Three Toposes for Probability & Randomness

Alex Simpson

FMF, University of Ljubljana

IMFM, Ljubljana

Topos Institute Colloquium

6th June 2024

Topos 1

(Probability)

Random variables form a <u>sheaf</u>

<u>Presheaf property</u>

$$
\begin{tikzcd}
\Omega' \arrow[d, "p"'] \arrow[r, dashed, "X \cdot p"] & \mathbb{R} \\
\Omega \arrow[ur, "X"'] &
\end{tikzcd}
$$

<u>Sheaf property</u>

$$
\begin{tikzcd}
\Omega' \arrow[d, "p"'] \arrow[r, "Y"] & \mathbb{R} \\
\Omega \arrow[ur, dashed] &
\end{tikzcd}
$$

Y a.s. constant on p fibres

$\exists! X$ s.t. $Y = X \cdot p$

(w.r.t. <u>atomic Grothendieck topology</u>)

The topos of probability sheaves

Base categories of 'nice' sample spaces

$\mathbf{SBP}$ Standard Borel probability spaces + measure-preserving functions

$\mathbf{SBP}_0$ " " " " " " " mod 0

The topos

$$\boxed{\underline{\mathcal{P}} := Sh_{at}(\mathbf{SBP}_0) \equiv Sh_{at}(\mathbf{SBP})}$$

(atomic sheaves)

$\boxed{\text{The sheaf } \underline{RV}(A)}$

(of A-valued random variables)

For any <u>standard Borel space</u> A (with σ-algebra $\mathcal{B}_A$)

$$\begin{array}{ccc} \Omega & & \underline{RV}(A)(\Omega) := \text{measurable functions } \Omega \to A \text{ mod } 0 \\ \big\uparrow [p] & \mapsto & \big\downarrow \; [X] \mapsto [X \circ p] \\ \Omega' & & \underline{RV}(A)(\Omega') \end{array}$$

defines $\underline{RV}(A) : \mathrm{SBP}_0^{op} \to \underset{\sim}{\mathrm{Set}}$ that is a sheaf in $\underset{\sim}{\mathbb{P}}$.

The $\underline{RV}$ functor

The mapping $A \mapsto \underline{RV}(A)$ defines a functor

$$\boxed{\underline{RV} : \mathbf{SBS} \to \underline{P}}$$

Category of standard Borel spaces and measurable functions.

(The domain category can be expanded to universally measurable functions between universally measurable subsets of standard Borel spaces.)

Properties:

- faithful
- preserves countable limits.

The distribution functor

$$A \longmapsto \{\mu : \mathcal{B}_A \to [0,1] \mid \mu \text{ a probability measure}\}$$

defines a functor

$$\boxed{\underline{D} : \mathbb{SBS} \longrightarrow \underline{Set} \xrightarrow{\Delta} \underline{P}}$$

discrete (pre)sheaf

Properties:

- faithful
- taut

The law of a random variable

$$\mathbb{P}_A : \underline{RV}(A)(\Omega) \longrightarrow \underline{D}(A)(\Omega)$$

$$X \longmapsto \left(B \in \mathcal{B}_A \mapsto \mathbb{P}[X \in B] \right)$$

the <u>law</u> of X

defines a natural transformation

$$\begin{tikzcd} \mathcal{SBS} \arrow[r, bend left, "\underline{RV}"] \arrow[r, bend right, "\underline{D}"'] \arrow[r, phantom, "\mathbb{P} \Downarrow"] & \underline{P} \end{tikzcd}$$

- $\mathbb{P}$ is taut

Internal definitions of probabilistic concepts

Using the internal logic of $\underline{\mathcal{P}}$, which is classical because atomic toposes are boolean.

For $X, Y : \underline{RV}(A)$

$$X \sim Y \quad :\Leftrightarrow \quad \mathbb{P}_X = \mathbb{P}_Y$$

abbreviation for $\mathbb{P}_A(X)$

$$X =_{a.s.} Y \quad :\Leftrightarrow \quad \mathbb{P}_{(X,Y)}(\{(x,x) \mid x \in A\}) = 1$$

using product preservation
$\underline{RV}(A) \times \underline{RV}(A) \cong \underline{RV}(A \times A)$ to
consider (X,Y) as element of $\underline{RV}(A \times A)$

For $X : \underline{RV}(A)$, $Y : \underline{RV}(B)$

$$X \perp\!\!\!\perp Y \quad :\Leftrightarrow \quad \forall S \in \Delta\mathcal{B}_A,\ \forall T \in \Delta\mathcal{B}_B .\ \mathbb{P}_{(X,Y)}(S \times T) = \mathbb{P}_X(S) \cdot \mathbb{P}_Y(T).$$

Internal definitions of probabilistic concepts

Using the internal logic of $\underline{\mathcal{P}}$, which is classical because atomic toposes are boolean.

