm{wkn}]]\!]} \arrow[d, "\mathrm{dep}"] \\
 & {[\![\Gamma, a : A]\!]}
\end{tikzcd}$$

Exercise:
action of substitution
$M[\sigma]$

Dependent Pairs

$$\frac{\Gamma, a : A \vdash B}{\Gamma \vdash \coprod_{a \in A} B}$$

$$\left[\!\left[\coprod_{a \in A} A \right]\!\right] := \left[\!\left[\Gamma, a : A \vdash B \right]\!\right]$$

$$\begin{array}{c}
\left[\!\left[\Gamma, a : A \vdash B \right]\!\right] \\
\Big\downarrow \mathrm{dep}_B \\
\left[\!\left[\Gamma, a : A \right]\!\right] \\
\| \\
\left[\!\left[\Gamma \vdash A \right]\!\right] \\
\Big\downarrow \\
\left[\!\left[\Gamma \right]\!\right]
\end{array}$$

$$\mathrm{dep}_{\coprod} := \ \left[\!\left[\coprod_{a \in A} A \right]\!\right] \longrightarrow \left[\!\left[\Gamma \right]\!\right]$$

Dependent Products

aha: $(a:A) \to B$

$$\frac{\Gamma, a:A \vdash B}{\Gamma \vdash \prod_{a \in A} B}$$

$$\left\lfloor [\![\Gamma \vdash \prod_{a \in A} B]\!] \right\rfloor :=$$

$$\left\{ \left(r_0, f : \{a \in [\![A]\!] \mid \mathrm{dep}\, a = r_0\} \to [\![\Gamma, a:A \vdash B]\!]\right) \middle| \begin{array}{l} \\ \forall a \in [\![\Gamma, a:A]\!].\ \mathrm{dep}\, a = r_0 \Rightarrow \mathrm{dep}(f\, a) = a \end{array} \right\}$$

Exercise: find the random elements.

Full model

type : Qbs $\quad W := [0,\infty] \quad \mathbb{B}X \cong \mathbb{B}^X$

$DX := (\text{Fri})$

$PX := \{\mu \in DX \mid \underset{\mu}{\text{Ce}}[X] = 1\}$

$\underset{\mu}{\text{Ce}}[E] := (\text{Fri}) \qquad \delta_x := (\text{Fri})$

$\oint \mu k := (\text{Fri})$

Plan:

1) type-driven probability: discrete case (Mon + Tue)

2) Borel sets & measurable spaces (Wed)

3) Quasi Borel spaces (Wed) Simple type structure (Thu)

4) Dependent type structure & standard Borel spaces (Thu)

5) Integration & random variables (Fri)

Please ask questions!

Course web page

Scribble

Foundations for type-driven probabilistic modelling

Ohad Kammar
University of Edinburgh

Logic Summer School
Australian National University
4–16 December, 2023
Canberra, ACT, Australia

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

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: $e[\lfloor X \rfloor] \in B_Y$
- e is Strong: $\alpha \in R_X \iff e \circ \alpha \in R_Y$

Examples

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

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

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

Its domain of definition

$$f : (Y \amalg \{\bot\})^X \vdash \mathrm{Dom}\, f := \{x \in X \mid f x \neq \bot\} : \mathrm{Type}$$

Dependent-type interpretation

$$\begin{tikzcd}
{[\![\mathrm{Dom}\, f]\!]} \arrow[rr] \arrow[d, "\mathrm{dep}"'] \arrow[drr, phantom, very near start, "\lrcorner"] & & \{y \in Y \mid y \in E\} \arrow[d, "\mathrm{dep}"] \\
{[\![f : (Y \amalg \{\bot\})^X]\!]} \arrow[rr, "{[E \mapsto \lambda x.\, f x \neq \bot]}"] & & {[\![E : B_Y]\!]}
\end{tikzcd}$$

Plan:

1) type-driven probability: discrete case (Mon + Tue)

2) Borel sets & measurable spaces (Wed)

3) Quasi Borel spaces (Wed) Simple type structure (Thu)

4) Dependent type structure & standard Borel spaces (Thu)

5) Integration & random variables (Fri)

Please ask questions!

Course web page

Scribble

Full model

type : Qbs $W := [0, \infty]$ $BX \cong \mathbb{B}^{X}$

$DX := (\text{Fri})$

$PX := \{ \mu \in DX \mid \mathrm{Ce}_{\mu}[X] = 1 \}$

$\mathrm{Ce}_{\mu}[E] := (\text{Fri})$ $\delta_{x} := (\text{Fri})$

$\oint \mu k := (\text{Fri})$

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

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

satisfying the measure axioms:

$$E_- : B^{W} \vdash$$

$$\mu\emptyset = 0, \quad \mu E = \mu(E \cap F) + \mu(E \cap F^c), \quad \mu\left(\bigcup_n E_n\right) = \sup_n \mu E_n$$

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

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

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

Thm (Lebesgue measure):

There is a unique measure $\lambda \in \lfloor G\mathbb{R} \rfloor$ s.t.:

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

Thm (Lebesgue measure):

There is a unique measure $\lambda \in \lfloor G\mathbb{R} \rfloor$ s.t.:

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

Proof sketch (standard analysis textbook):

1) restrict attention to $(0,1]$ & extend via σ-additivity

2) Take $\Sigma_0 \subseteq \mathcal{B}_{(0,1]}$ $E \in \Sigma_0 \Leftrightarrow E = \bigcup_{i=1}^{n} (a_i, b_i]$

3) Defining $\lambda : \Sigma_0 \to \mathbb{W}$, $\lambda \bigcup_{i=1} (a_i, b_i] := \sum_i (b_i - a_i)$ independent of

4) $\lambda \phi = 0$, $\lambda E = \lambda(E \cap F) + \lambda(E \cap F^c)$ (straightforward)

$\hookrightarrow$

5) Technical gadget: $\forall (E_n \supseteq E_{n+1})$ in Σ_0,

$$\inf \lambda E_n > 0 \implies \bigcap E_n \neq \emptyset.$$

6) λ is continuous on Σ_0: If $(E_n \subseteq E_{n+1})_n$ in Σ_0

and $\bigcup_n E_n \in \Sigma_0$ then $\lambda \bigcup E_n = \sup_n \lambda E_n$

7) Noting that: Σ_0 is a Boolean algebra

& $\sigma(\Sigma_0) = \mathcal{B}_{(0,1]}$

we use Caratheodory's extension thm:

λ extends uniquely to $\lambda : \mathcal{B}_{(0,1]} \to \mathbb{W}$.

The unrestricted Giry spaces

Equip $\lfloor GV \rfloor$ with two qbs structures:

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

↪ α is a kernel.

