

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



supported by:



THE ROYAL
SOCIETY

The
Alan Turing
Institute

Facebook Research NCSC

Language of Distribution & Probability

Recap

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: $[0, \infty]$

$\mu: DX, E: BX \vdash c_\mu[E] : W$

→ type judgement

↳ measure μ assigns to E

Axioms for measures/distributions

Recap

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

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

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

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

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

kernels & their koch integral

Recap

Kernel from Γ to X : $k: (DX)^\Gamma$ or $k: \Gamma \rightarrow DX$

Dirac kernel: $\delta_-: X \rightarrow DX$

Koch integral: $\mu: D\Gamma, k: (DX)^\Gamma \vdash \oint \mu k : DX$
or $\oint \mu(dx) k(x)$ (dx binding occurs in $k(x)$)

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

Discrete model

$\text{type} : \text{set} \quad \mathbb{W} := [0, \infty] \quad \mathcal{B}X := \mathcal{P}X$

$\mathcal{D}X := \{ \mu : X \rightarrow \mathbb{W} \mid \text{supp } \mu \text{ countable} \}$

$\mathcal{P}X := \{ \mu \in \mathcal{D}X \mid \sum_{\mu} \mathcal{C}_{\mu}[X] = 1 \}$

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

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

Ex distributions

Counting measure (χ_{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!

smile



Course
web
page

Product measures

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

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

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

↑
discrete
model

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

build measures
compositionally

Indeed:

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

Notation: $\lambda: D(x \times Y), \kappa: (Dz)^{x \times Y} \vdash \iint \lambda(z_2, dy) \kappa(z, y)$
 $:= \oint \lambda \kappa$

Fubini - Tonelli Thm:

Integrate in any order:

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

$$\int \mu(dx) \int \nu(dy) u(x, y) = \iint (\mu \otimes \nu)(dx, dy) \\ = \int \nu(dy) \int \mu(dx) u(x, y)$$

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 $(+)$:

$(\#_{\underline{6}} \otimes \#_{\underline{6}})_{(+)} : \mathbb{D}\mathbb{N}$

is the number of
rolls whose sum is given

build measures
compositionally

Scaling a measure

$$(\cdot) : \mathbb{W} \times \mathcal{D}X \longrightarrow \mathcal{D}X$$

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

$$\text{NB: } \text{supp}(a \cdot \mu) = \begin{cases} a=0: \emptyset \\ a \neq 0: \text{supp } \mu \\ \quad \checkmark c+|b| \end{cases}$$

$(\cdot) : \mathbb{W} \times \mathcal{D}X \rightarrow \mathcal{D}X$ is an action of monoid $(\mathbb{W}, (\cdot), 1)$ on $\mathcal{D}X$:

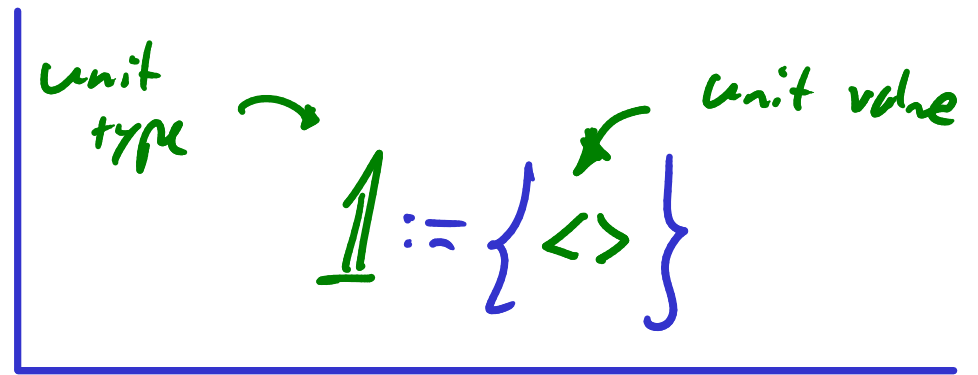
$$\mu : \mathcal{D}X \vdash$$

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

$$a, b : \mathbb{W}, \mu : \mathcal{D}X \vdash$$

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

Normalisation



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

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

Ex:

$$\emptyset \neq A \subseteq_{\text{fin}} X : \bigcup_{A \subseteq X} := \|(\#_A)\| : PX$$

