My name is stochastic calculus but everybody calls me calculus

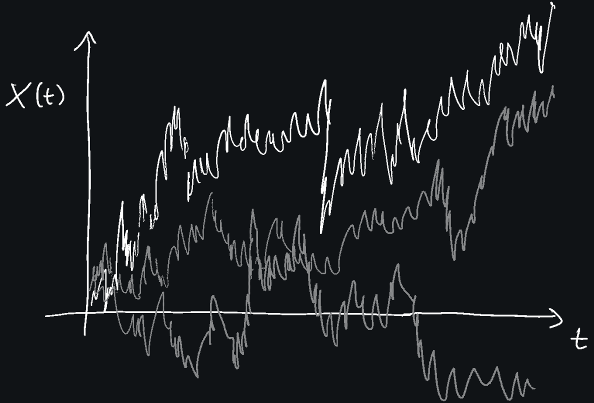
Matteo Capucci, University of Strathclyde

March 11th, 2021

(Day 434 of the COVID Era)

Idea

"random variables"



"almost everywhere"

$\bigcup_n \bigcap_m A_n$ events as propositions

Preliminaries

Let's fix a probability space $W = (W, \mathcal{F}, \mathbb{P})$. Recall that

W : **Set**

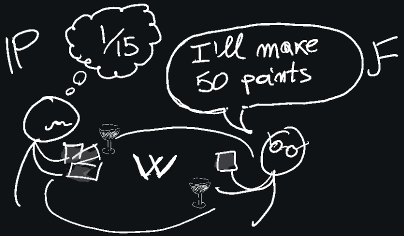
Space of 'outcomes' of a stochastic experiment

$\mathcal{F}$: σ-field

Space of 'events', i.e. σ-algebra of subsets of W

$\mathbb{P} : \mathcal{F} \to \mathbb{R}^+$

'Probability measure', i.e. an assignment of beliefs to events



Preliminaries

We call

$$\ker \mathbb{P} = \{A \in \mathcal{F} \mid \mathbb{P}(A) = 0\}$$

the σ-ideal of null sets.

Let $V = (V, \mathcal{G}, \mathbb{Q})$.

$$\underline{\text{measurable maps}}$$

$$\mathbf{Msbl}(W, V) = \{f : W \to V \mid f^{-1}\mathcal{G} \subseteq \mathcal{F}\}$$

$$\underline{\text{null-reflecting maps}}$$

$$\mathbf{Msbl}_0(W, V) = \{f \in \mathbf{Msbl}(W, V) \mid f^{-1} \ker \mathbb{Q} \subseteq \ker \mathbb{P}\}$$

Preliminaries

Definition

An E-valued **stochastic process** on W is a collection of measurable maps

$$X_t : W \to E, \qquad t \in I$$

where I is a poset of **time**s and E is a measurable space of **values**.

Example

Toss a coin and bet 1$ on heads: $W = \{H, T\}^{\mathbb{N}}$, $E = \mathbb{R}$, $I = \mathbb{N}$,

$$X_n(\omega) = \sum_{i=0}^{n} (\delta_H^{\omega_i}(\omega) - \delta_T^{\omega_i}(\omega))$$

Example

You have a financial portfolio, your profit are a stochastic process X_t which is a function of another stochastic process M_t, the market.

Preliminaries

A central concept in stochastic calculus is that of filtration:

Definition

A **filtration** on W is a sequence of σ-fields $\{\mathcal{F}_t\}_{t\in I}$ such that $\mathcal{F}_t \subseteq \mathcal{F}_s$ when $t \leq s$. One sets $\mathcal{F}_\infty := \mathcal{F}$.

It gives a 'causal structure' to the observed events.

Definition

An **adapted process** is a process $\{X_t\}_{t\in I}$ such that

$$\text{for each } t \in I, \qquad X_t \text{ is } \mathcal{F}_t\text{-measurable.}$$

i.e. the value of X at t only depends on info available until that point.

