Foundations for type-driven probabilistic modelling

Ohad Kammar
University of Edinburgh

TutorialFest
ACM-SIGPLAN 51st Symposium on
Principles of Programming Languages (POPL'24)
15 January, 2024
City of Westminister, England

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

Discrete model

Recap

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

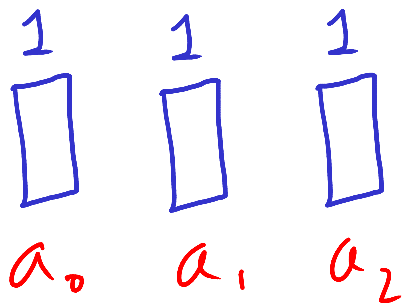
$$DX := \{\mu : X \to W \mid \mathrm{supp}\,\mu \text{ countable}\}$$

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

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

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

Ex. measures

1) 

. . . Counting measure $\#_X : DX$

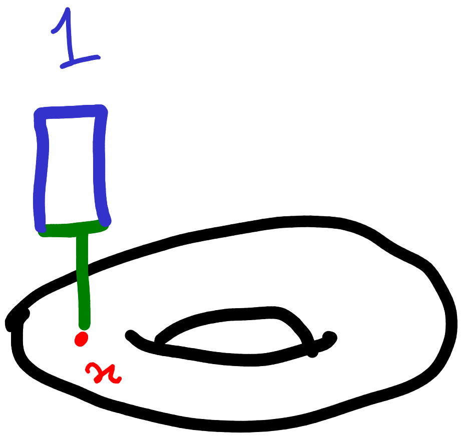
. . . (X ctbl)

$\#_X := \lambda x:X.1$

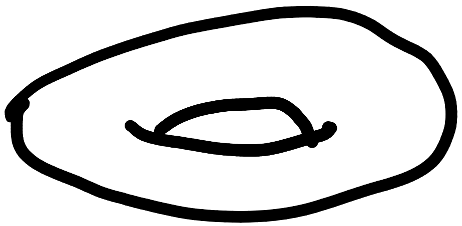
2) Dirac measure:

$x:X \vdash \delta_x : DX$

$$:= \lambda x'. \begin{cases} x=x' : & 1 \\ o.w. : & 0 \end{cases}$$



δ_x

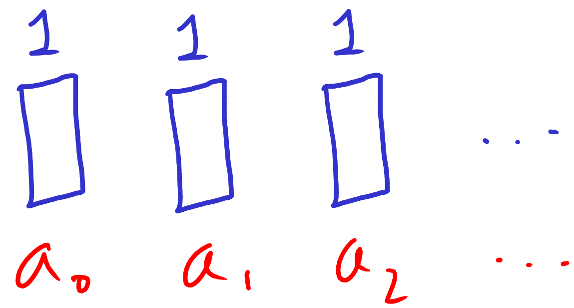
3) 

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

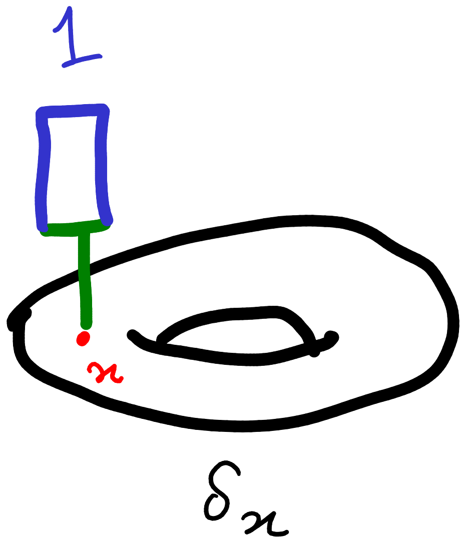
Ex distributions

Recap

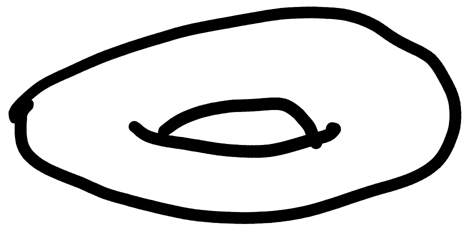
Counting measure (X ctbl): $\#_X := \lambda x. 1$



Dirac measure δ_x (prev slide)