✓ $GV \hookrightarrow W^{B_X}$

- Fewer random elements $R_{GV} \subseteq R_{G'V}$
- Lebesgue integral measurable in both arguments. (upcoming)

Farewell Meas

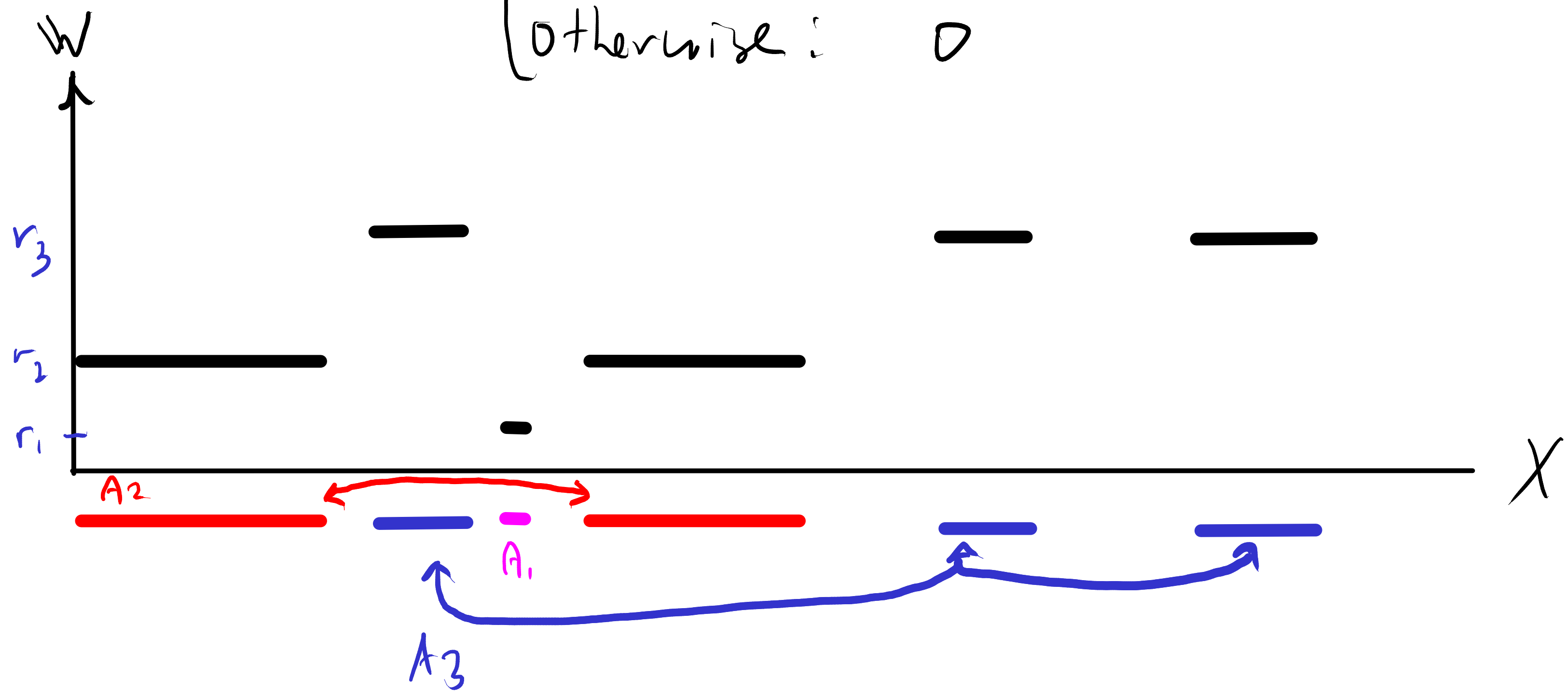
Now on:

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:

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

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

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 → remember! QBS morphism!

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

$f_n \in \mathrm{Simple}$.

Moreover, we have measurable such choice:

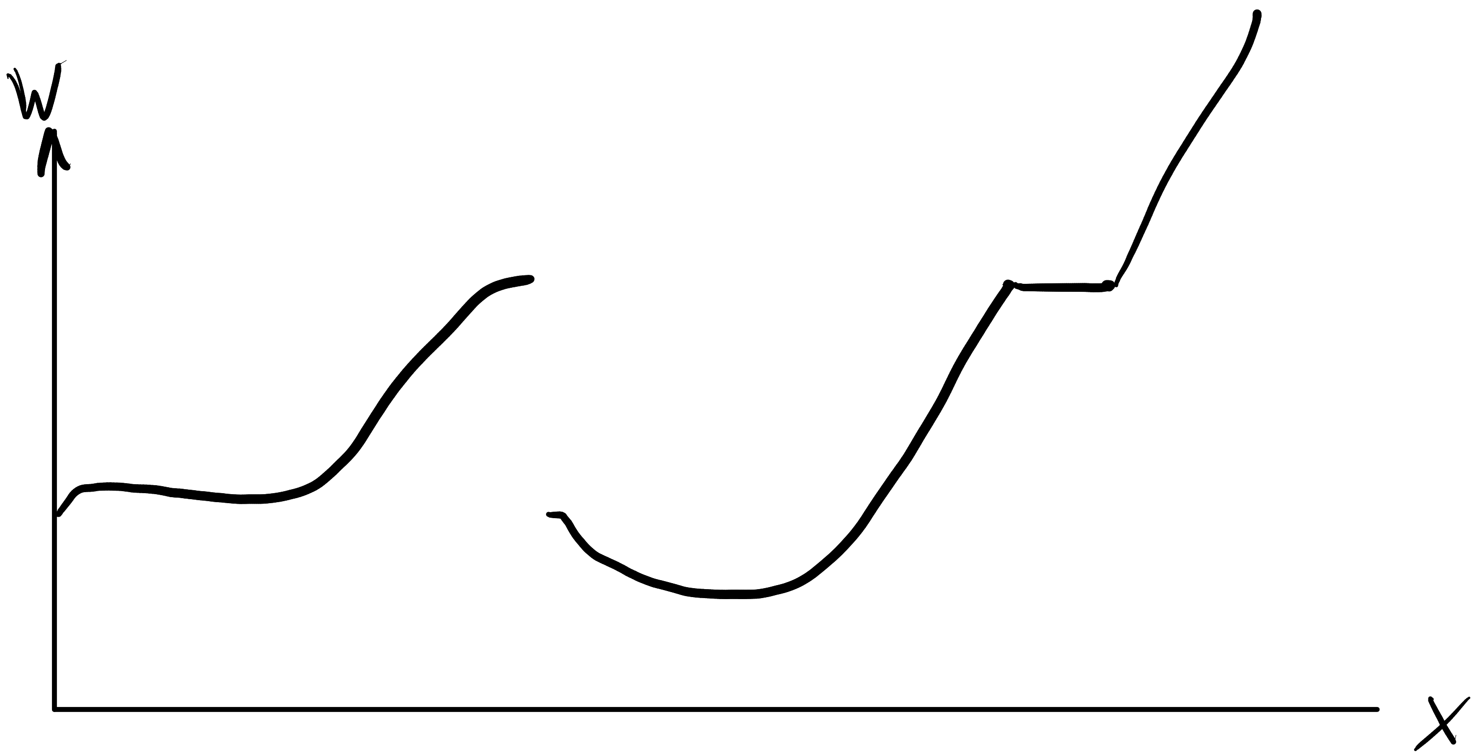
Simple Approx :

$$\{\vec{\Delta} \in \mathbb{R}^+ \mid \Delta_n \to 0\} \times \{\vec{\alpha} \in W^{\mathbb{N}} \mid \vec{\alpha} \text{ monotone}, \alpha_n \to \infty\} \times W^X \longrightarrow \mathrm{SimpleCode}$$