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

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

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

$$\mu_{\pi_i} : D X_i$$

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

integrate out y

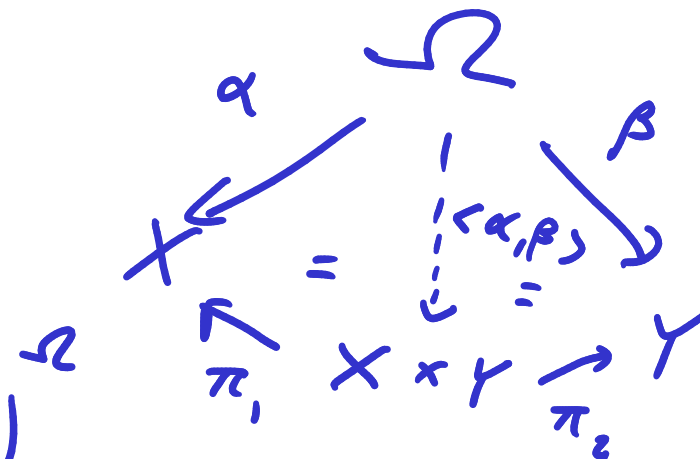
Exercise: $\mu : P X, \nu : D X \vdash (\mu \otimes \nu)_{\pi_2} = \nu$

independence

Pairing r.e.s:

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

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



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

: Prop

α, β independent w.r.t. λ

^(Dummett)
Ex represent outcomes of 3 coin tosses:

$$C := \{T, H\} \quad \Omega := C \times C \times C \quad \lambda: \bigcup_C \otimes \bigcup_C \otimes \bigcup_C : P_\Omega$$

$$\pi_i: \Omega \rightarrow C \quad \text{outcome of } i^{\text{th}} \text{ toss}$$

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

$$\text{where } (\stackrel{?}{=}) : C^2 \rightarrow B := \{ \text{True}, \text{False} \}$$

$$x \stackrel{?}{=} y := \begin{cases} x = y : \text{True} \\ x \neq y : \text{False} \end{cases}$$

^(Dummett)
Ex represent outcomes of 3 coin tosses:

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

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

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

marginalisation

$$\lambda_{\text{Same}_{12}}^T = (U_C \otimes U_C)^T_{(\stackrel{?}{=})} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$\begin{matrix} U_C(T) \cdot U_C(T) \\ \downarrow \\ \frac{1}{4} \\ \uparrow \\ U_C(H) \cdot U_C(H) \end{matrix}$

so $\lambda_{\text{Same}_{12}}^F = \frac{1}{2}$ too

^(Dummett)
Ex represent outcomes of 3 coin tosses:

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

$$\pi_i: \Omega \rightarrow C \quad \text{outcome of } i^{\text{th}} \text{ toss}$$

$$\begin{aligned} \underline{i \neq j}: \lambda_{\text{same}_{ij}} &= \bigcup_{\mathcal{B}} \\ \text{Same}_{ij}: \Omega &\xrightarrow{\langle \pi_i, \pi_j \rangle} C \times C \xrightarrow{(\cdot)} \mathcal{B} \end{aligned}$$

$$\begin{aligned} \lambda: \quad (T, T) &\mapsto \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} \\ \langle \text{same}_{12}, \text{same}_{23} \rangle &\quad \hookrightarrow \lambda(H, H, H) \quad \hookrightarrow \lambda(T, T, T) \end{aligned}$$

$$\begin{aligned} (T, F) &\mapsto \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} \\ &\quad \hookrightarrow \lambda(H, H, T) \quad \hookrightarrow \lambda(T, T, H) \end{aligned}$$

^(Dummett)
Ex represent outcomes of 3 coin tosses:

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

$$\pi_i: \Omega \rightarrow C \quad \text{outcome of } i^{\text{th}} \text{ toss}$$

$$\underline{i \neq j}: \lambda_{\text{same}_{ij}} = \bigcup_{\mathbb{B}} \quad \text{same}_{ij}: \Omega \xrightarrow{\langle \pi_i, \pi_j \rangle} C \times C \xrightarrow{(\cdot)} \mathbb{B}$$

$$\lambda_{\langle \text{same}_{12}, \text{same}_{23} \rangle} = \bigcup_{\mathbb{B} \times \mathbb{B}} = \bigcup_{\mathbb{B}} \otimes \bigcup_{\mathbb{B}} = \lambda_{\text{same}_{12}} \otimes \lambda_{\text{same}_{13}}$$