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



Product measures

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

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

discrete model

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

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: represent configurations of 2 dice using

$\mathrm{Die}_6 := \{1, 2, \ldots, 6\}$ $\quad \mathrm{Die}_6^2$

Letting $(+) : \mathrm{Die}_6^2 \longrightarrow \mathbb{N}^2 \xrightarrow{(+)} \mathbb{N}$

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

$$\mu := \left(\#_{\mathrm{Die}} \otimes \#_{\mathrm{Die}}\right)_{(+)} : D\mathbb{N}$$

build measures compositionally

$\mu s =$ number of outcomes whose sum is s

Scaling a measure

$$(\cdot) : \mathbb{W} \times DX \longrightarrow DX$$

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

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

$\mu : DX \vdash$ $\quad 1 \cdot \mu = \mu$ $\qquad\qquad$ $a, b : \mathbb{W}, \mu : DX \vdash$ $\quad a \cdot (b \cdot \mu) = (a \cdot b) \cdot \mu$

Normalisation

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

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

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

Ex:

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

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

I.e.

$$U_{A \subseteq X} := \lambda x. \begin{cases} x \in A : & \frac{1}{|A|} \\ x \notin A : & 0 \end{cases} \quad \text{so} \quad U_{\{x\} \subseteq X} = \delta_x$$

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.v.s:

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

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

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

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

$$: Prop$$

α,β independent w.r.t. λ

$\underline{\underline{Ex}}$ (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

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

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

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

$\underline{Ex}$ (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

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

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

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

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

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

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

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

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

__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

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

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

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

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

independence

Pairing r.e.s:

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

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

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

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

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:

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

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

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

$i \neq j$: $\lambda_{Same_{ij}} = U_{\mathbb{B}}$

$i \neq j$ ($\neq k$): $Same_{ij} \perp_{\lambda} Same_{jk}$

$$\not\perp_{\lambda} \{Same_{12}, Same_{23}, Same_{13}\}$$

Intuition: $Same_{13} = IFF(Same_{12}, Same_{23})$

Calc:

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

Vocabulary

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

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

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

$\mu : DX,\ f : X \to Y \vdash$ μ 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:

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

mutually inverse

Proof:

$$
\begin{tikzcd}
\mu \arrow[r, mapsto] \arrow[d, equal] & \mu(\langle\rangle) \arrow[dl] \\
\lambda\langle\rangle.\mu\langle\rangle &
\end{tikzcd}
\qquad A \qquad
\begin{tikzcd}
\lambda\langle\rangle.a \arrow[dr] & a \arrow[l, mapsto] \arrow[d, equal] \\
& (\lambda\langle\rangle.a)\langle\rangle
\end{tikzcd}
$$

$\blacksquare$

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:

$$\begin{array}{l} I \text{ ctbl} \\ \mu_- : (D\Gamma)^I, k : (DX)^\Gamma, \\ a_- : \mathbb{W}^I \end{array} \vdash \oint \Big(\sum_{i \in I} a_i \cdot \mu_i\Big) k = \sum_{i \in I} a_i \cdot \oint \mu_i \cdot k$$

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

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

Prop: $\mathbb{W} \cong D\mathbb{1}$ 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 module:

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

Corollary: $\int$ 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 $\alpha = \alpha^{+} - \alpha^{-}$