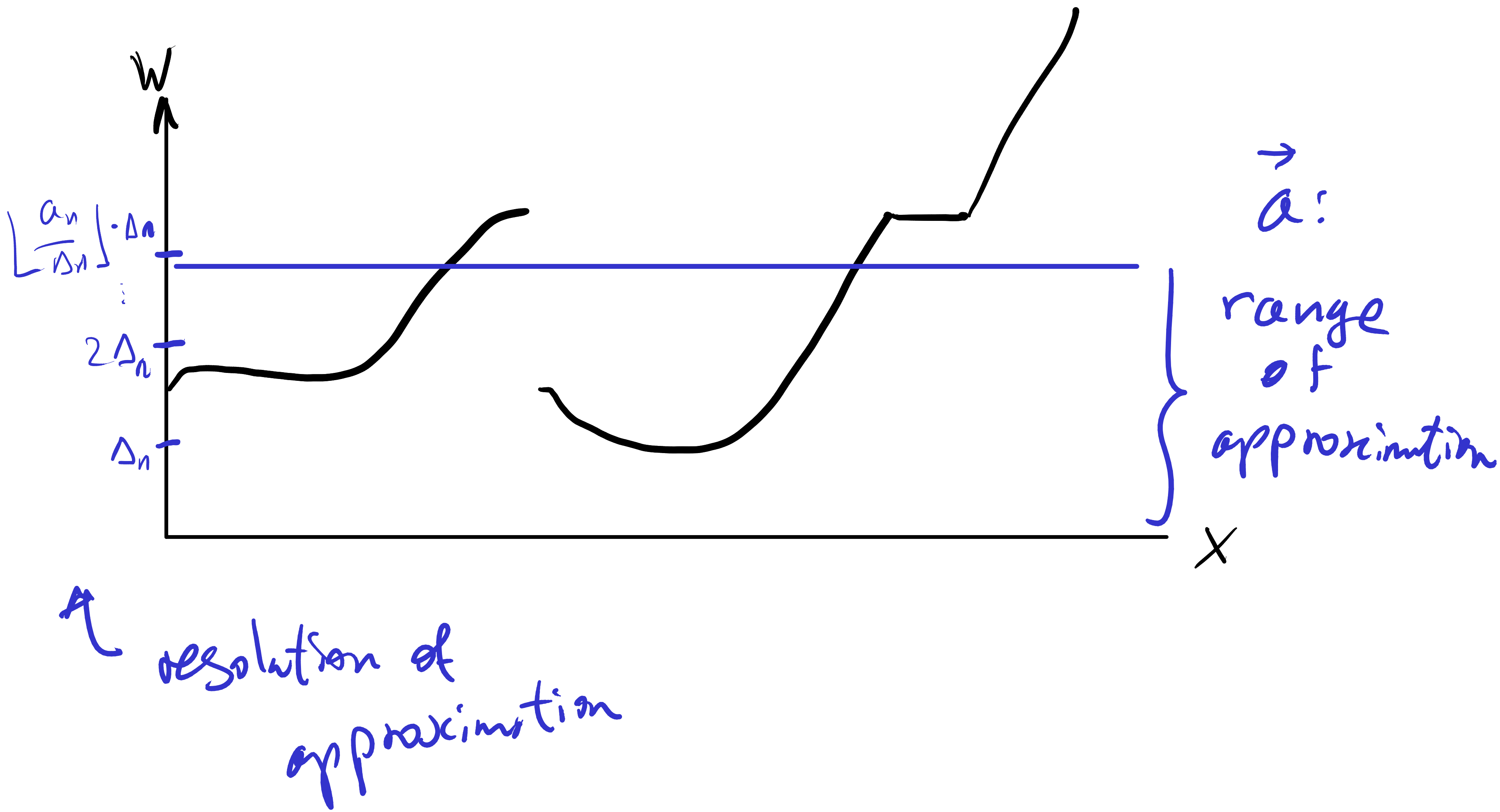
rate of convergence

range of approximation

Why?

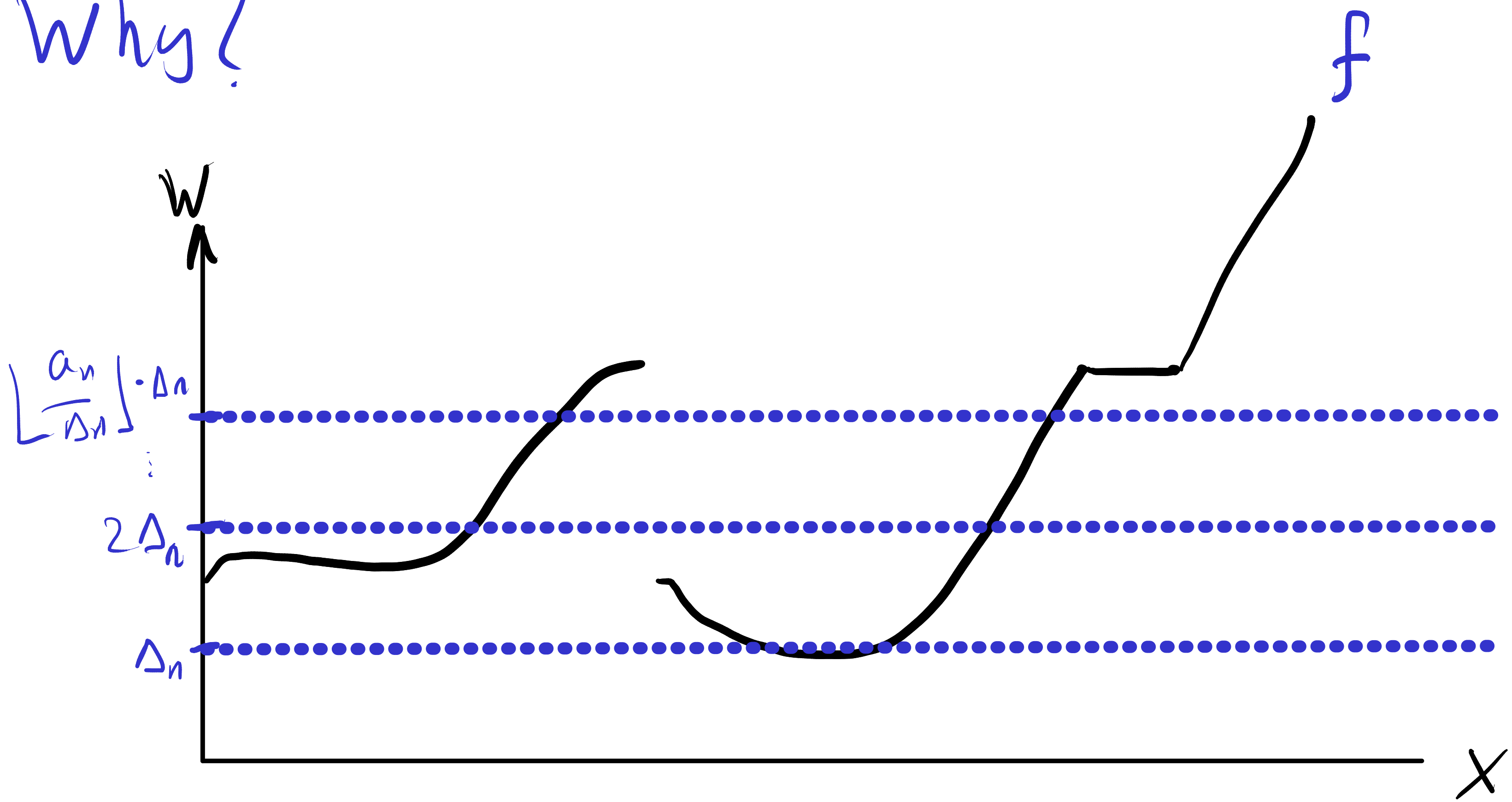


Why?

