Synthetic Probability Theory

Alex Simpson

Faculty of Mathematics and Physics
University of Ljubljana, Slovenia

PSSL, Amsterdam,
October 2018

Gian-Carlo Rota (1932-1999):

> *"The beginning definitions in any field of mathematics are always misleading, and the basic definitions of probability are perhaps the most misleading of all."*

Twelve Problems in Probability Theory No One Likes to Bring Up, The Fubini Lectures, 1998 (published 2001)

Random variables

An A-valued random variable is:

$$X : \Omega \to A$$

where:

- the value space A is a measurable space (set with σ-algebra of measurable subsets);
- the sample space Ω is a probability space (measurable space with probability measure $\mathbf{P}_\Omega$); and
- X is a measurable function.

Features of random variables

σ-algebras:

- used to avoid non-measurable sets;
- used to implement dependencies, conditioning, etc.

The sample space Ω :

- the carrier of all probability.

Neither σ-algebras nor sample spaces are inherent in the concept of probability.

Technical manipulations involving σ-algebras and sample spaces seem a cumbersome distraction when carrying out probabilistic arguments.

Features of random variables

σ-algebras:

- used to avoid non-measurable sets;
- used to implement dependencies, conditioning, etc.

The sample space Ω :

- the carrier of all probability.

Neither σ-algebras nor sample spaces are inherent in the concept of probability.

Technical manipulations involving σ-algebras and sample spaces seem a cumbersome distraction when carrying out probabilistic arguments.

Goal: Reformulate probability theory with "random variable" as a primitive notion, governed by axioms formulated in terms of concepts with direct probabilistic meaning.

David Mumford:

> *"The basic object of study in probability is the random variable and I will argue that it should be treated as a basic construct, like spaces, groups and functions, and it is artificial and unnatural to define it in terms of measure theory."*

The Dawning of the Age of Stochasticity, 2000

Guiding model

Let $(\mathbb{K}, J_\omega)$ be the site with:

- objects — standard Borel probability spaces;
- morphisms — nullset reflecting measurable functions;
- Grothendieck topology — the countable cover topology J_ω.

Then $\mathcal{R}\mathrm{an} = \mathrm{Sh}(\mathbb{K}, J_\omega)$ is a Boolean topos validating DC in which $\mathbb{R}^n$ carries a translation-invariant measure on all subsets.

Internally in $\mathcal{R}\mathrm{an}$, let $(\mathbb{P}, J_{\mathrm{at}})$ be the site with:

- objects — standard Borel probability spaces;
- morphisms — measure preserving functions;
- Grothendieck topology — the atomic topology J_{at}.

Then $\mathrm{Sh}_{\mathcal{R}\mathrm{an}}(\mathbb{P}, J_{\mathrm{at}})$ is again a Boolean topos validating DC in which $\mathbb{R}^n$ carries a translation-invariant measure on all subsets. It also models primitive random variables.

Axiomatic setting

We assume a Boolean elementary topos $\mathcal{E}$ validating DC.

For any object A there is an object $\mathcal{M}_1(A)$ of all probability measures defined on all subsets.

$$\mathcal{M}_1(A) \;=\; \{\mu \colon \mathcal{P}(A) \to [0,1] \mid \mu \text{ is a probability measure}\} \;.$$

The operation $A \mapsto \mathcal{M}_1(A)$ is functorial with $f \colon A \to B$ lifting to the function that sends any probability measure $\mu \colon \mathcal{M}_1(A)$ to its pushforward $f[\mu] \colon \mathcal{M}_1(B)$:

$$f[\mu](B') \;=\; \mu(f^{-1}B') \;.$$

This functor further carries a (commutative) monad structure.

Primitive random variables

Motivation: a random variable determines a probability distribution (but not vice versa).

Axiom There is a functor $\mathrm{RV}\colon \mathcal{E} \to \mathcal{E}$ together with a natural transformation $\mathbf{P}\colon \mathrm{RV} \Rightarrow \mathcal{M}_1$.

Interpretation:

- The object $\mathrm{RV}(A)$ is the set of A-valued random variables.
- $\mathbf{P}_A\colon \mathrm{RV}(A) \to \mathcal{M}_1(A)$ maps $X \in \mathrm{RV}(A)$ to its probability distribution $\mathbf{P}_X\colon \mathcal{P}(A) \to [0,1]$ (we omit A from the notation).
- Naturality says that, for $X \in \mathrm{RV}(A)$, $f\colon A \to B$ and $B' \subseteq B$,

$$\mathbf{P}_{f[X]}(B') \;=\; \mathbf{P}_X(f^{-1}B')$$

Probability for a single random variable

The equidistribution relation for $X, Y \in \mathrm{RV}(A)$

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

For $X \in \mathrm{RV}\,[0, \infty]$, define the expectation $\mathbf{E}(X) \in [0, \infty]$ by:

$$\begin{aligned}
\mathbf{E}(X) \;&=\; \int_x x \,\mathrm{d}\mathbf{P}_X \\
&=\; \sup_{n \geq 0,\, 0 < c_1 < \cdots < c_n < \infty} \quad \sum_{i=1}^{n} c_i \,.\, \mathbf{P}_X(c_i, c_{i+1}] \qquad (c_{n+1} = \infty)
\end{aligned}$$

Similarly, define variance, moments, etc.