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

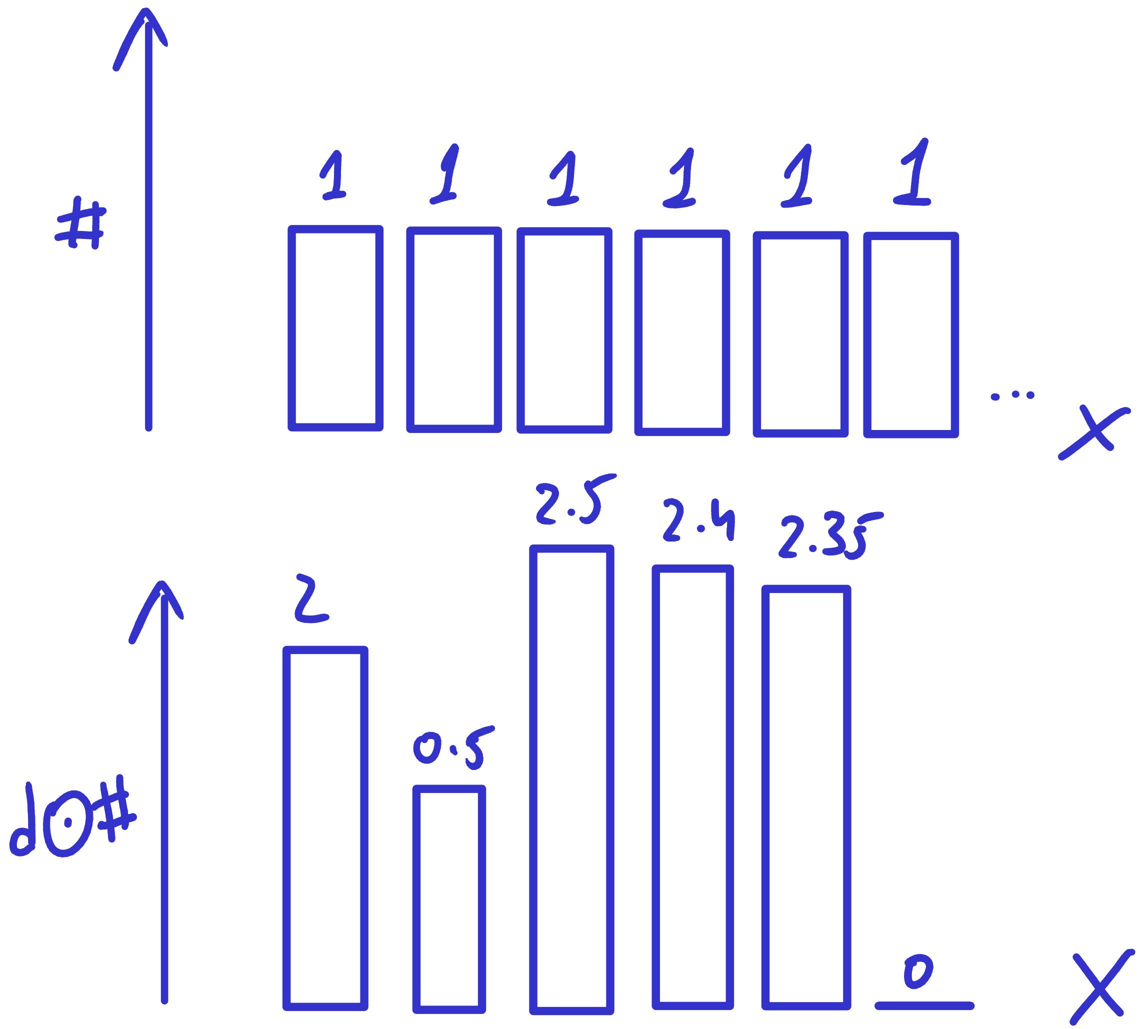
a density over X : $d : X \to W$

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

$$:= \oint \mu(dx)(d_x \cdot \delta_x)$$

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

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

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

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

When $\mu : PX$ we say instead

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

Absolute continuity

d is a density of μ w.r.t. ν or

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

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

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

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

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

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

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

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

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

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

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

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

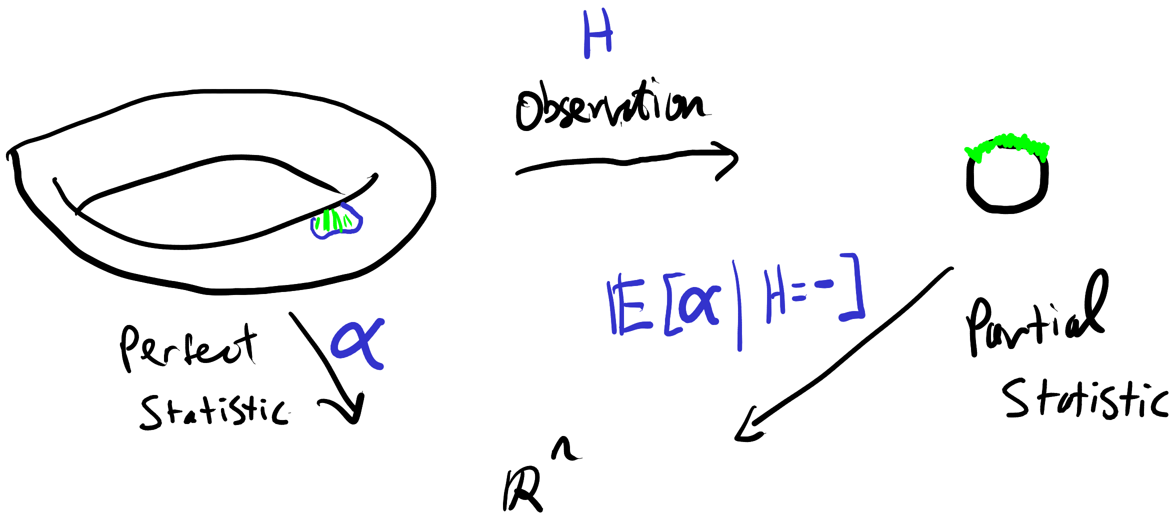
Conditional expectation

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

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

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

Conditional Probability (discrete version):

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

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

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

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

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

Obs 1: $d_X : W^X$ then $d_X = \dfrac{d\lambda_\alpha}{d\mu}$

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

& 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):