For $X, Y : \underline{RV}(A)$

abbreviation for $\mathbb{P}_A(X)$

$$X \sim Y \quad :\Leftrightarrow \quad \mathbb{P}_X = \mathbb{P}_Y$$

Proposition $X =_{a.s.} Y \Leftrightarrow X = Y$!

(follows from tautness of $\mathbb{P}$)

$$X =_{a.s.} Y \quad :\Leftrightarrow \quad \mathbb{P}_{(X,Y)}(\{(x,x) \mid x \in A\}) = 1$$

using product preservation

$\underline{RV}(A) \times \underline{RV}(A) \cong \underline{RV}(A \times A)$ to

consider (X,Y) as element of $\underline{RV}(A \times A)$

For $X : \underline{RV}(A)$, $Y : \underline{RV}(B)$

$$X \perp\!\!\!\perp Y \quad :\Leftrightarrow \quad \forall S \in \Delta\mathcal{B}_A, \forall T \in \Delta\mathcal{B}_B \,.\, \mathbb{P}_{(X,Y)}(S \times T) = \mathbb{P}_X(S) \cdot \mathbb{P}_Y(T).$$

Two logical laws

Invariance principle

For any subsheaf $\Phi \rightarrowtail \underline{RV}(A)$,

$$\forall X, Y : \underline{RV}(A) \quad X \sim Y \wedge \Phi(X) \to \Phi(Y)$$

Independence principle

$$\forall X : \underline{RV}(A), Y : \underline{RV}(B) \;\; \exists Z : \underline{RV}(A) \quad Z \sim X \wedge Z \perp\!\!\!\perp Y$$

Dependent choice

A topos with countable limits enjoys the principle of (internal countable) <u>dependent choice</u> (DC) if

every ω^{op}-diagram of epimorphisms

$$\cdots \overset{e_4}{\twoheadrightarrow} X_4 \overset{e_3}{\twoheadrightarrow} X_3 \overset{e_2}{\twoheadrightarrow} X_2 \overset{e_1}{\twoheadrightarrow} X_1 \overset{e_0}{\twoheadrightarrow} X_0$$

has a (w.l.o.g. limit) cone of epimorphisms

Lemma A sufficient condition for an atomic topos $\mathrm{Sh}_{at}(\mathbb{C})$ to satisfy DC is that every ω^{op}-diagram in $\mathbb{C}$ has a cone.

Proposition The topos $\underline{P}$ of probability sheaves satisfies DC.

Proof outline Consider any ω^{op}-diagram in $\mathbb{SBP}_0$

$$\cdots \xrightarrow{[p_4]} \Omega_4 \xrightarrow{[p_3]} \Omega_3 \xrightarrow{[p_2]} \Omega_2 \xrightarrow{[p_1]} \Omega_1 \xrightarrow{[p_0]} \Omega_0$$

Define $\Omega_\omega := \{(\omega_n)_{n\geq 0} \mid \forall i.\ p_i(\omega_{i+1}) = \omega_i\}$ (limit in $\mathbb{SBS}$)

By Daniell-Kolmogorov extension Ω_ω carries a unique probability measure that projects to each Ω_n, giving the required cone (limit in $\mathbb{SBP}$). $\square$

iid sequences

<u>Proposition</u> For any $X : \underline{RV}(A)$ there exists $S : (\underline{RV}(A))^{\mathbb{N}}$ s.t. $\forall n : \mathbb{N}.\ S_n \sim X$ and S is a sequence of independent random variables.

<u>Proof</u> Define $S_0 := X$. Suppose we have $S_0, \ldots, S_{n-1}$, where $n \geq 1$.

By the independence principle, there exists $S_n \in \underline{RV}(A)$ s.t. $S_n \sim X$ and $S_n \perp\!\!\!\perp (S_0, \ldots, S_{n-1})$.

By DC, the S_n above can be found by a function $S : \mathbb{N} \to \underline{RV}(A)$. $\square$

By countable product preservation $(\underline{RV}(A))^{\mathbb{N}} \cong \underline{RV}(A^{\mathbb{N}})$

So $S : \underline{RV}(A)^{\mathbb{N}}$ gives us $S' \cdot \underline{RV}(A^{\mathbb{N}})$ allowing us to express many properties, e.g.,

if $T \in \mathcal{B}_{A^{\mathbb{N}}}$ is <u>tail</u> then $\mathbb{P}_{S'}(T) = 0$ or $\mathbb{P}_{S'}(T) = 1$ (0-1 law)

We have a good setting for <u>discrete-time</u> stochastic processes.