Restriction of random variables

Axiom If $m \colon A \to B$ is a monomorphism then the naturality square below is a pullback.

$$\begin{tikzcd}
\mathrm{RV}(A) \arrow[r, "X \mapsto \mathbf{P}_X"] \arrow[d, "\mathrm{RV}(m)"'] & \mathcal{M}_1(A) \arrow[d, "\mathcal{M}_1(m)"] \\
\mathrm{RV}(B) \arrow[r, "Y \mapsto \mathbf{P}_Y"] & \mathcal{M}_1(B)
\end{tikzcd}$$

Equivalently:

- Given $Y \in \mathrm{RV}(B)$ and $A \subseteq B$ with $\mathbf{P}_Y(A) = 1$, there exists a unique $X \in \mathrm{RV}(A)$ such that $Y = i[X]$, where i is the inclusion function.

Finite limits

Proposition The functor $\mathrm{RV}\colon \mathcal{E} \to \mathcal{E}$ preserves equalisers.

Axiom The functor $\mathrm{RV}\colon \mathcal{E} \to \mathcal{E}$ preserves finite products.

I.e., every $X \in \mathrm{RV}(A)$ and $Y \in \mathrm{RV}(B)$ determines $(X, Y) \in \mathrm{RV}(A \times B)$, and every $Z \in \mathrm{RV}(A \times B)$ is of this form.

Axiom Equality of measures (finite products):

$$\mathbf{P}_{(X,Y)} = \mathbf{P}_{(X',Y')} \\ \Leftrightarrow \quad \forall A' \subseteq A, B' \subseteq B.\ \mathbf{P}_{(X,Y)}(A' \times B') = \mathbf{P}_{(X',Y')}(A' \times B')$$

Almost sure equality

Proposition

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

$$X = Y \quad \Leftrightarrow \quad \begin{array}{ll} \mathbf{P}_{(X,Y)}\{(x,y) \mid x = y\} = 1 & \text{(official notation)} \\ \mathbf{P}(X = Y) = 1 & \text{(informal notation)} \end{array}$$

That is, equality of random variables is given by almost sure equality.

This is an extensionality principle for random variables.

Independence

Independence between $X \in \mathrm{RV}(A)$ and $Y \in \mathrm{RV}(B)$:

$$X \perp\!\!\!\perp Y \quad \Leftrightarrow \quad \forall A' \subseteq A, B' \subseteq B.\ \mathbf{P}_{(X,Y)}(A' \times B') = \mathbf{P}_X(A') \,.\, \mathbf{P}_Y(B')$$

Mutual independence

$$\perp\!\!\!\perp X_1, \ldots, X_n \quad \Leftrightarrow \quad \perp\!\!\!\perp X_1, \ldots, X_{n-1} \text{ and } (X_1, \ldots, X_{n-1}) \perp\!\!\!\perp X_n$$

Infinite mutual independence

$$\perp\!\!\!\perp (X_i)_{i \geq 1} \quad \Leftrightarrow \quad \forall n \geq 1.\ \perp\!\!\!\perp X_1, \ldots, X_n$$

Nonstandard axioms for independence

Axiom Preservation of independence:

$$(\exists Z.\ \Phi(Y,Z)) \ \rightarrow\ \forall X.\ (\, X \perp\!\!\!\perp Y \ \rightarrow\ \exists Z.\ (\, X \perp\!\!\!\perp Z \wedge \Phi(Y,Z)\,)\,)$$

Axiom Existence of independent random variables:

$$\exists Z.\ X \perp\!\!\!\perp Z \wedge Y \sim Z$$

Proposition For every random variable $X \in \mathrm{RV}(A)$ there exists an infinite sequence $(X_i)_{i \geq 0}$ of mutually independent random variables with $X_i \sim X$ for every X_i.

Proposition For every random variable $X \in \mathrm{RV}(A)$ there exists an infinite sequence $(X_i)_{i \geq 0}$ of mutually independent random variables with $X_i \sim X$ for every X_i.

Proof

Let $X_0 = X$.

Proposition For every random variable $X \in \mathrm{RV}(A)$ there exists an infinite sequence $(X_i)_{i \geq 0}$ of mutually independent random variables with $X_i \sim X$ for every X_i.