Preliminaries

Itô calculus: a popular way to do stochastic calculus

1. We can **integrate** wrt stochastic processes:

$$\int_0^t X_t \, \mathrm{d}B_t = \lim_n \sum_i X_{t_i}^{(n)} (B_{t_{i+1}} - B_{t_i})$$

2. We can discuss **differential problems**:

$$\mathrm{d}X_t = \sigma_t X_t \mathrm{d}B_t + \mu_t X_t \mathrm{d}t$$

Here B_t is a **Brownian motion** $\rightsquigarrow$ archetypal diffusion process

It's not exactly like elementary calculus though, e.g.

$$\mathrm{d}f(B_t) = f'(B_t)\, \mathrm{d}B_t + \frac{1}{2} f''(B_t) \mathrm{d}t$$

because B_t has 'unbounded variation'

Idea

Original goal:

'elementary'
stochastic calculus
topos

It turned out to be much harder than I thought

and probably not that simple

...still, plenty of interesting stuff along the way!

Warning: work still in progress!

Essential algebra

Define the preorder $(\mathcal{F}, \leq)$:

$$A \leq B \quad \underline{\text{iff}} \quad \mathbb{P}(A \Rightarrow B) = 1$$

i.e. if A *is essentially contained in* B:

The **essential algebra** of W is the *posetal reflection* of $(\mathcal{F}, \leq)$:

$$\mathbb{F} := (\mathcal{F}/\ker \mathbb{P}, \leq)$$

Terminology: *informally, we use 'essential' to say 'up to a null set'.*

Tripos approach

Theorem

$\mathbb{F}$ *is a* **complete** *Boolean algebra.*

Proof.

$\mathbb{F}$ is σ-complete and every essential partition of W can be refined to a countable essential partition (CCC). □

In [Sco67], Dana Scott forces CH using a Boolean-valued model whose algebra of truth values is an essential algebra.
Therefore in [Cap20], we call

$$\mathbf{Scott}[W] := \mathbb{F}^{(-)} : \mathbf{Set}^{\mathrm{op}} \to \mathbf{Cat}$$

the **Scott tripos** associated to W.

Scott tripos

This tripos captures 'logic valued in W':

1. A predicate on X : **Set** is a map $\varphi : X \to \mathbb{F}$ (*W-predicates*).
2. $\wedge, \vee, \Rightarrow$ are pointwise essential intersection/union/implication.
3. true $= [W] =$ *almost everywhere*,
   false $= [\varnothing] =$ *almost nowhere*.
4. Quantifiers:
   $$\forall x{:}X \varphi(x) := \bigwedge_{x \in X} \varphi(x), \qquad \exists x{:}X \varphi(x) := \bigvee_{x \in X} \varphi(x)$$
5. Power objects:
   $$(\pi(X), \in_X) = (\mathbb{F}^X, \mathrm{ev} : \mathbb{F}^{\mathbb{F}^X} \times \mathbb{F}^X \to \mathbb{F}).$$

Scott tripos

The tripos should be be 'hybridized' with quantitative notions from probability theory:

1. Given a proposition φ, we can talk about its probability $\mathbb{P}(\varphi)$.
2. Given a family of propositions $\{\varphi_i\}_{i\in I}$, we can talk about their independence.

Open question

Can this structure be captured by the tripos?

I have some ideas about this... reach out if you want to talk about it.

Tripos to topos

Since **Scott**$[W]$ is localic, its associated topos can be presented as

$$\mathrm{Sh}\,\mathbb{F} \simeq \mathrm{Sh}_{\mathrm{a.e.}}\,\mathcal{F} \hookrightarrow \mathrm{Sh}\,\mathcal{F}$$

where 'a.e.' is the almost-everywhere topology on $\mathcal{F}$:

$$\{A_i\}_{i\in I} \text{ covers } A \quad \underline{\text{iff}} \quad \{A_i\}_{i\in I} \text{ 'essentially covers' } A$$

Definition

$$\mathrm{Sh}_{\mathrm{a.e.}}\,W := \mathrm{Sh}\,\mathbb{F} \simeq \mathrm{Sh}_{\mathrm{a.e.}}\,\mathcal{F} \simeq \mathbf{Pitts}[\mathbf{Scott}[W]]$$

It's the topos of 'random sets and random functions' wrt to W.

It's a classical universe: we have both LEM and AC.