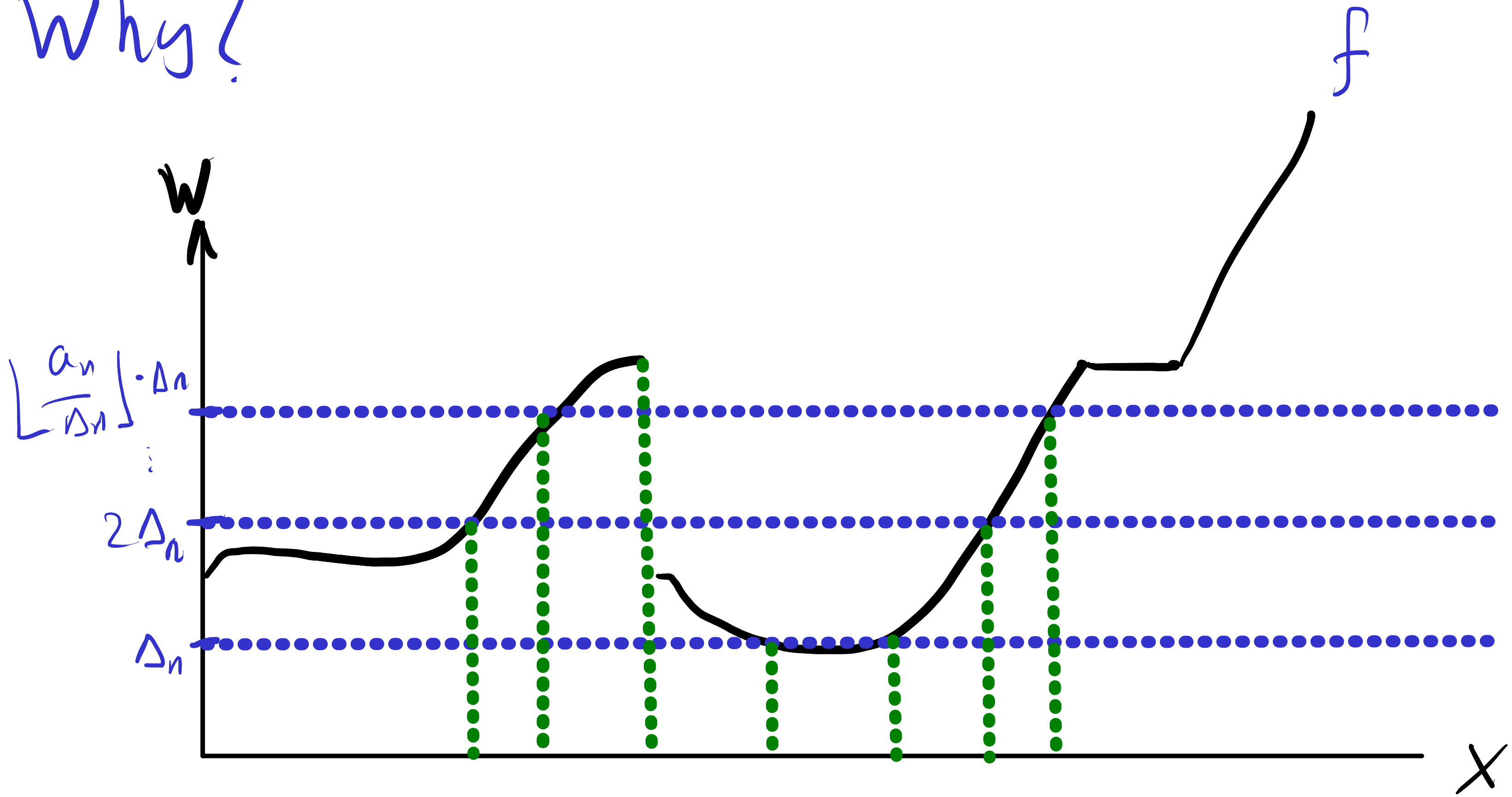


f
W
$\lfloor \frac{a_n}{\Delta_n} \rfloor \cdot \Delta_n$
$2\Delta_n$
Δ_n
X
$\vec{a}$:
range of f
approximation
resolution of approximation

Why?

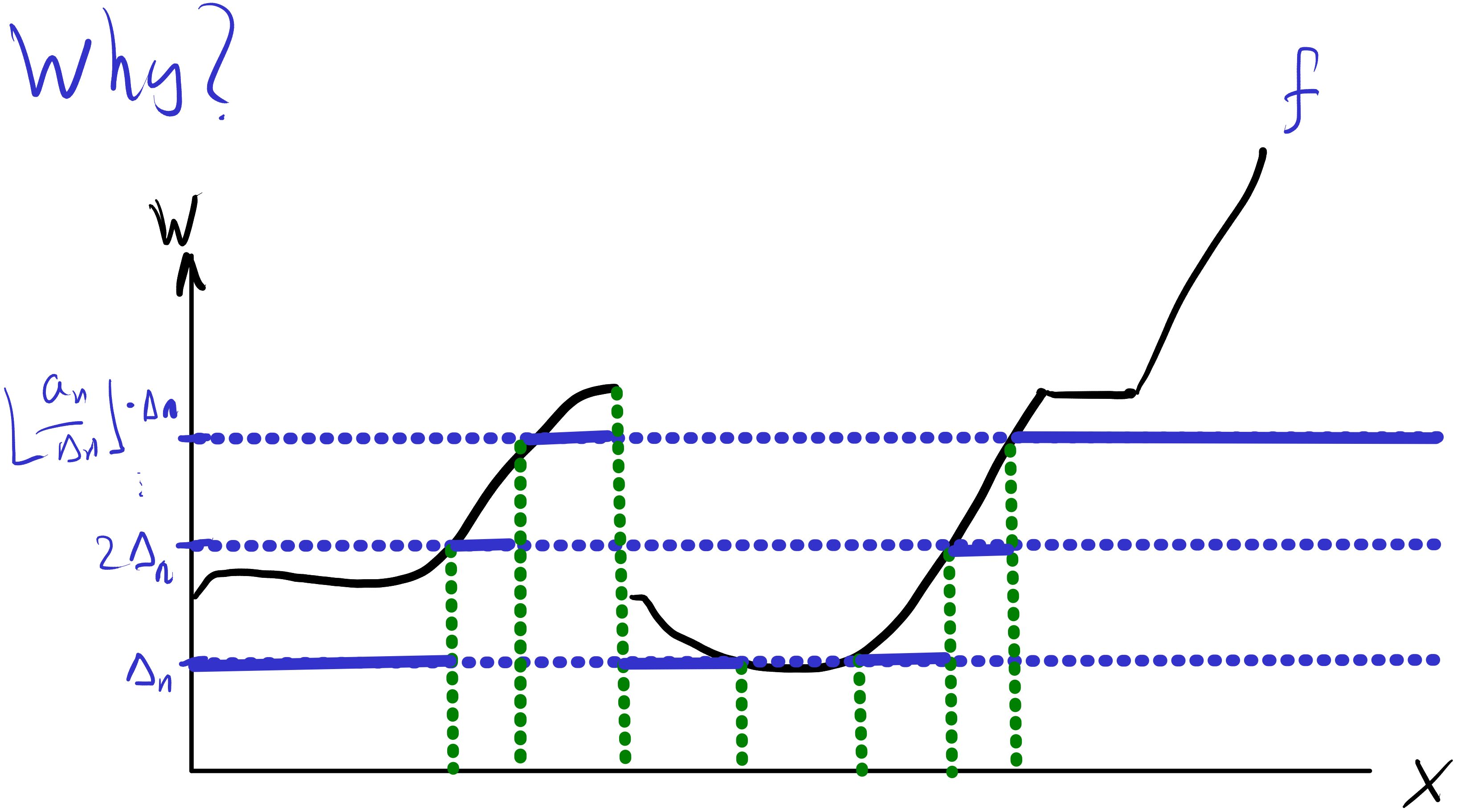

f
W
$\lfloor \frac{a_n}{\Delta_n} \rfloor \cdot \Delta_n$
$2\Delta_n$
Δ_n
X

Why?