<u>Continuous time</u> is more problematic. E.g.,

- What do we mean by $\underline{RV}(\mathbb{R}^{[0,\infty)})$?

  We need to choose an SBS for $\mathbb{R}^{[0,\infty)}$ requiring pre-commitment to continuous processes (or similar) and a particular choice of σ-algebra, all of which encumbers the natural probabilistic development.

Topos 2

(Randomness)

Random elements

Probability concerns random variables. The value of an A-valued random variable X varies according to the probability law $\mathbb{P}_X$.

Randomness concerns random elements. A random element is a single fixed value of A obtained by sampling an A-valued random variable.

Random sequences

A <u>random sequence</u> is a random element in $2^{\mathbb{N}}$, obtained via an infinite sequence of fair coin tosses (i.e., by sampling the uniform probability distribution λ on $2^{\mathbb{N}}$).

$(01)^{\omega} := 010101010101010101\ldots$ is <u>not</u> random

$011010100010100010\ldots$ <u>is</u> random!

$s \in 2^{\mathbb{N}}$ is <u>naively random</u> if, for every measurable $T \subseteq 2^{\mathbb{N}}$,

$$\lambda(T) = 1 \implies s \in T.$$

<u>Problem</u> No $s \in 2^{\mathbb{N}}$ is naively random.

<u>Proof</u> Take $T := 2^{\mathbb{N}} \setminus \{s\}$.

This problem is circumvented in approaches to randomness by, e.g.,:

- restricting to T satisfying computability restrictions (<u>algorithmic randomness</u>)
- restricting to T definable in a given countable model (<u>set theory</u>)

Randomness-preserving functions

Even if inconsistent in itself, the notion of naive randomness suggests a sensible notion of randomness-preserving function.

A measurable function $f : \Omega' \to \Omega$ between standard Borel probability spaces is <u>randomness preserving</u> if

$$\forall T \subseteq \mathcal{B}_\Omega . \quad \mathbb{P}_\Omega(T) = 1 \Rightarrow \mathbb{P}_{\Omega'}(f^{-1}T) = 1$$

or $\equiv$ly

$$\forall T \subseteq \mathcal{B}_\Omega \quad \mathbb{P}_\Omega(T) = 0 \Rightarrow \mathbb{P}_{\Omega'}(f^{-1}T) = 0$$

Base categories

$\mathbf{SBR}$ Standard Borel probability spaces + randomness-preserving functions

$\mathbf{SBR}_0$ " " " . " " " " mod 0

Covers

Let $f : \Omega' \longrightarrow \Omega$ be randomness preserving

$T \in \mathcal{B}_\Omega$ is a subimage of f if, for all $S \in \mathcal{B}_\Omega$, $S \subseteq T$ and $\mathbb{P}_\Omega(S) > 0 \Rightarrow \mathbb{P}_{\Omega'}(f^{-1}S) > 0$.

An image of f is a subimage of maximum measure (amongst subimages)

A countable family $(f_n : \Omega_n \longrightarrow \Omega)$ is covering if

$$\mathbb{P}_\Omega\left(\bigcup_n T_n\right) = 1$$ where each T_n is an image of f_n.

The <u>random topos</u>

Countable covering families form a Grothendieck topology both on $\mathbb{SBR}$ and on $\mathbb{SBR}_0$: the <u>countable cover topology</u>.

The <u>random topos</u> is defined by:

$$\boxed{\underline{\mathcal{R}} := Sh_{cc}(\mathbb{SBR}_0) \simeq Sh_{cc}(\mathbb{SBR})}$$

(sheaves for the countable cover topology)

- On $\mathbb{SBR}_0$, countable cover = canonical = dense.
- $\underline{R}$ is therefore a boolean topos (its internal logic is classical)
- DC holds
- The subobject classifier $\underset{\sim}{2}$

$$\underline{2}(\Omega) := B_\Omega \bmod 0 \qquad \text{(measure algebra)}$$

- The real numbers $\underline{\mathbb{R}}$

$$\underline{\mathbb{R}}(\Omega) := \text{measurable functions } \Omega \to \mathbb{R} \bmod 0$$

- Sequences $\underline{2}^{\mathbb{N}}$

$$\underline{2}^{\mathbb{N}}(\Omega) := \text{measurable functions } \Omega \to 2^{\mathbb{N}} \text{ mod } 0$$