Geometry

We can look at this topos geometrically:

algebraic geometry	**probability theory**
base scheme S	base probability space W
regular functions	essentially bounded r.v.
structure sheaf $\mathcal{O}_S$	'structure sheaf' $L^\infty(-, \mathbb{R})$
commutative rings	commutative Von Neumann algebras (kind of)
Grothendieck topos $\mathrm{Sh}\, S$	Kolmogorov topos $\mathrm{Sh}_{\mathrm{a.e.}}\, W$

Geometry

The rabbit hole runs deep...

Gelfand duality

Localizable measurable spaces

Riesz duality

Structured topoi

σ-locales

Geometry

There is a **measurability structure**

$$\mathcal{O}_W : \mathbf{Msbl} \to \mathrm{Sh}_{\text{a.e.}}\, W$$

given by **essential functors of points**

$$\mathcal{O}_W(E) = \mathbf{Msbl}(-, E)/\underset{\text{a.e.}}{=}$$

and any null-reflecting map $f : W \to V$ induces a morphism

$$\begin{tikzcd}
\mathrm{Sh}_{\text{a.e.}}\, W \arrow[rr, bend right=10, "f_*"'] \arrow[rr, phantom, "\perp"] & & \mathrm{Sh}_{\text{a.e.}}\, V \arrow[ll, bend right=10, "f^*"'] \\
& \mathbf{Msbl} \arrow[ul, "\mathcal{O}_W"] \arrow[ur, "\mathcal{O}_V"'] \arrow[u, phantom, "\Leftarrow"] &
\end{tikzcd}$$

$$f^*\mathcal{O}_V \Longrightarrow \mathcal{O}_W \quad \text{or, equivalently} \quad \mathcal{O}_V \Longrightarrow f_*\mathcal{O}_W$$

Internal notions

We can start looking for the objects of probability theory inside $\mathrm{Sh}_{\mathrm{a.e.}}\, W$ (with the help of [Jac06]).

Random variables

The internal real numbers are given by

$$\mathcal{R}_W = \mathcal{O}_W(\mathbb{R})$$

Hence this confirms 'random variables' are indeed 'variables' in a 'random' setting.

Then we can prove things like

Theorem

A sequence of random variables $\{X_n\}_{n\in\mathbb{N}}$ converges almost surely iff it converges in the usual sense as a sequence of internal real numbers.

Also, L^p_{loc} spaces become subsets of the real numbers.

Integration

There's a sheaf $\mathcal{M}$ of measures on W, which is definable internally

$$\mathcal{M} = \{\mu \in \text{a.e.}(P\mathbb{Q}^+) \mid \mu \text{ is an `additive' upper cut}\}$$

Thus integration is straightforward: first, on simple rational functions:

$$\begin{aligned} \int - \,\mathrm{d}- : \mathcal{Q}^+ \times \mathcal{M} &\longrightarrow \mathcal{M} \\ (q, \mu) &\longmapsto q\mu \end{aligned}$$

then we extend by continuity:

$$\begin{aligned} \int - \,\mathrm{d}- : \mathcal{R}^+ \times \mathcal{M} &\longrightarrow \mathcal{M} \\ (X, \mu) &\longmapsto \underbrace{\bigvee_{q \le X} q\mu}_{\int\_ X \,\mathrm{d}\mu} \end{aligned}$$

In particular, we defined $\mathbb{E}[X]$.

Radon–Nikodym

Now recall

Theorem (Radon–Nikodym)

Let ν be a measure on W such that $\mu(A) = 0$ implies $\nu(A) = 0$. Then there exists a positive measurable random variable $\frac{\mathrm{d}\nu}{\mathrm{d}\mu}$ such that

$$\int_{-} \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \, \mathrm{d}\mu = \nu(-)$$

It says every absolutely continuous measure wrt μ has a μ-density.