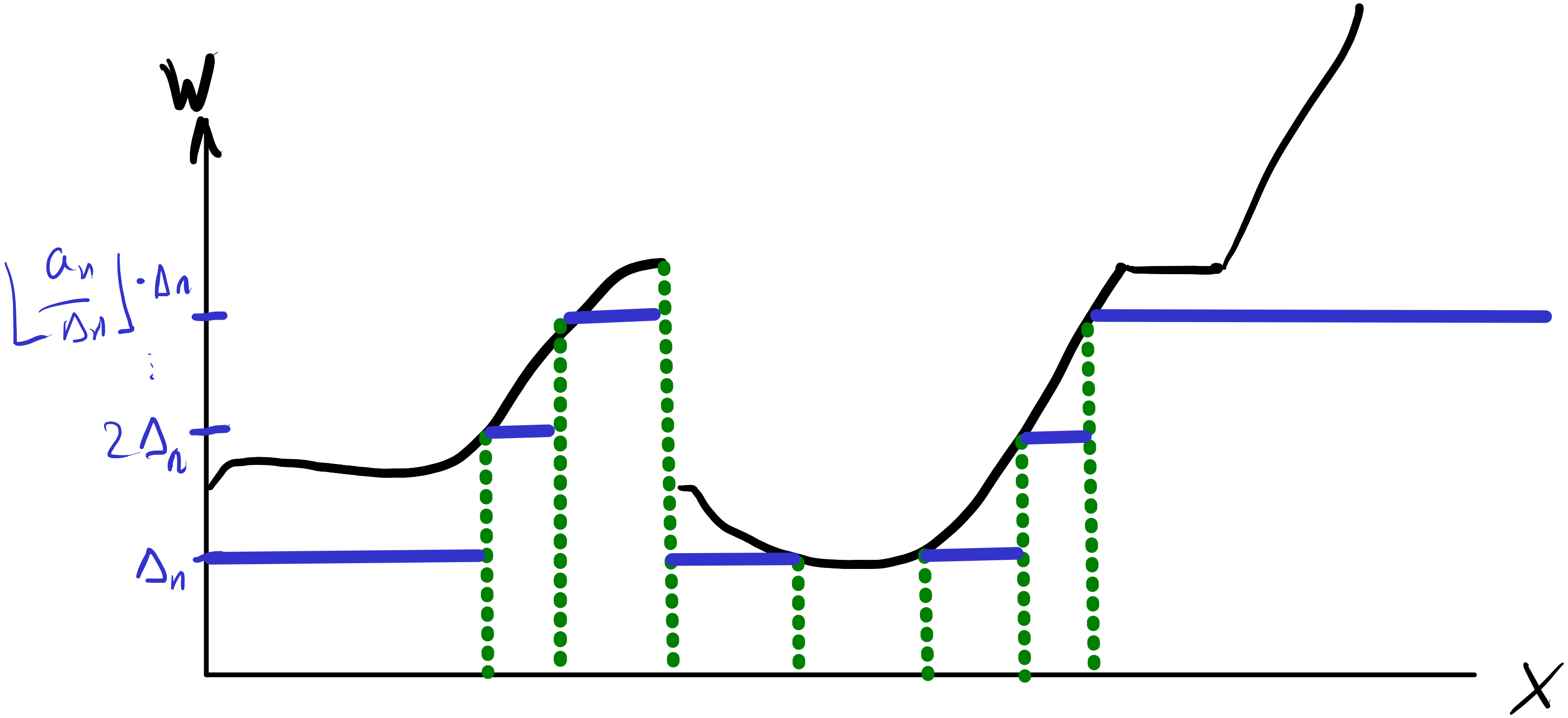
f
W
$\lfloor \frac{a_n}{\Delta_n} \rfloor \cdot \Delta_n$
$2\Delta_n$
Δ_n
X

Why?



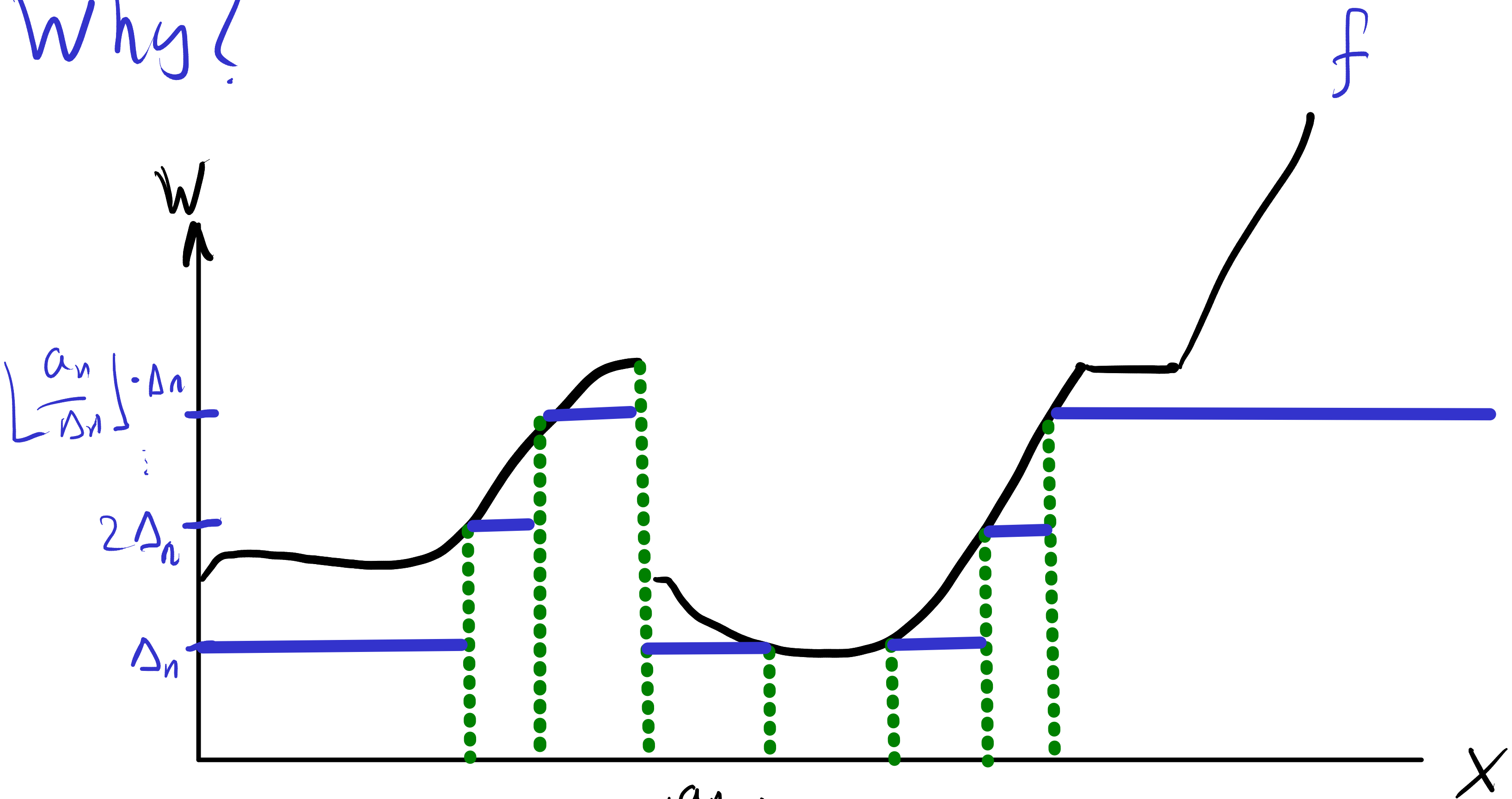
f
W
$\lfloor \frac{a_n}{\Delta_n} \rfloor \cdot \Delta_n$
$2\Delta_n$
Δ_n
X

Why?