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

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 ✓

2) Borel sets & measurable spaces

3) Quasi Borel spaces

Lecture 1

4) Type structure & standard Borel spaces

5) Integration & random variables

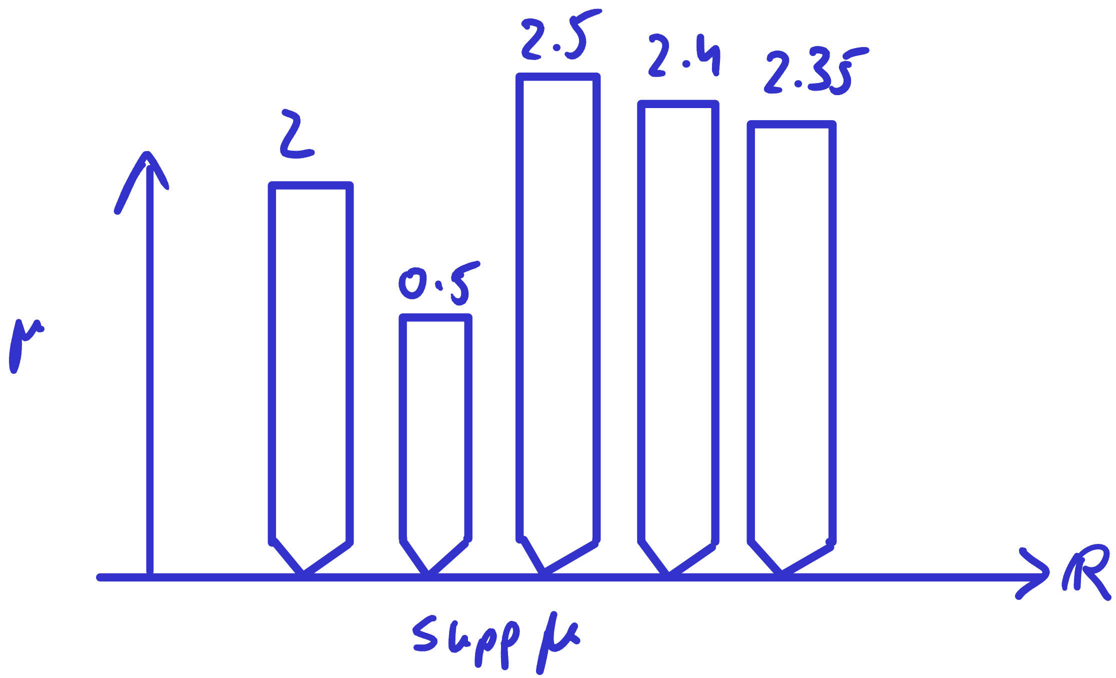
Lecture 2

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

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

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

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:

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

empty set

complements

countable unions

Idea: structure all spaces after the worst-case scenario

Examples

- Discrete spaces $\overset{\ulcorner meas \urcorner}{X} = (X, PX)$
- 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: $A \in P\lfloor V \rfloor$ $A := (A, [B_V] \cap A)$
- Products: $A \times B := (\lfloor A \rfloor \times \lfloor B \rfloor, \sigma([B_A] \times [B_B]))$

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:

$$f^{-1}[A] \in \mathcal{B}_{V_1} \Longleftarrow A \in \mathcal{B}_{V_2}$$

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 $f : \ulcorner X \urcorner \longrightarrow V$

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-algebras $B_{B_{\mathbb{R}}}$, $B_{\mathbb{R}^{\mathbb{R}}}$ for measurable

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

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

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

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

Sequential Higher-order Structure:

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

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

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

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

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

Compose higher-order building blocks: lim is measurable!