$$\text{So } \text{same}_{12} \perp_{\lambda} \text{same}_{13}$$

independence

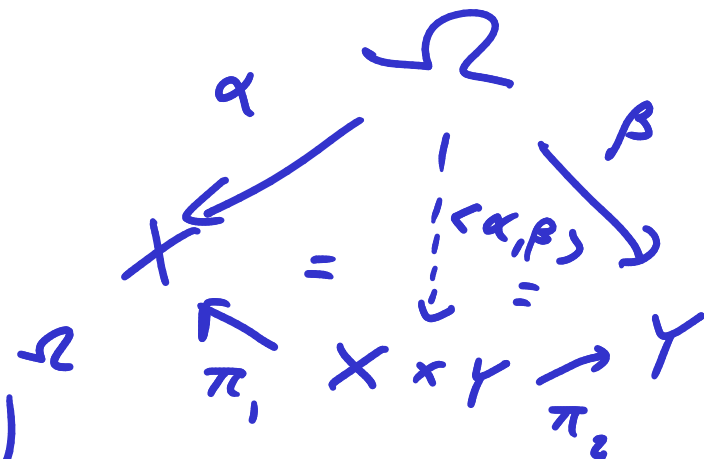
Pairing r.e.s:

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

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

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

α, β independent w.r.t. λ



I-any version:

$$\lambda : D\Omega, \alpha_i : \prod_{i \in I} X_i^\Omega \vdash \bigwedge_{i \in I} \alpha_i :=$$

α_i independent
w.r.t. λ

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

(Durrett)
Ex represent outcomes of 3 coin tosses:

$$C := \{T, H\} \quad \Omega := C \times C \times C \quad \lambda: \bigcup_C \otimes \bigcup_C \otimes \bigcup_C : P_\Omega$$

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

$$\text{Same}_{ij}: \Omega \xrightarrow{\langle \pi_i, \pi_j \rangle} C \times C \xrightarrow{(\cdot)} \mathbb{B}$$

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

$$\underset{\substack{* \\ \wedge \\ *}}{i \neq j}: \text{Same}_{ij} \perp \text{Same}_{jk} \quad \frac{1}{\lambda} \{ \text{Same}_{12}, \text{Same}_{23}, \text{Same}_{13} \}$$

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

Calc:

$$\lambda \left(\text{Same}_{12}, \text{Same}_{23}, \text{Same}_{13} \right) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} \neq \frac{1}{2^3} = \lambda_{\text{Same}_{12}} \otimes \lambda_{\text{Same}_{23}} \otimes \lambda_{\text{Same}_{13}}$$

$\hookrightarrow \lambda(H, H, H) \quad \hookrightarrow \lambda(T, T, T)$

Vocabulary

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

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

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

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

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

Ex:

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

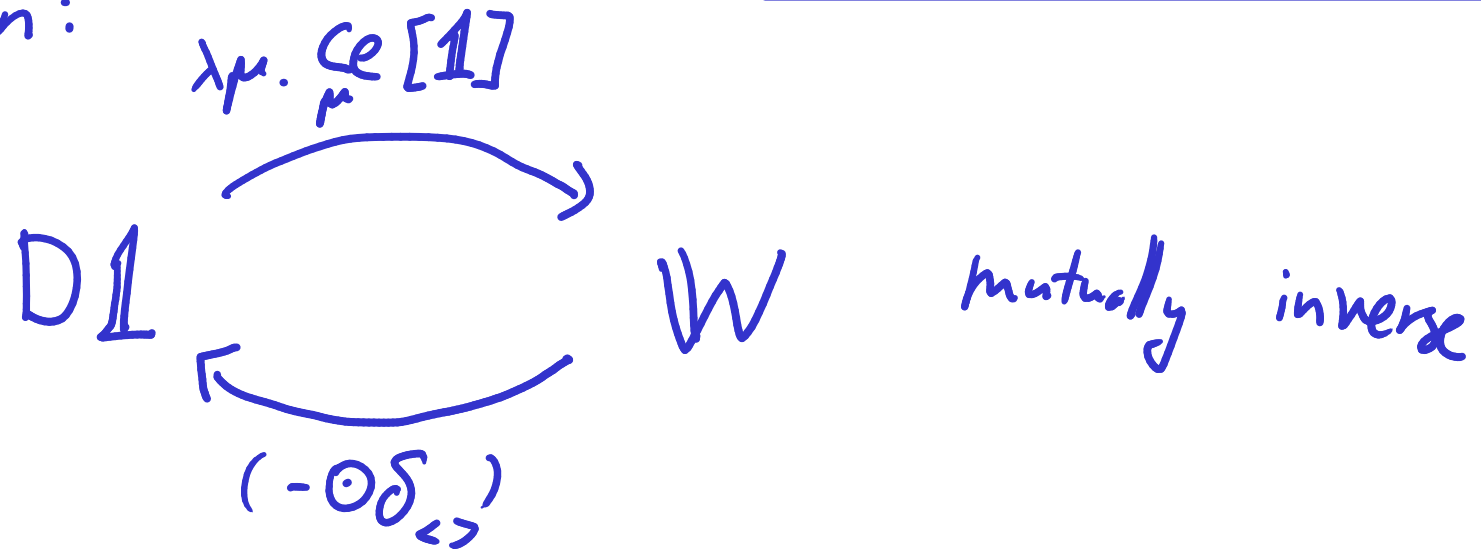
$$\text{Swap}: (X \times Y, \mu \otimes \nu) \longrightarrow (Y \times X, \nu \otimes \mu) \quad \text{so}$$