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

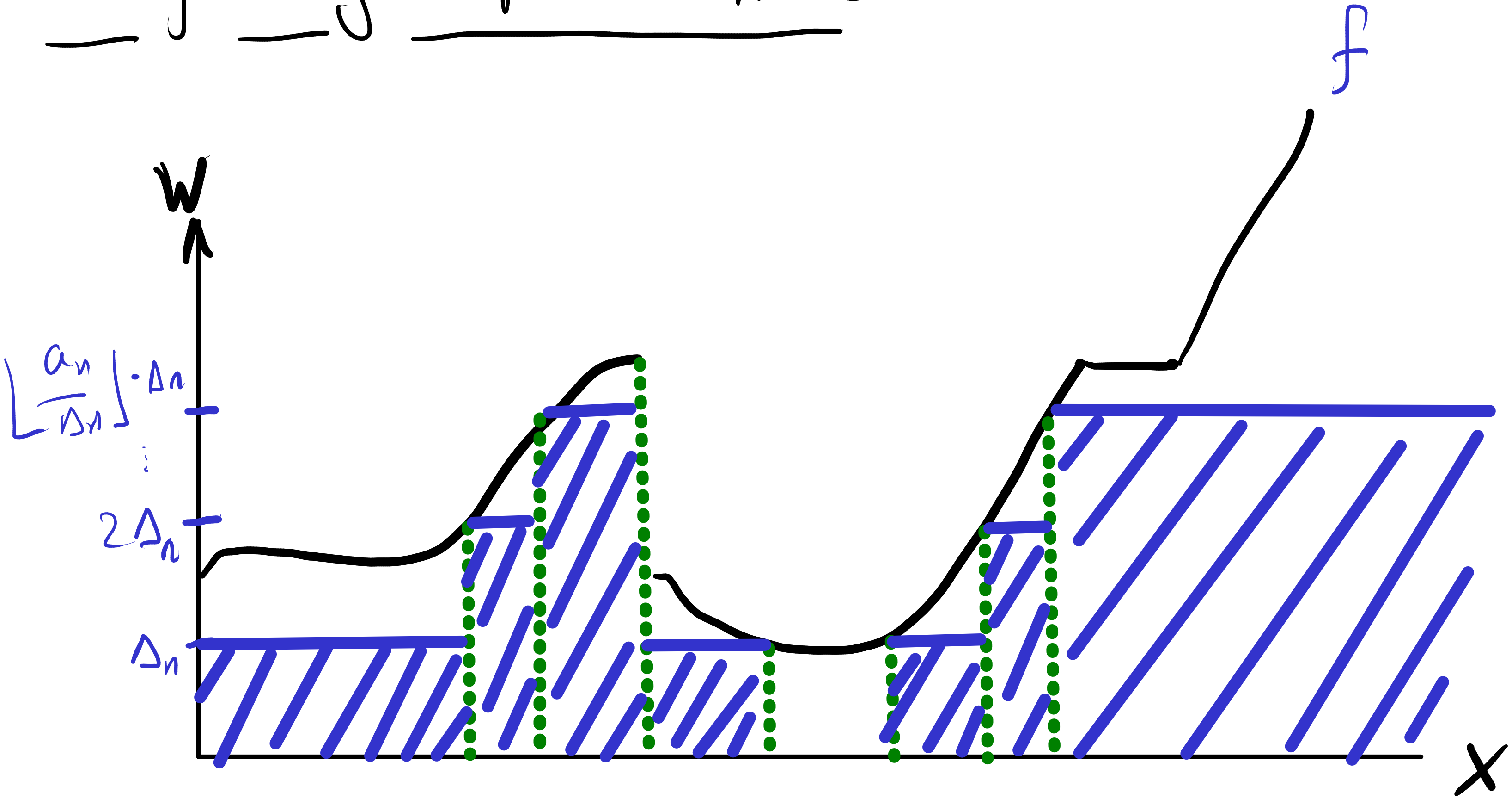
Why?



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

$$+ \lfloor \frac{a_n}{\Delta_n} \rfloor \Delta_n \cdot [\, f \geq \lfloor \frac{a_n}{\Delta_n} \rfloor \cdot \Delta_n \,] \in \mathrm{Simple}$$

Integrating Simple Functions



$$\int : GX \times \text{Simple Code} \longrightarrow W$$

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

Integration

Proper higher-order operation

$$\int : GX \times W^X \longrightarrow W$$

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

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

measurable by type

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

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:

$$\delta : X \to GX$$

$$x \mapsto \lambda A. \begin{cases} x \in A : & 1 \\ x \notin A : & 0 \end{cases}$$

but: non-commutative

(Fubini fails, just like in Meas)

Kleisli extension / Kock integral:

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

$$\oint \mu f := \lambda A. \int \mu(dx)\, f(x; A)$$

Fubini-Tonelli fails

$\# \in G\mathbb{R}$ $\qquad \#E := \begin{cases} E \text{ finite}: & |E| \\ \text{o.w.}: & \infty \end{cases}$

$\lambda \in G\mathbb{R}$ $\qquad$ lebesgue $\qquad k : \mathbb{R}\times\mathbb{R} \to W \cong G\mathbb{1}$

$k(x,y) := [x=y]$

$$\oint \#(dr) \underbrace{\oint \lambda(dx) k(x,y)} = \oint \# \underline{0} = \underline{0} \cong 0$$

$y : \mathbb{R} \vdash \{\langle\rangle\} \mapsto \lambda\{y\}\cdot 1 + \lambda\{y\}^c \cdot 0 = 0$

$\#$

$$\oint \lambda(dx) \underbrace{\oint \#(dr)\ K(x,y)} = \oint \lambda(dx)\, \delta_{\langle\rangle} \cong \infty$$

$x : \mathbb{R} \vdash \{\langle\rangle\} \mapsto \#\{x\}\cdot 1 + 0 = 1$

Randomisable measures monad

$\lambda A. \int_{\mathrm{Dom}\,\alpha} \lambda(\mathrm{Dom}\,\alpha)$

$$D \rightarrowtail G$$

$$\lfloor DX \rfloor := \{ \lambda_\alpha \mid \alpha : \mathbb{R} \to X \}$$

Lebesgue measure

$$R_{DX} := \{ \lambda x.\, \lambda_{\alpha x} \mid \alpha : \mathbb{R} \times \mathbb{R} \to X \}$$

$\delta : X \to DX$ $\quad \oint : D\Gamma \times (DX)^\Gamma \to DX$ lift along $D \rightarrowtail G$.

D validates our measure axioms including Fubini-Tonelli

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

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

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

and agree with their counterparts on Meas.

$$\mathcal{D}S_{\mathcal{J}} = \{\mu \mid \mu \text{ s-finite}\}$$

See [Staton'16]

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

Open: Is there a counterpart to D in Meas?

More modestly, is $DS \in Sbs$?

(Hypothesis: no)

Distribution Submonads

$$P X := \{\mu \in D X \mid \mu X = 1\}$$

$$\downarrow$$

$$P_{\leq 1} X := \{\mu \in D X \mid \mu X \leq 1\}$$

$$\downarrow$$

$$P_{<\infty} X := \{\mu \in D X \mid \mu X < \infty\}$$

$$\downarrow$$

$$D X$$

A measure space

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

is a qbs Ω with

$\mu \in D X$.

Similarly: - finite measure space

- (sub) probability space.