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

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

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

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 1^{st} order fragment $\Rightarrow$ non compositional

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 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"

$$
\begin{tikzcd}
 & \text{Standard Borel spaces} \arrow[dl] \arrow[dr] & \\
\mathbf{Meas} & & \mathbf{Qbs}
\end{tikzcd}
$$

abstract spaces (auxiliary?)

Plan:

1) type-driven probability: discrete case ✓
2) Borel sets & measurable spaces ✓
3) Quasi Borel spaces

Lecture 1

4) Type structure & standard Borel spaces
5) Integration & random variables

Lecture 2

Please ask questions!

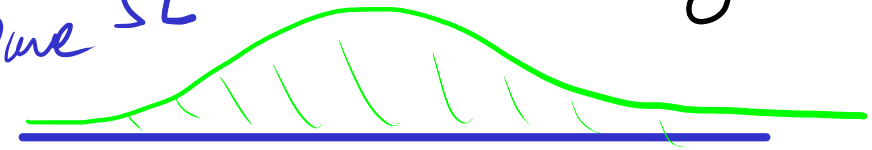
Course web page

scribble

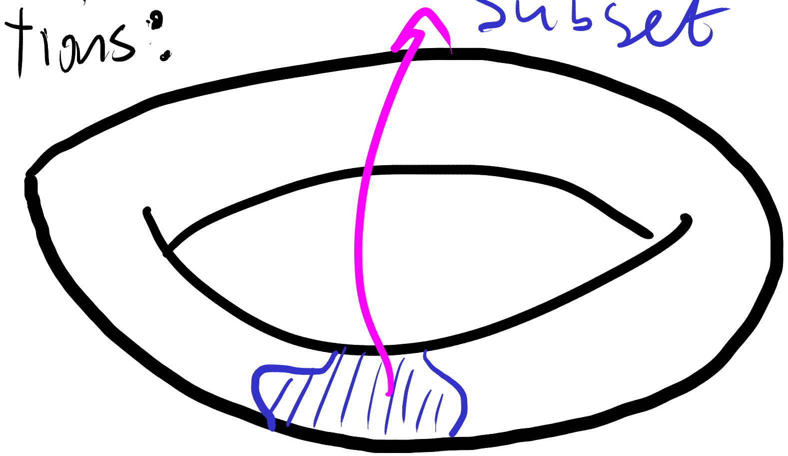
Core idea

Measure Theory	Qbs Theory

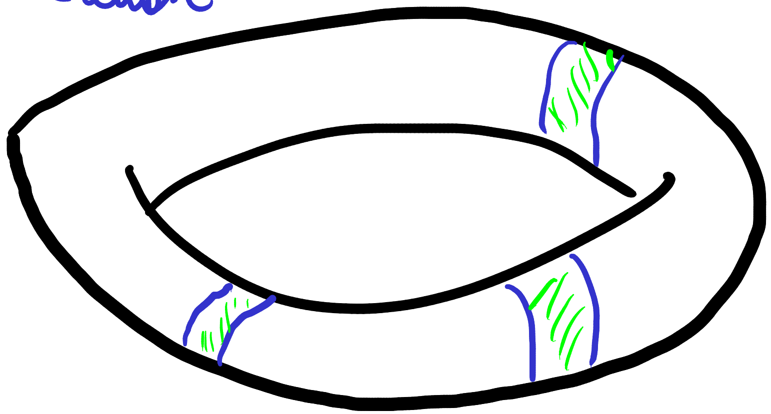
sample space Ω



Primitive notions: measurable subset



random element $\downarrow \alpha$



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

measure

Events $E \in B_X$

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

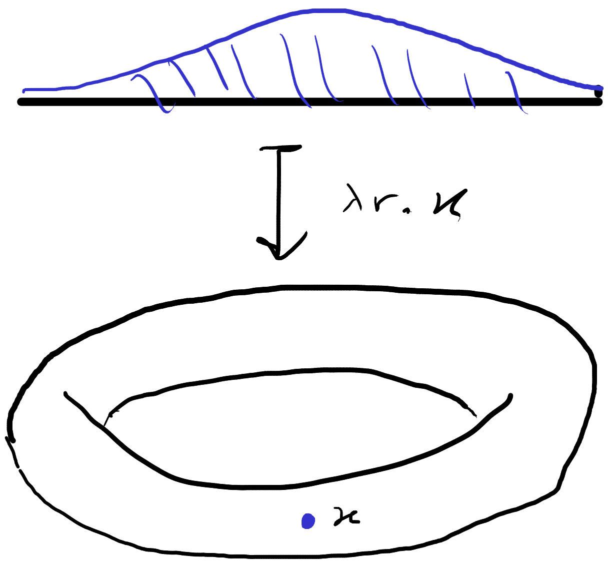
set "carrier"

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

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

- Constants:

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



- precomposition:

- recombination

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

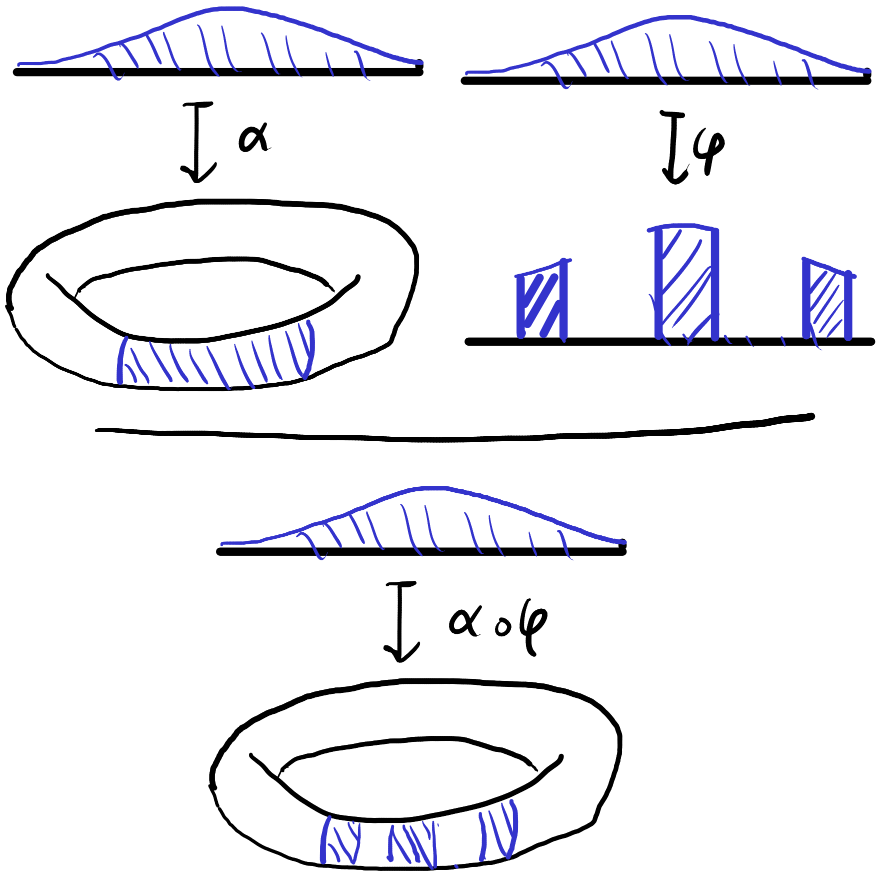
set "carrier"