Proof

Let $X_0 = X$.

Given $X_0, \ldots, X_{i-1}$, the axiom for the existence of independent random variables gives us X_i with $X \sim X_i$ such that $(X_0, \ldots, X_{i-1}) \perp\!\!\!\perp X_i$.

Proposition For every random variable $X \in \mathrm{RV}(A)$ there exists an infinite sequence $(X_i)_{i \geq 0}$ of mutually independent random variables with $X_i \sim X$ for every X_i.

Proof

Let $X_0 = X$.

Given $X_0, \ldots, X_{i-1}$, the axiom for the existence of independent random variables gives us X_i with $X \sim X_i$ such that $(X_0, \ldots, X_{i-1}) \perp\!\!\!\perp X_i$.

This defines a sequence $(X_i)_{i \geq 0}$ by DC. $\square$

Infinite sequence of coin tosses

Axiom There exists $K \in \mathrm{RV}\{0,1\}$ with $\mathbf{P}_K\{0\} = \frac{1}{2} = \mathbf{P}_K\{1\}$.

By the proposition there exists an infinite sequence $(K_i)_{i\geq 0}$ of independent random variables identically distributed to K.

We would like to be able to prove the strong law of large numbers

$$\mathbf{P}\left(\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} K_i \;=\; \frac{1}{2}\right) \;=\; 1$$

Infinite sequence of coin tosses

Axiom There exists $K \in \mathrm{RV}\{0,1\}$ with $\mathbf{P}_K\{0\} = \frac{1}{2} = \mathbf{P}_K\{1\}$.

By the proposition there exists an infinite sequence $(K_i)_{i \geq 0}$ of independent random variables identically distributed to K.

We would like to be able to prove the strong law of large numbers

$$\mathbf{P}\left(\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} K_i \;=\; \frac{1}{2}\right) \;=\; 1$$

At the moment this cannot even be stated.

Countable products

Axiom The functor $\mathrm{RV}\colon \mathcal{E} \to \mathcal{E}$ preserves countable products.

I.e., every $(X_i)_i \in \prod_{i \geq 0} \mathrm{RV}(A_i)$ determines $(X_i)_i \in \mathrm{RV}(\prod_{i \geq 0} A_i)$, and every $Z \in \mathrm{RV}(\prod_{i \geq 0} A_i)$ is of this form.

Axiom Equality of measures (countable products):

$$\begin{aligned} &\mathbf{P}_{(X_i)_i} = \mathbf{P}_{(X'_i)_i} \\ &\quad \Leftrightarrow \quad \forall n.\ \ \forall A'_0 \subseteq A_0,\ \ \ldots,\ \ A'_{n-1} \subseteq A_{n-1}. \\ &\qquad \mathbf{P}_{(X_0,\ldots,X_{n-1})}(A'_0 \times \ldots \times A'_{n-1}) = \mathbf{P}_{(X'_0,\ldots,X'_{n-1})}(A'_0 \times \ldots \times A'_{n-1}) \end{aligned}$$

Countable products

Axiom The functor $\mathrm{RV}\colon \mathcal{E} \to \mathcal{E}$ preserves countable products.

I.e., every $(X_i)_i \in \prod_{i\geq 0} \mathrm{RV}(A_i)$ determines $(X_i)_i \in \mathrm{RV}(\prod_{i\geq 0} A_i)$, and every $Z \in \mathrm{RV}(\prod_{i\geq 0} A_i)$ is of this form.

Axiom Equality of measures (countable products):

$$\begin{aligned}
&\mathbf{P}_{(X_i)_i} = \mathbf{P}_{(X'_i)_i} \\
&\quad \Leftrightarrow \quad \forall n.\ \ \forall A'_0 \subseteq A_0,\ \ \ldots,\ \ A'_{n-1} \subseteq A_{n-1}. \\
&\qquad\quad \mathbf{P}_{(X_0,\ldots,X_{n-1})}(A'_0 \times \ldots \times A'_{n-1}) = \mathbf{P}_{(X'_0,\ldots,X'_{n-1})}(A'_0 \times \ldots \times A'_{n-1})
\end{aligned}$$

The strong law of large numbers can now be stated.

Let $(K_i)_i$ be an infinite sequence of independent coin tosses.

Proposition The measure $\mathbf{P}_{(K_i)_i} : \mathcal{P}(\{0,1\}^{\mathbb{N}}) \to [0,1]$ is translation invariant.

Let $(K_i)_i$ be an infinite sequence of independent coin tosses.

Proposition The measure $\mathbf{P}_{(K_i)_i} : \mathcal{P}(\{0,1\}^{\mathbb{N}}) \to [0,1]$ is translation invariant.