Full model

type : Qbs $\quad W := [0,\infty] \quad \mathbb{B}X \cong \mathbb{B}^X$

$$DX := \left(\{\lambda_\alpha \mid \alpha : \mathbb{R} \to X\}, \{\lambda r.\lambda_{\alpha(r,-)} \mid \alpha : \mathbb{R}\times\mathbb{R} \to X\} \right)$$

$$PX := \{\mu \in DX \mid \underset{\mu}{Ce}[X] = 1\}$$

$$\underset{\mu}{Ce}[E] := \mu E \qquad \delta_x := E \mapsto \begin{cases} x \in E : 1 \\ x \notin E : 0 \end{cases}$$

$$\oint \mu k := \lambda E. \int \mu(dx) k(x; E)$$

Plan:

1) type-driven probability: discrete case (Mon + Tue)

2) Borel sets & measurable spaces (Wed)

3) Quasi Borel spaces (Wed) Simple type structure (Thu)

4) Dependent type structure & standard Borel spaces (Thu)

5) Integration & random variables (Fri)

Please ask questions!

Course web page

scribble

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

Θ:

$$
\begin{tikzcd}
 & & \mathbb{W} \arrow[dr] & \\
\mathbb{I} \arrow[r] & \mathbb{R}_{\geq 0} \arrow[ur] \arrow[dr] & & \overline{\mathbb{R}} \\
\mathbb{R}^{+} \arrow[ur] & & \mathbb{R} \arrow[ur] &
\end{tikzcd}
$$

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

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

$$\alpha\xi + \zeta := \lambda\omega.\ \alpha\cdot\xi\omega + \zeta\omega$$

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

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

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

$p : P\Omega \vdash$ "$p\,x$ holds $\lambda(dx)$-almost surely"

for some $Q \hookrightarrow \Omega$, $p \supseteq Q$, $[- \in Q] \cdot \lambda = \lambda$

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

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

Integrating random variables (as discretely) in Qbs!

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

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

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

$$\lambda : P\Omega \vdash \mathrm{ASConverge}(\overline{\mathbb{R}})^{\Omega} : \mathcal{B}\left(\overline{\mathbb{R}}^{\mathbb{N}\times\Omega}\right)$$

$$:= \left\{ \vec{\xi} \in \overline{\mathbb{R}}^{\mathbb{N}\times\Omega} \;\middle|\; \Pr_{\omega \sim \lambda}\left[\lim \xi_n \omega \neq \bot\right] \right\}$$

So:

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

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

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

Thm (monotone convergence):

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

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

$$\Downarrow$$

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

Lebesgue space

(Ω Prob. Space, $p \in [1, \infty)$)

$$p : [1, \infty), \lambda : P\Omega \vdash \mathcal{L}^{p}_{(\Omega,\lambda)} : B(\mathbb{R}^{\Omega})$$

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

Ensemble $\mathcal{L}_{\Omega} := \prod_{\substack{\lambda \in P\Omega \\ p \in [1,\infty)}} \mathcal{L}^{p}_{(\Omega,\lambda)}$

$p \leq q \Rightarrow \mathcal{L}^{p}_{\Omega} \supseteq \mathcal{L}^{q}_{\Omega}$

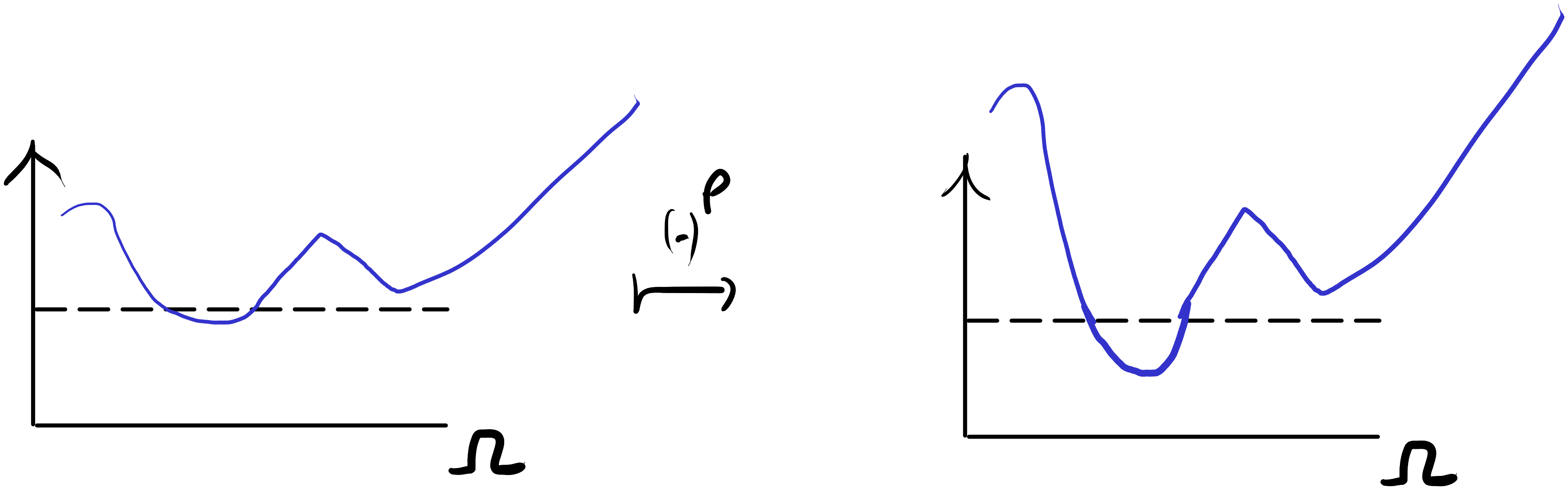
$\mathcal{L}^p$ semi norms

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

$\mathcal{L}^2$ inner product

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

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



Statistics

Expectation

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

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

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

$p : [1,\infty), \vdash : p \times \vdash$ $\mathrm{Cauchy}\,\mathcal{L}^p_{(\Omega,\mu)} : \mathcal{B}(\mathcal{L}^p_{(\mathbb{R},\mu)})^{\mathbb{N}}$

$$:= \left\{ \vec{\xi} \;\middle|\; \forall \varepsilon \in \mathbb{Q}^+\, \exists k \in \mathbb{N}\, \forall m,n \geq k.\; \|\xi_{k+n} - \xi_{k+m}\|_p < \varepsilon \right\}$$

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

$$\lim^p : \mathrm{Cauchy}\,\mathcal{L}^p \longrightarrow \mathcal{L}^p$$ (convergence in mean)

Why?

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

Example

Thm (dominated convergence)

For $\vec{Z}, Z \in \mathcal{L}^1$ s.t. $Z_n \leq Z$ a.s.:

1. $\lim^{as} \vec{Z} \in \mathcal{L}^1$
2. $\lim^{1} \vec{Z} = \lim^{as} \vec{Z}$
3. $\lim_{n\to\infty} \int \lambda Z_n = \int \lambda \lim_{n\to\infty} Z_n$

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

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

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

Prop. - every sbs S measurably separable in

$$PS \times [1, \infty)$$

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

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

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

$\|\beta_n\|_2 = 1$

(β_n) dense

Example

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

orthogonal decomposition — linear in fact.