$$\mu: DX \vdash \mu \otimes \mu \text{ invariant under Swap}$$

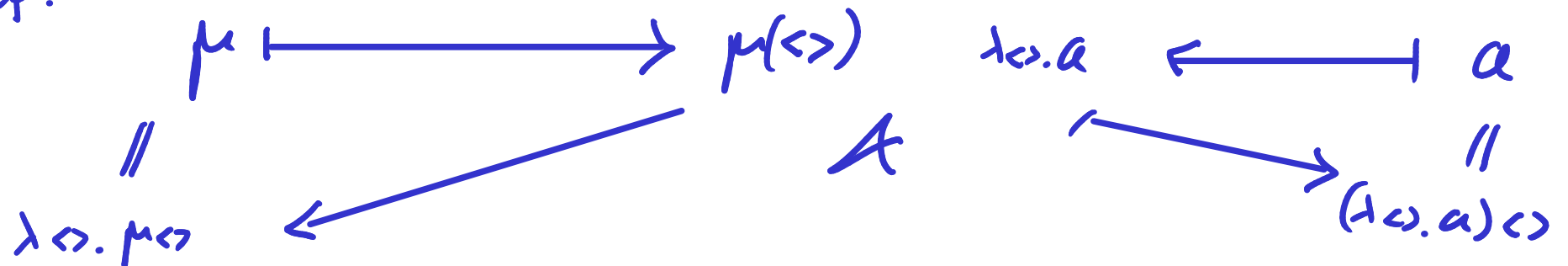
Weights as measures

NB: unit type \rightarrow $\underline{1} := \{<>\}$ \nwarrow unit value

Observation:



Proof:



Integration

$$\mu : DX, \varphi : W^X \vdash \int \mu \varphi : W$$
$$:= \sum_{x \in X} \mu x \cdot \varphi x$$

(Lebesgue
integral)

Can derive it:

$$\begin{array}{ccc} DX \times W^X & \xrightarrow{DX \times (\cong \circ -)} & DX \times (D\mathbb{1})^X \\ \int \downarrow & \cong & \downarrow \phi \\ W & \xleftarrow{\cong} & D\mathbb{1} \end{array}$$

Additivity:

$$\begin{aligned} I \text{ ctbl}, \mu : (DX)^I \vdash \sum_{i \in I} \mu_i : DX \\ := \lambda x. \sum_{i \in I} \mu_i x \end{aligned}$$

NB:

$$\begin{aligned} \text{supp} \sum_i \mu_i &\subseteq \\ &\cup_i \text{supp} \mu_i \\ &\checkmark \text{ctbl} \end{aligned}$$

Ex: Bernoulli distribution

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

$$\text{i.e. } B_p : \begin{aligned} \text{True} &\mapsto p \\ \text{False} &\mapsto 1-p \end{aligned}$$

Thm (affine-linearity):

ϕ is affine-linear in each argument:

$\mathcal{I} \text{ ctbl}$

$$\mu_- : (D\Gamma)^I, k_- : (Dx)^I, a_- : W^I \vdash \int (\sum_{i \in I} a_i \cdot \mu_i) k = \sum_{i \in I} a_i \cdot \int \mu_i k$$

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

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

Prop: $\mathbb{W} \cong D1$ is a σ -semi-ring isomorphism:

$$(\mathbb{W}, \Sigma, (\cdot), 1) \cong (D1, \Sigma', (\cdot), \delta_{\langle \rangle})$$

and $(\cdot): \mathbb{W} \times D\mathcal{X} \rightarrow D\mathcal{X}$ makes $D\mathcal{X}$ into a module:

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

Corollary: \int is affine-linear in each argument.