- Random sequences $\underline{\mathrm{Ran}} \subseteq \underline{2}^{\mathbb{N}}$

$$\underline{\mathrm{Ran}}(\Omega) := \text{randomness-preserving functions } \Omega \rightsquigarrow (2^{\mathbb{N}}, \lambda) \text{ mod } 0$$
$$\cong \$\mathbb{B}\mathbb{R}_0(\Omega, (2^{\mathbb{N}}, \lambda)) \qquad \text{representable!}$$

__Theorem__ Internally in $\underline{\mathcal{R}}$

$$\underline{\mathrm{Ran}} = \{ s : \underline{2}^{\mathbb{N}} \mid \forall T \subseteq \underline{2}^{\mathbb{N}}.\ \lambda(T) = 1 \text{ and } T \perp\!\!\!\perp s \Rightarrow s \in T \}$$

- Sequences $\underline{2}^{\mathbb{N}}$

$$\underline{2}^{\mathbb{N}}(\Omega) := \text{measurable functions } \Omega \to 2^{\mathbb{N}} \text{ mod } 0$$

- Random sequences $\underline{\mathrm{Ran}} \subseteq \underline{2}^{\mathbb{N}}$

$$\underline{\mathrm{Ran}}(\Omega) := \text{randomness-preserving functions } \Omega \rightsquigarrow (2^{\mathbb{N}}, \lambda) \text{ mod } 0$$
$$\cong \mathrm{\$BR}_0(\Omega, (2^{\mathbb{N}}, \lambda))$$

representable!

<u>Theorem</u> Internally in $\underline{\mathcal{R}}$

$$\underline{\mathrm{Ran}} = \{ s : \underline{2}^{\mathbb{N}} \mid \forall T \subseteq \underline{2}^{\mathbb{N}}.\ \lambda(T) = 1 \text{ and } T \perp\!\!\!\perp s \Rightarrow s \in T \}$$

a primitive relation of <u>independence</u> in $\underline{\mathcal{R}}$ related to Day convolution.

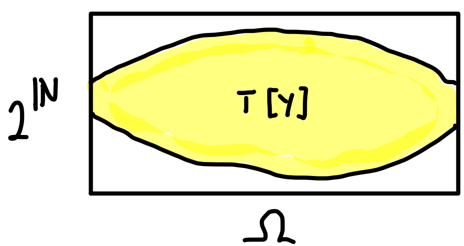
Theorem Internally in $\underline{\mathcal{R}}$, there exists a translation-invariant probability measure $\underline{\lambda} : \mathcal{P}(\underline{2}^{\mathbb{N}}) \to \underline{[0,1]}$ extending the uniform Borel measure.

Outline construction of $\underline{\lambda}$

Suppose $T \in \mathcal{P}(\underline{2}^{\mathbb{N}})(\Omega) \cong \left(\underline{2}^{2^{\mathbb{N}}}\right)(\Omega)$. Reindex T along $\Omega \otimes (2^{\mathbb{N}}, \lambda) \xrightarrow{\pi_1} \Omega$.

Consider $[Y] \in (\underline{2}^{\mathbb{N}})(\Omega \otimes (2^{\mathbb{N}}, \lambda))$ where $Y := \Omega \otimes (2^{\mathbb{N}}, \lambda) \xrightarrow{\pi_2} 2^{\mathbb{N}}$. (Generic random sequence)

Then $T[Y] \in \underline{2}(\Omega \otimes (2^{\mathbb{N}}, \lambda)) \cong \mathcal{B}_{\Omega \times 2^{\mathbb{N}}}$ mod 0



$$\underline{\lambda}(T) := \left[\omega \mapsto \lambda\{s \in 2^{\omega} \mid (\omega, s) \in T[Y]\}\right]$$

$$: \Omega \to [0,1] \text{ mod } 0 \quad = \underline{[0,1]}(\Omega)$$

More generally <u>all Borel probability measures</u> extend to canonical powerset measures. We state the precise theorem without explaining the underlined concepts, which rely on a general theory of randomness in $\mathcal{R}$.

<u>Theorem</u> Internally in $\mathcal{R}$, for every standard Borel probability space $(A, \mathcal{B}, \mu) : \mathcal{B} \to [0,1]$, there exists a unique probability measure $\mu^* : \mathcal{P}(A) \to [0,1]$ satisfying:

- μ^* extends μ,
- μ^* is <u>near Borel</u>,
- there are <u>enough μ^*-random elements</u>, and
- μ^*-random elements are <u>Borel testable</u>.