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

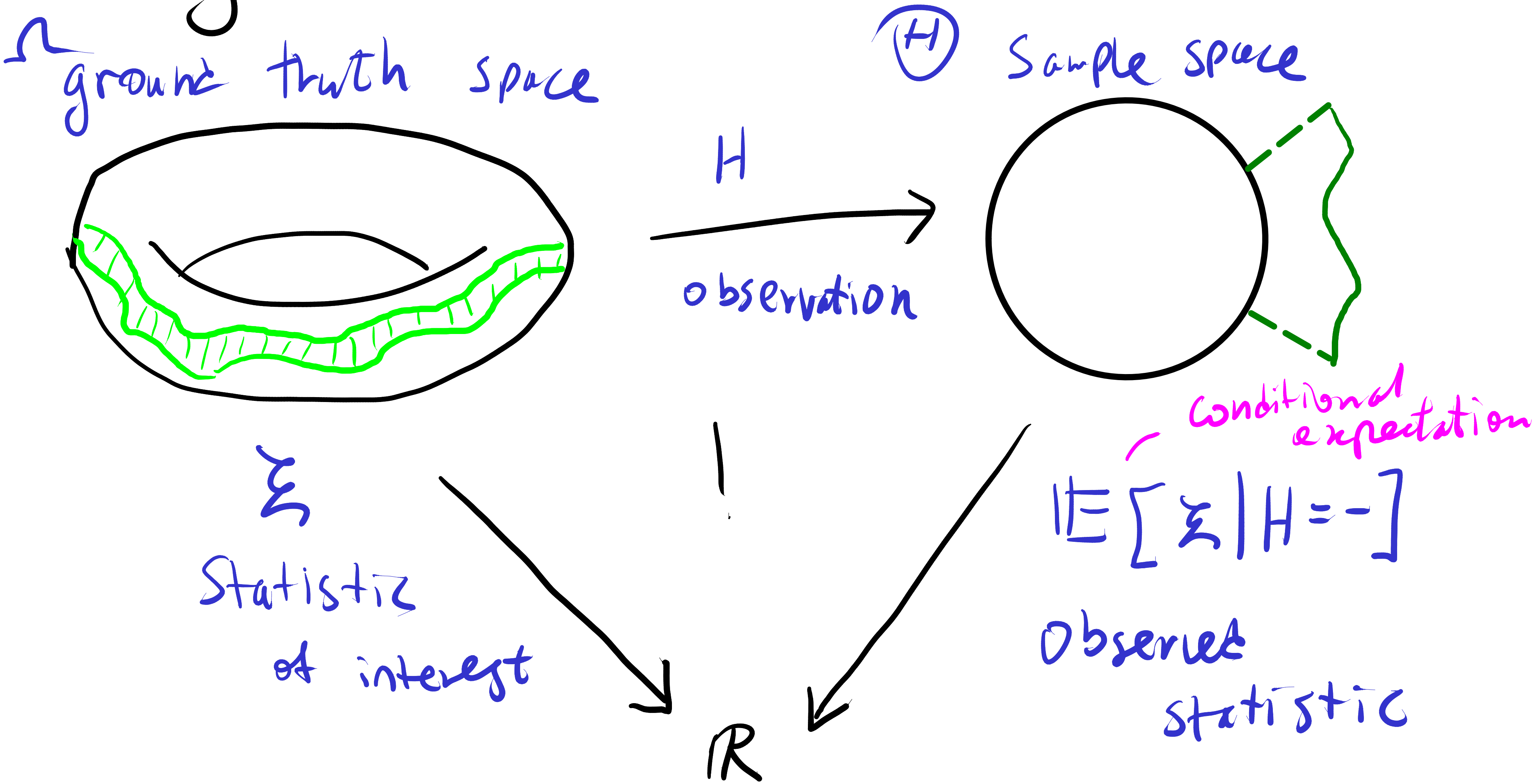
When S is separable with orthonormal system β

We have a measurable version of

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

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

Kolmogorov's Conditional Expectation



Kolmogorov's Conditional Expectation

A conditional expectation

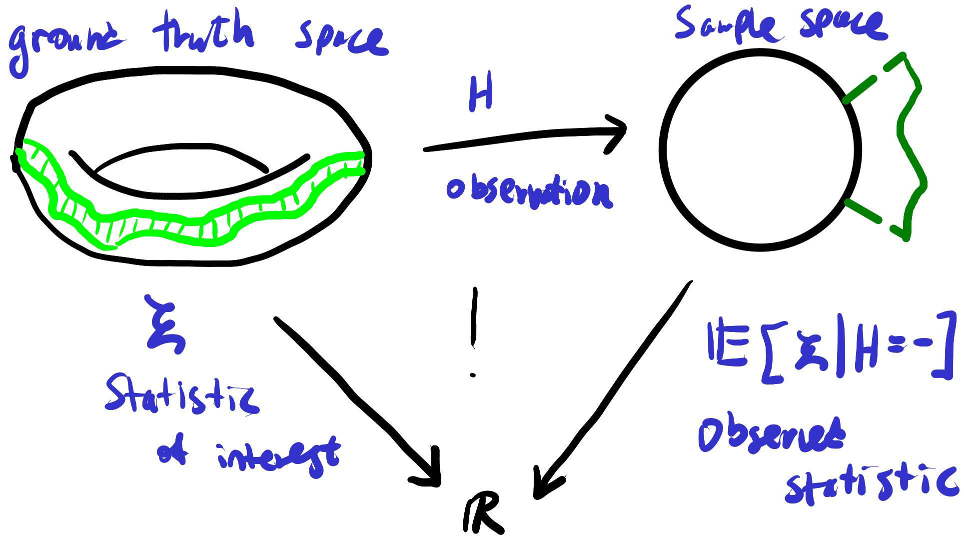
of $\xi \in \mathcal{L}^1_\Omega$ wrt

$H : \Omega \to H$ is

$Z \in \mathcal{L}^1_{H}$ s.t. for all $A \in \mathcal{B}_{H}$:

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

where $\mu := \lambda_H$



Conditional expectations

1. unique a.s.

2. fundamental to modern Probability, eg:

   a martingale

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

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

Thm (Existence)

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

- When (Ω,λ) is Separable

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

- When Θ is $\mathcal{I}'$-measurably separable

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

Plan:

1) type-driven probability: discrete case (Mon + Tue)
2) Borel sets & measurable spaces (Wed)
3) Quasi Borel spaces (Wed) Simple type structure (Thu)
4) Dependent type structure & standard Borel spaces (Thu)
5) Integration & random variables (Fri)

Course web page

Discrete model

$\text{type} : \text{Set} \qquad W := [0, \infty] \qquad BX := \mathcal{P}X$

$$DX := \{ \mu : X \to W \mid \text{supp}\, \mu \text{ countable} \}$$

$$PX := \{ \mu \in DX \mid \mathrm{Ce}_\mu [X] = 1 \}$$

$$\mathrm{Ce}_\mu [E] := \sum_{x \in E} \mu x \qquad \delta_x := \lambda x'. \begin{cases} x = x' : & 0 \\ x \neq x' : & 1 \end{cases}$$

$$\oint \mu k := \lambda x. \sum_{r \in \Gamma} \mu r \cdot k(r; x)$$

Full model

$\text{type} : \text{Qbs} \qquad \mathbb{W} := [0,\infty] \qquad BX \cong \mathbb{B}^X$

$$DX := \left(\{ \lambda_\alpha \mid \alpha : \mathbb{R} \to X \}, \{ \lambda r. \lambda_{\alpha(r,-)} \mid \alpha : \mathbb{R} \times \mathbb{R} \to X \} \right)$$

$$PX := \{ \mu \in DX \mid \underset{\mu}{\mathcal{E}} [X] = 1 \}$$

$$\underset{\mu}{\mathcal{E}} [E] := \mu E \qquad \delta_x := E \mapsto \begin{cases} x \in E : 1 \\ x \notin E : 0 \end{cases}$$

$$\oint \mu k := \lambda E. \int \mu(dx) k(x; E)$$