Random variable :

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

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

Can add, multiply r.v.'s.

To integrate r.v.'s:

$$(-)^{\pm}: \bar{\mathbb{R}}^\Omega \longrightarrow \mathbb{W}^\Omega$$

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

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

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

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

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

Ex. The (discrete) Lebesgue p -space:

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

$$\left\{ \alpha: \overline{\mathbb{R}}^{\Omega} \mid \int_{\mu} |\alpha|^p < \infty \right\}$$

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

$\mathcal{L}_2(\Omega, \mu)$ has an inner product $\langle \alpha, \beta \rangle := \int_{\mu} \alpha \cdot \beta$ almost Hilbert

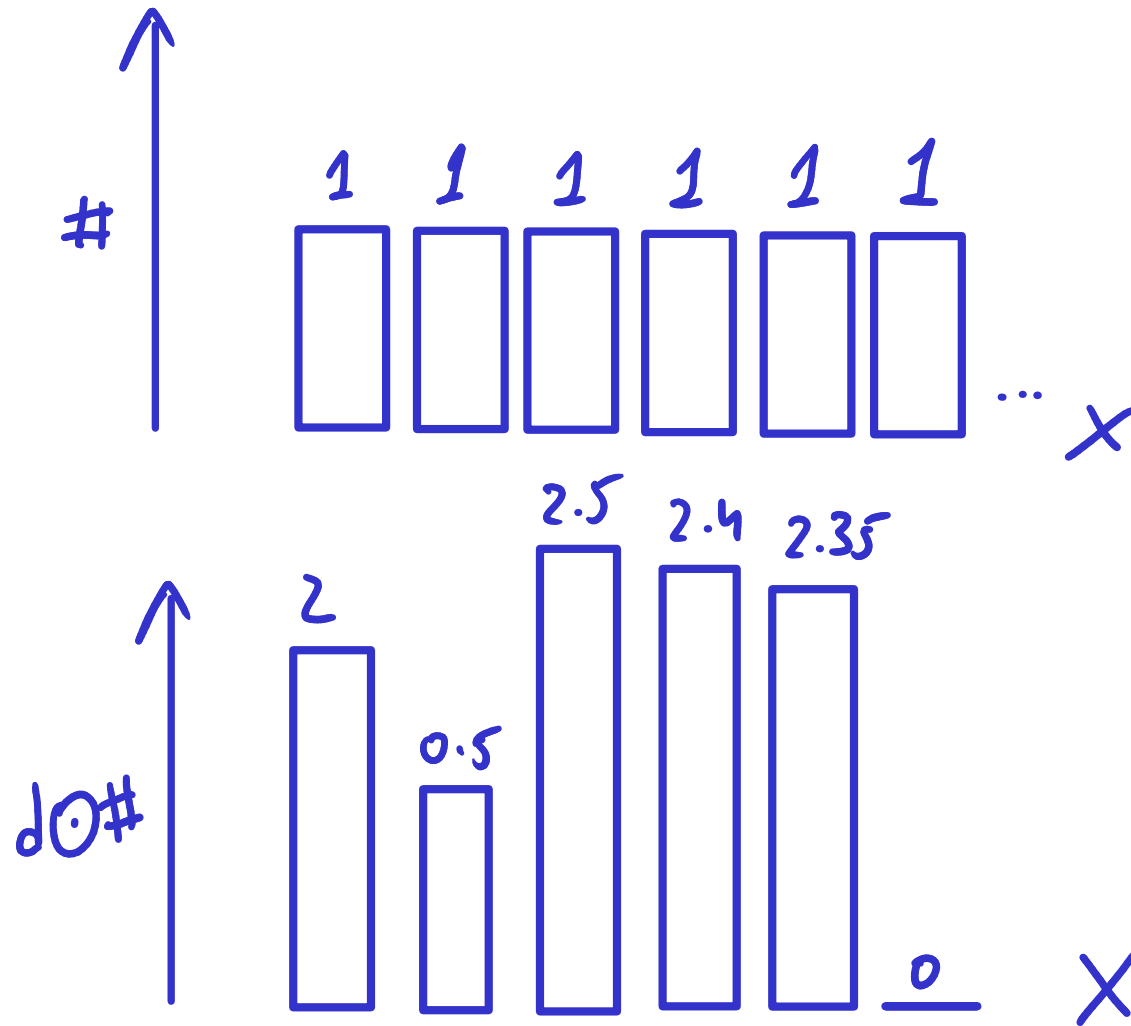
Density

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

$$d : W^X, \mu : DX \vdash d \odot \mu : DX \\ := \int \mu(dx) (dx \cdot \delta_x)$$

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

$$E: \mathcal{B}X, \mu: DX \vdash \text{\textcolor{green}{\mu(dx)-almost certainly } } x \in E : \text{Prop} \\ := [- \in E] \odot \mu = \mu$$

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

When $\mu: P X$ we say instead