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

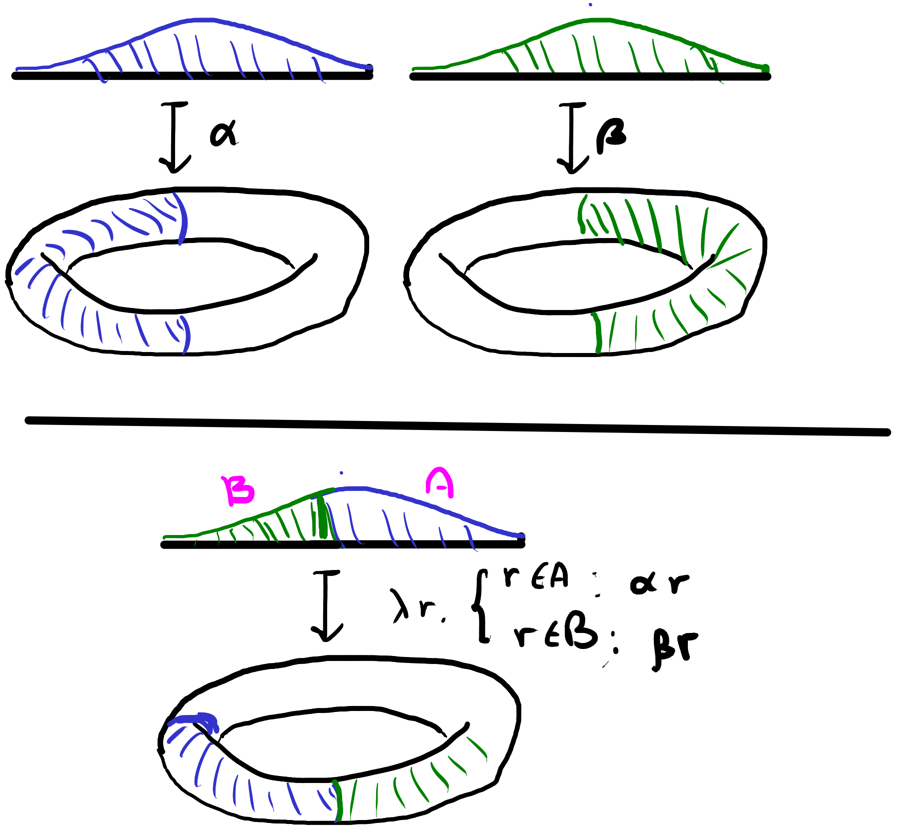
set "carrier"

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

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



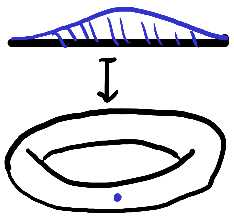
Def: Quasi-Borel space $X = (\lfloor X \rfloor, R_X)$

set "carrier"

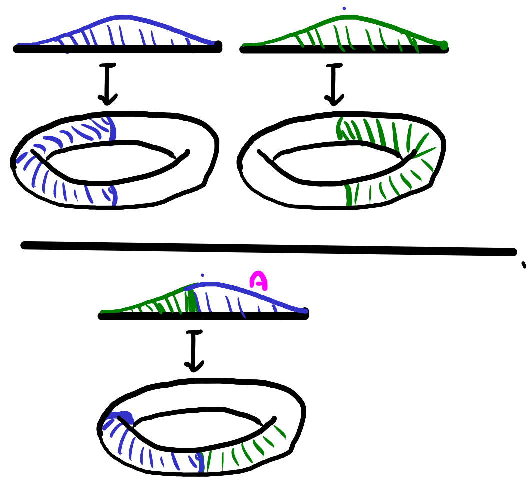
set of functions $\alpha : \mathbb{R} \to \lfloor X \rfloor$ "random elements"

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

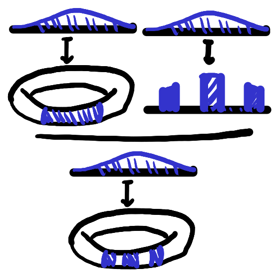
- constants:



- 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 funtions

  Indiscrete qbs on X

Qbs morphism $f : X \to Y$

- function $f : \lfloor X \rfloor \to \lfloor Y \rfloor$

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

Example

- Constant functions are qbs morphism

- σ-simple functions are qbs morphisms

Category Qbs $\Leftarrow$ - identity, composition

Full model

type : Qbs $\qquad W := [0,\infty] \qquad BX := \square$

$$DX := \square$$

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

$$Ce_{\mu}[E] := \square \qquad \delta_x := \square$$

$$\oint \mu k := \square$$

Plan:

1) type-driven probability: discrete case ✓
2) Borel sets & measurable spaces ✓
3) Quasi Borel spaces ✓

Lecture 1

4) Type structure & standard Borel spaces
5) Integration & random variables

Lecture 2

Please ask questions!

Course web page

Scribble