Radon–Nikodym

We can prove the theorem internally ([Jac06]). Fix $\mu : \mathcal{M}$:

1. μ-integration in an arrow $\int - \mathrm{d}\mu : \mathcal{R}^+ \to \mathcal{M}$.
2. Its image $\mathcal{M}_{\int \mu}$ is contained in $\mathcal{M}^{\ll \mu}$, the object of absolutely continuous measures with respect to μ.
3. Then we want to show $\mathcal{M}_{\int \mu} = \mathcal{M}^{\ll \mu}$ so that $\frac{\mathrm{d}-}{\mathrm{d}\mu}$ is the splitting of the image of $\int - \mathrm{d}\mu$:

$$\frac{\mathrm{d}-}{\mathrm{d}\mu} : \mathcal{M}^{\ll \mu} \longrightarrow \mathcal{R}^+$$

To do so, one fixes $\nu : \mathcal{M}^{\ll \mu}$ and defines

$$\frac{\mathrm{d}\nu}{\mathrm{d}\mu} = \bigvee \{ q \in \mathcal{Q}^+ \mid q\mu \leq \nu \}$$

up to technicalities, this concludes the proof!

Martingales

For stochastic calculus, the most useful application of Radon–Nikodym is given by the definition of **conditional expectations**:

Definition

Given a filtration $\{F_t\}_{t \in I}$, the conditional expectation of X at t is defined as

$$\mathbb{E}[X|\mathcal{F}_t] = \frac{\mathrm{d}(\int X \, \mathrm{d}\mathbb{P})|_{\mathcal{F}_t}}{\mathrm{d}\mathbb{P}|_{\mathcal{F}_t}}$$

Definition

A **martingale** is an adapted stochastic process such that

$$\mathbb{E}[X_{t+s}|\mathcal{F}_t] = X_t$$

It's a process not expected to change, e.g. the coin tosses bet, B_t.

Internalizing martingales

To define a martingale internally we need

1. An internal definition of ‘stochastic process’
2. An internal definition of ‘adaptedness’ $\rightsquigarrow$ filtrations
3. An internal definition of ‘conditional expectation at t’.

Stochastic processes

Theorem

An E-valued stochastic process $\{X_t\}_{t \in I}$ on W corresponds to a map

$$X : \Delta I \to \mathcal{O}_W(E)$$

in $\mathrm{Sh}_{a.e.}\, W$.

If X is **measurable** as a map $W \times I \to E$, then this map can be lifted to

$$X : \mathcal{O}_W(I) \to \mathcal{O}_W(E).$$

By extension, any map $X : I \to E$ in $\mathrm{Sh}_{\mathrm{a.e.}}\, W$ is a 'stochastic process'.

Stochastic processes

Two facts:

Theorem

Two stochastic processes $X, Y : \Delta I \to \mathcal{O}_W(E)$ *are equal in the logic of* $\mathrm{Sh}_{a.e.}\, W$ *iff they are* **indistinguishable**, *i.e.*

$$\mathbb{P}(\forall t \in I,\, X_t = Y_t) = 1$$

Therefore we can state & prove theorems like:

Theorem

A stochastic process is almost surely continuous iff it's continuous from the internal point of view.

Filtrations

We need a way to express '*measurability at t*' for elements in $\mathcal{O}_W(E)$.

A filtration induces a chain of null-reflecting maps (each carried by 1_W)

$$W_\infty \longrightarrow \cdots \longrightarrow W_t \longrightarrow \cdots \longrightarrow W_0$$

where $W_t = (W, \mathcal{F}_t, \mathbb{P}|_{\mathcal{F}_t})$.

which in turn becomes a chain of (surjections of) Kolmogorov topoi:

$$\begin{tikzcd} \mathrm{Sh}_{\text{a.e.}}\, W_\infty \arrow[r, shift right] \arrow[r, phantom, "\perp"] & \cdots \arrow[l, hook', shift right] \arrow[r, shift right] \arrow[r, phantom, "\perp"] & \mathrm{Sh}_{\text{a.e.}}\, W_t \arrow[l, hook', shift right] \arrow[r, shift right] \arrow[r, phantom, "\perp"] & \cdots \arrow[l, hook', shift right] \arrow[r, shift right] \arrow[r, phantom, "\perp"] & \mathrm{Sh}_{\text{a.e.}}\, W_0 \arrow[l, hook', shift right] \end{tikzcd}$$