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

Exercise Look up the def. of a normed space

and modify the definition so that $L_p(\Omega, \mu)$ is a normed space up to almost sure equality.

Absolute continuity

d is a density of μ w.r.t. V or

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

$$\mu, V : DX, d : W^X \vdash d = \frac{d\mu}{dV} : \text{Prop}$$

$$:= \mu = d \odot V$$

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

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

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

$$\text{Lemma: } \begin{array}{l} \mu, V : DX, \\ \mu \ll V, \\ h : (DY)^X \end{array} \quad \int V(dx) \frac{d\mu}{dV}(x) \cdot kx = \int \mu(dx) kx$$

$$\underline{Ex}: V_{A \subseteq X} \ll (\#_A)_{\text{Cust}: A \subseteq X}$$

$$\frac{dV_{A \subseteq X}}{d(\#_A)_{\text{Cust}}} = \lambda x. \left\{ \begin{array}{ll} x \in A: & \frac{1}{|A|} \\ \text{o.w.}: & 0 \end{array} \right.$$

but also:

$$\frac{dV_{A \subseteq X}}{d(\#_A)_{\text{Cust}}} = \lambda x. \frac{1}{|A|}$$

Radon-Nikodym Thm: (discrete version)

$\mu, \nu: \mathcal{P}X \vdash \mu \ll \nu$ iff $\forall x. \nu x = 0 \Rightarrow \mu x = 0$

i.e. $\text{Supp } \mu \subseteq \text{Supp } \nu$

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

$\forall (d_1) \text{-a.s. } d_1 x = d_2 x$

Ex: for ctbl X , $\forall \mu: \mathcal{D}X. \mu \ll \#_X$. Proof: vacuously, as $\#_X x \neq 0$.