Topos 3

(Probability again)

The topos of <u>random probability sheaves</u>

<u>The topos</u>

$$\boxed{\underline{P}_{\underline{R}} \quad := \quad \text{the topos of probability sheaves } \underline{P} \text{ relative to the random topos } \underline{R}}$$

One can externalise this to a Grothendieck topos over <u>Set</u>, but we shall study $\underline{P}_{\underline{R}}$ from the <u>internal perspective</u> of $\underline{R}$.

The $\underline{RV}$ endofunctor

$$\underline{RV} : \underline{P}_{\underline{R}} \to \underline{P}_{\underline{R}}$$

$$\begin{tikzcd}
\Omega & & \underline{RV}(\underline{A})(\Omega) \arrow[d] & := \text{ arbitrary functions } \Omega \to \underline{A}(\Omega) \text{ mod } 0 \\
\Omega' \arrow[u, "{[p]}"] & \longmapsto & \underline{RV}(\underline{A})(\Omega') & {[f]} \mapsto [\omega' \mapsto f(p(\omega')) \cdot [p]]
\end{tikzcd}$$

$f \equiv f' \bmod 0 \;:\Longleftrightarrow\; \mathbb{P}^*_\Omega(\{\omega \mid f(\omega) \neq f'(\omega)\}) = 0$

is well-defined because $\mathbb{P}^*_\Omega : \mathcal{P}(\Omega) \to [0,1]$

The law of a random variable

The distributions endofunctor $\underline{D} : \underline{P}_{\underline{R}} \to \underline{P}_{\underline{R}}$ (in fact monad)

$$\underline{D}(\underline{A}) := \{ \mu : \mathcal{P}(\underline{A}) \to [0,1] \mid \mu \text{ a probability measure} \}$$

(powerset probability measures, internally defined in $\underline{P}_{\underline{R}}$)

Probability law natural transformation $\mathbb{P} : \underline{RV} \Rightarrow \underline{D}$

$$\mathbb{P}_{\underline{A}} : \underline{RV}(\underline{A})(\Omega) \longrightarrow \underline{D}(\underline{A})(\Omega)$$

$$[X : \Omega \to \underline{A}(\Omega)] \longmapsto (B \in \mathcal{P}(\underline{A})(\Omega)) \longmapsto \mathbb{P}^*_\Omega(X^{-1}B_\Omega)$$

Properties of the set-up in $\underline{P}_{\underline{R}}$

- $\underline{RV}$ is faithful and preserves countable limits
- $\underline{D}$ is faithful and taut
- $\mathbb{P} \cdot \underline{RV} \Rightarrow \underline{D}$ is taut
- The independence principle & invariance principle
- DC

This supports the development of an axiomatic <u>synthetic probability theory</u> based on the $\underline{RV}$ functor.

Random variables can be valued in arbitrary sets and their probability laws are powerset measures.

We now have a very natural route to continuous-time stochastic processes; e.g.,

$\underline{RV}(\mathbb{R})^{[0,\infty)}$ — processes up to <u>modification equivalence</u>

$\underline{RV}(\mathbb{R}^{[0,\infty)})$ — processes up to <u>indistinguishability</u>

Given $X : \underline{RV}(\mathbb{R}^{[0,\infty)})$,

$$\exists Y : \underline{RV}(\mathbb{R}^{[0,\infty)}) .\ \mathbb{P}_Y(C[0,\infty)) = 1 \wedge \forall t \in [0,\infty)\ X_t = Y_t$$

says that X has a continuous modification.

A fourth topos!

- A nominal approach to probabilistic separation logic, Li, Ahmed, Aytac, Holtzen, Johnson-Freyd, LICS 2024.

Considers the atomic topos on the subcategory of $\mathbb{SBR}_0$ consisting of maps that are singleton covers. This is proven equivalent to a category of continuous group actions, and several probabilistically-interesting constructions are related across this equivalence.

(cf. Schanuel topos $\simeq$ nominal sets)

Paper + on-line talks

- Equivalence and conditional independence in atomic sheaf logic, LICS 2024.

  Study of the independence and invariance principles in atomic sheaf categories.

- Probability sheaves, Topos à l'IHES, 2015.

  An early talk on probability sheaves.

- A mathematical theory of true randomness, Parts 1 & 2. UNISA seminar, 2023.

  An axiomatic theory of randomness & measure, modelled by the random topos.

- Synthetic probability theory. Talk at Categorical Probability & Statistics, 2020

  An axiomatic development of the synthetic probability theory modelled by $\underline{P}_{\underline{R}}$.