Filtrations

Remember that null-reflecting maps induce maps

$$
\begin{tikzcd}
\mathrm{Sh}_{\text{a.e.}}\, W_\infty \arrow[rr, phantom, "\perp"] \arrow[rr, shift right=2, "1_{t*}"'] & & \mathrm{Sh}_{\text{a.e.}}\, W_t \arrow[ll, shift right=2, hook', "1_t^*"'] \\
& \mathbf{Msbl} \arrow[ul, "\mathcal{O}_\infty"'] \arrow[ur, "\mathcal{O}_t"] \arrow[u, phantom, "\Leftarrow"] &
\end{tikzcd}
$$

Hence for every $t \in I$, $E : \mathbf{Msbl}$, we have

$$1_t^* \mathcal{O}_t(E) =: \mathcal{O}_t^*(E) \longrightarrow \mathcal{O}_\infty(E)$$

Filtrations

The inclusion $\mathbb{F}_t \hookrightarrow \mathbb{F}_\infty$ has both adjoints, in particular

$$\begin{aligned} 1_t^{-1} \dashv \Diamond_t : \mathbb{F}_\infty &\longrightarrow \mathbb{F}_t \\ A &\longmapsto \bigwedge \{B \in \mathcal{F}_t \mid A \leq B\} \end{aligned}$$

We can interpret it as a '*possibly at t*' modal operator.

Thus for each $A \in \mathbb{F}_\infty$:

$$\mathcal{O}_t^*(E)(A) := \mathcal{O}_t(E)(\Diamond_t A) = \mathbf{Msbl}(\Diamond_t A, E)/\underset{\text{a.e.}}{=}$$

and

$$\begin{aligned} \mathcal{O}_t^*(E)(A) &\longrightarrow \mathcal{O}_\infty(E)(A) \\ \Diamond_t A \xrightarrow{f} E &\longmapsto A \hookrightarrow \Diamond_t A \xrightarrow{f} E \end{aligned}$$

Notice: $\Diamond_t A \subseteq W_t$, while $A \subseteq W_\infty$

$\rightsquigarrow$ $\mathcal{O}_t^*(E)$ contains $\mathcal{F}_t$-measurable elements.

Filtrations

For each $t \in I$, let $\mathsf{msbl}(t)$ be the image of $\mathcal{O}^*_t(E)$ in $\mathcal{O}_\infty(E)$.

Sheafification extends this definition: given $\tau : \Delta I$,

$$W \vDash e \in \mathsf{msbl}(\tau) \quad \underline{\text{iff}} \quad \text{for all } s \in I,\ \{\tau = s\} \vDash e \in \mathsf{msbl}(s)$$

$$\begin{tikzcd} \mathsf{const}(I) \arrow[d, "\eta_{\mathsf{a.e.}}"'] \arrow[dr, "\mathsf{msbl}"] & \\ \Delta I \arrow[r, dashed, "\mathsf{msbl}"] & P\mathcal{O}_\infty(E) \end{tikzcd}$$

We can extend the definition of msbl to $\mathcal{O}_\infty(I)$:

$$W \vDash \forall \tau : \mathcal{O}_\infty(I),\ \mathsf{msbl}(\tau) \Leftrightarrow \forall \sigma : \Delta I,\ \sigma \leq \tau \Rightarrow \mathsf{msbl}(\sigma)$$

This recovers $\mathcal{F}_\tau$-measurability for a **random time** $\tau : W \to I$.

Filtrations

Definition

We call $X : \Delta I \to \mathcal{O}_\infty(E)$ **adapted** iff

$$W \vDash \forall t : \Delta I,\ X(t) \in \mathsf{msbl}(t)$$

Theorem

External adaptedness coincides with internal adaptedness

Open question

Is there a better **internal/synthetic** notion of adaptedness?

Alternative construction suggested by Morgan Rogers (priv. comm.):

1. The object $\mathcal{F}_\bullet = \{\mathcal{F}_t\}_{t \in I}$ is naturally in $\mathrm{Psh}(I, \leq)$
2. Adapted processes are maps of internal sheaves over $\mathcal{F}_\bullet$!