(I.e., for any $A' \subseteq \{0,1\}^{\mathbb{N}}$ and $\alpha \in \{0,1\}^{\mathbb{N}}$, it holds that $\mathbf{P}_{(K_i)_i}(A') = \mathbf{P}_{(K_i)_i}(\alpha \oplus A')$, where $\oplus$ is pointwise xor.)

Let $(K_i)_i$ be an infinite sequence of independent coin tosses.

Proposition The measure $\mathbf{P}_{(K_i)_i} : \mathcal{P}(\{0,1\}^{\mathbb{N}}) \to [0,1]$ is translation invariant.

(I.e., for any $A' \subseteq \{0,1\}^{\mathbb{N}}$ and $\alpha \in \{0,1\}^{\mathbb{N}}$, it holds that $\mathbf{P}_{(K_i)_i}(A') = \mathbf{P}_{(K_i)_i}(\alpha \oplus A')$, where $\oplus$ is pointwise xor.)

Proof

$$\mathbf{P}_{(K_i)_i}(\alpha \oplus A') \;=\; \mathbf{P}_{(\alpha_i \oplus K_i)_i}(A')$$

Let $(K_i)_i$ be an infinite sequence of independent coin tosses.

Proposition The measure $\mathbf{P}_{(K_i)_i} \colon \mathcal{P}(\{0,1\}^{\mathbb{N}}) \to [0,1]$ is translation invariant.

(I.e., for any $A' \subseteq \{0,1\}^{\mathbb{N}}$ and $\alpha \in \{0,1\}^{\mathbb{N}}$, it holds that $\mathbf{P}_{(K_i)_i}(A') = \mathbf{P}_{(K_i)_i}(\alpha \oplus A')$, where $\oplus$ is pointwise xor.)

Proof

$$\mathbf{P}_{(K_i)_i}(\alpha \oplus A') \;=\; \mathbf{P}_{(\alpha_i \oplus K_i)_i}(A')$$

But $\mathbf{P}_{(\alpha_i \oplus K_i)_i} \;=\; \mathbf{P}_{(K_i)_i}$ by the criterion on previous slide. $\square$

Convergence in probability

Suppose (A, d) is a metric space. A sequence $(X_i)_{i \geq 0}$ of A-valued random variables converges in probability to $X_\infty \in \mathrm{RV}(A)$ if, for any $\epsilon > 0$,

$$\lim_{n \to \infty} \mathbf{P}(d(X_n, X_\infty) > \epsilon) \;=\; 0 \;.$$

Proposition

The function

$$(X, Y) \;\mapsto\; \mathbf{E}(\min(d(X, Y), 1))$$

defines a metric on $\mathrm{RV}(A)$ whose convergence relation is convergence in probability. This metric is complete whenever (A, d) is itself a complete metric space.

Conditional expectation

We say that $Z \in \mathrm{RV}(B)$ is functionally dependent on $Y \in \mathrm{RV}(A)$ (notation $Z \leftarrow Y$) if there exists $f : A \to B$ such that $Z = f[Y]$.

Proposition

For any $X \in \mathrm{RV}\,[0, \infty]$ and $Y \in \mathrm{RV}(A)$, there exists a unique random variable $Z \in \mathrm{RV}\,[0, \infty]$ satisfying:

- $Z \leftarrow Y$, and
- for all $A' \subseteq A$

$$\mathbf{E}(Z \,.\, \mathbf{1}_{A'}(Y)) = \mathbf{E}(X \,.\, \mathbf{1}_{A'}(Y))$$

The unique such Z defines the conditional expectation $\mathbf{E}(X \mid Y)$.

A meta-theoretic property

Every sentence of the form

$$\forall X_1, Y_1 \in \mathrm{RV}(A_1), \ \dots \ , \forall X_n, Y_n \in \mathrm{RV}(A_n)$$
$$\Phi(X_1, \dots, X_n) \ \wedge \ (X_1, \dots, X_n) \sim (Y_1, \dots, Y_n)$$
$$\rightarrow \ \Phi(Y_1, \dots, Y_n)$$

is true.

A meta-theoretic property

Every sentence of the form

$$\forall X_1, Y_1 \in \mathrm{RV}(A_1), \ \ldots \ , \forall X_n, Y_n \in \mathrm{RV}(A_n)$$
$$\Phi(X_1, \ldots, X_n) \ \wedge \ (X_1, \ldots, X_n) \sim (Y_1, \ldots, Y_n)$$
$$\rightarrow \ \Phi(Y_1, \ldots, Y_n)$$

is true.

Definable properties are equidistribution invariant.

To do

Develop a good amount of probability theory (e.g., including continuous time Markov processes, and stochastic differential equations).

Constructive and (hence) computable versions.

Type theoretic version.