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

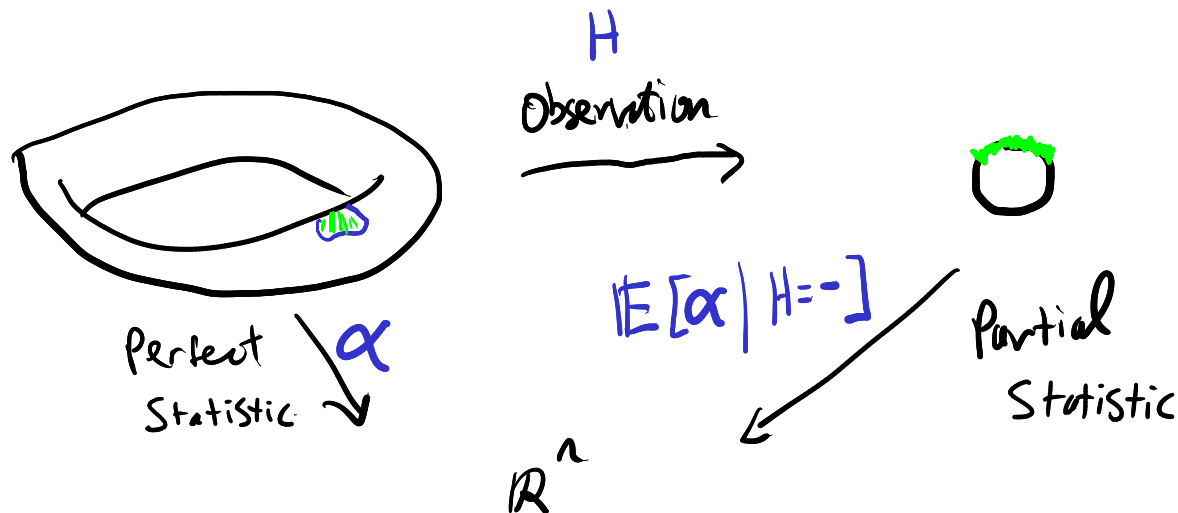
Conditional expectation

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

$$\mu: \Omega \rightarrow \mathbb{R}, H: X \rightarrow \mathbb{R}, \alpha: \mathcal{L}_1(\Omega, \mu), \beta: \mathcal{L}_1(X, \mu_H)$$

$$\vdash \beta = \mathbb{E}[\alpha | H = -] \quad : \text{Prop}$$

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



Thm (Kolmogorov): (discrete version)

There is a function

$$\underline{\mathbb{E}}[-|-:-] \in \prod_{\mu: P_{\Omega}} \prod_{H: X^{\Omega}} L_1(\Omega, \mu) \rightarrow L_1(X, \mu_H)$$

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

Conditional Probability (discrete version):

$$H: X^\Omega, \mu: P_X \vdash \prod_{\mu} [- | H = -] : (P_\Omega)^X$$

$$:= \lambda x_0: X. \lambda \omega_0: \Omega. \prod_{\omega \sim \mu} [(\omega_0 = \omega) | H\omega = x_0]$$

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

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

Assume $\mu: D_X$, $\nu: D_H$ s.t. $\lambda \ll \mu \otimes \nu$.

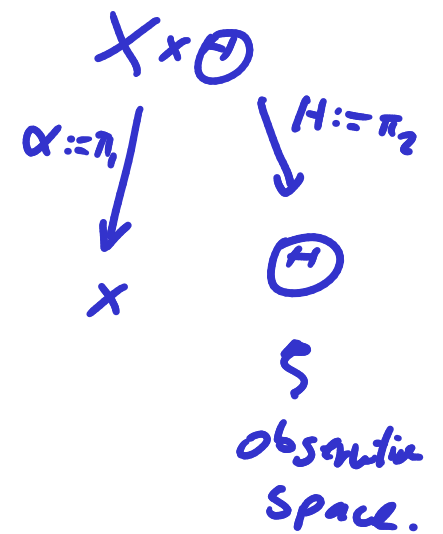
with $d_{X,H} = \frac{d\lambda}{d(\mu \otimes \nu)}$.

obs 1: $d_X: W^X$

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

$$\text{then } d_X = \frac{d\lambda_x}{d\mu}$$

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

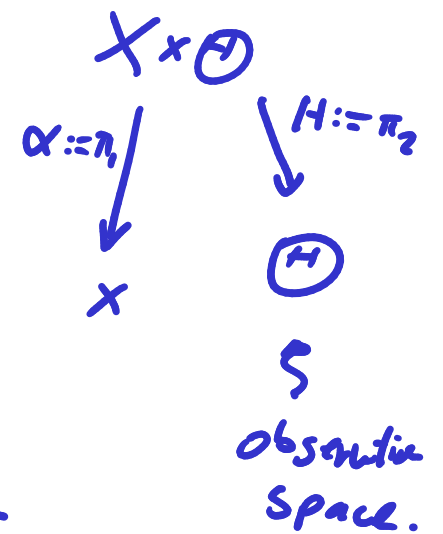


Bayes's Thm (discrete version, adapted from Williams):

Let $\lambda: P(X \times \mathcal{H})$ joint probability distribution.

Assume $\mu: D_X, \nu: D_{\mathcal{H}}$ s.t. $\lambda \ll \mu \otimes \nu$.

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



Let $d_{X/H}(-|-): X \times \mathcal{H} \rightarrow W$

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

$$\lambda_{X|H=-}: \mathcal{H} \rightarrow P_X$$

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

Bayes's formula:

$$P_{\lambda}[-|H=-] = \lambda_{X|H=-}$$

Summary

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

μ_H push-forward / law

$(Dx, \Sigma, (\cdot))$ module structure over a finite linearity of \mathcal{F}

\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 (^{Wed} 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!

snibble



Course
web
page