Conditional expectation

Restrictions of measures is given by the canonical inclusion

$$1_{t*}\mathcal{M}_\infty \xhookrightarrow{-|_t} \mathcal{M}_t$$

Then we define

$$\mathbb{E}[-|\mathcal{F}_t]_t = 1_{t*}\mathcal{R}^+_\infty \xrightarrow{(\int - \mathrm{d}\mathbb{P})|_t} \mathcal{M}^{\ll \mathbb{P}|_t} \xrightarrow{\frac{\mathrm{d}}{\mathrm{d}\mathbb{P}|_t}} \mathcal{R}^+_t$$

Finally:

$$\mathbb{E}[-|\mathcal{F}_t]_\infty := \mathcal{R}^+_\infty \xrightarrow{\text{ext. by } 0} 1_t^* 1_{t*}\mathcal{R}^+_\infty \xrightarrow{1_t^*\mathbb{E}[X|\mathcal{F}_t]_t} 1_t^*\mathcal{R}^+_t \xrightarrow{A \hookrightarrow \Diamond_t A} \mathcal{R}^+_\infty$$

If we unpack the definitions, this is what's happening (hand-wavingly):

$$X \mapsto \bigvee \{q_t \in 1_t^*\mathcal{Q}^+_t \mid q_t \leq (\textstyle\int X \, \mathrm{d}\mathbb{P})/\mathbb{P}\}$$

best approximation at t of X's average.

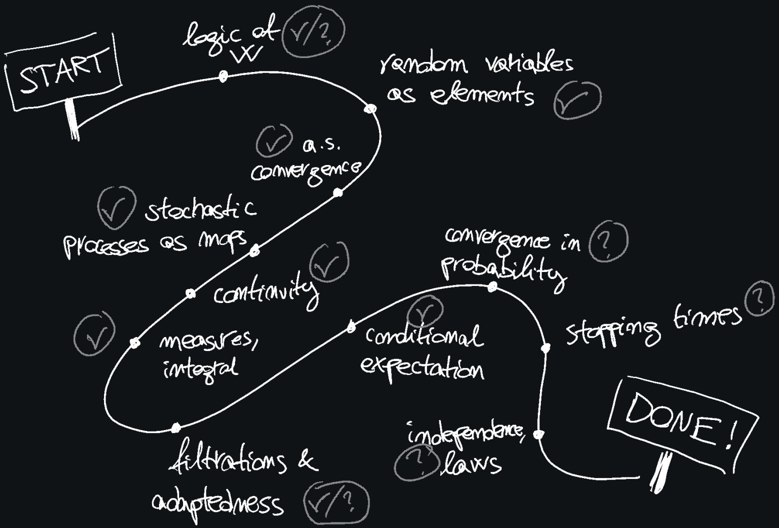
Martingales, finally

Definition

A map $X : \Delta I \to \mathcal{O}_\infty(E)$ is a **martingale** if it is adapted and

$$W \vDash \forall t{:}\Delta I\ \forall s \geq t,\ \mathbb{E}[X(t+s)|\mathcal{F}_t]_\infty = X(t),$$

Roadmap



START
logic of W
random variables as elements
a.s. convergence
stochastic processes as maps
continuity
measures, integral
filtrations & adaptedness
conditional expectation
convergence in probability
stopping times
independence, laws
DONE!

Thanks for your attention!

Questions?

References I

Dana Scott. "A Proof of the Independence of the Continuum Hypothesis". In: *Mathematical systems theory* 1.2 (1967), pp. 89–111. URL: https://perma.cc/38WZ-KJTP.

Matteo Capucci. "Internal mathematics for stochastic calculus: a tripos-theoretic approach". MA thesis. Università degli Studi di Padova, 2020. URL: https://perma.cc/CX5B-S94T.

Matthew Tobias Jackson. "A sheaf theoretic approach to measure theory". PhD thesis. University of Pittsburgh, 2006.

Alex Simpson. "Measure, randomness and sublocales". In: *Annals of Pure and Applied Logic* 163.11 (2012), pp. 1642–1659.

References II

Ingo Blechschmidt. “Using the internal language of toposes in algebraic geometry”. PhD thesis. Universität Augsburg, 2018.

Jacob Lurie. “Derived algebraic geometry V: Structured spaces”. In: *arXiv preprint arXiv:0905.0459* (2009).

Asgar Jamneshan and Terence Tao. “Foundational aspects of uncountable measure theory: Gelfand duality, Riesz representation, canonical models, and canonical disintegration”. In: *arXiv preprint arXiv:2010.00681* (2020).