An overview of categorical probability theory

Ruben Van Belle

September 2022
1. What are *the* objects and morphisms of probability theory?
2. (Incomplete) overview of the current literature.
1. What are the objects and morphisms of probability theory?
2. (Incomplete) overview of the current literature.
What are the objects and morphisms of probability theory?

1. Probability spaces?
2. Random variables?
3. Markov kernels?
What are *the* objects and morphisms of probability theory?

1. Probability spaces?
What are *the* objects and morphisms of probability theory?

1. Probability spaces?
2. Random variables?
What are the objects and morphisms of probability theory?

1. Probability spaces?
2. Random variables?
3. Markov kernels?
1. Probability spaces

- The category $\textbf{Prob}$:
  - objects: probability spaces $(\Omega, \mathcal{F}, \mathbb{P})$
  - morphisms: measure-preserving maps, i.e. measurable maps $f : \Omega_1 \to \Omega_2$ such that $\mathbb{P}_1 \circ f^{-1} = \mathbb{P}_2$. 

$\text{Prob}$ does not have many (co)limits:
- In general no (co)product or equalizers.
- There is no initial object (unbounded randomness).
- There is a terminal objects and coequalizers exist.
1. Probability spaces

- The category **Prob**:  
  - objects: probability spaces \((\Omega, \mathcal{F}, \mathbb{P})\)  
  - morphisms: measure-preserving maps, i.e. measurable maps \(f : \Omega_1 \to \Omega_2\) such that \(\mathbb{P}_1 \circ f^{-1} = \mathbb{P}_2\).

- **Prob** does not have many (co)limits  
  - In general no (co)product or equalizers.
  - There is no initial object (unbounded randomness).
  - There is a terminal objects and coequializers exist.
1. Probability spaces

- (many couplings) For $\mathbb{P}_1$ and $\mathbb{P}_2$ on $\Omega$, there exists many $\mathbb{P}$ on $\Omega^2$ such that:

$$\mathbb{P} \circ \pi_1^{-1} = \mathbb{P}_1 \quad \text{and} \quad \mathbb{P} \circ \pi_2^{-1} = \mathbb{P}_2.$$
1. Probability spaces

- *(many couplings)* For \( P_1 \) and \( P_2 \) on \( \Omega \), there exists many \( P \) on \( \Omega^2 \) such that:

\[
P \circ \pi_1^{-1} = P_1 \quad \text{and} \quad P \circ \pi_2^{-1} = P_2.
\]

But we want to talk about the *interaction* between stochastic events.
1. Probability spaces

- *(many couplings)* For $\mathbb{P}_1$ and $\mathbb{P}_2$ on $\Omega$, there exists many $\mathbb{P}$ on $\Omega^2$ such that:

$$\mathbb{P} \circ \pi_1^{-1} = \mathbb{P}_1 \quad \text{and} \quad \mathbb{P} \circ \pi_2^{-1} = \mathbb{P}_2.$$ 

But we want to talk about the \textbf{interaction} between stochastic events. We can not do this using probability spaces (e.g. it does not make sense to say that two probability spaces are independent).
1. Probability spaces

- *(many couplings)* For $P_1$ and $P_2$ on $\Omega$, there exists many $P$ on $\Omega^2$ such that:

  $$P \circ \pi_1^{-1} = P_1 \quad \text{and} \quad P \circ \pi_2^{-1} = P_2.$$ 

But we want to talk about the *interaction* between stochastic events. We can not do this using probability spaces (e.g. it does not make sense to say that two probability spaces are independent).

- Probability spaces are an important aspect of probability theory, but *not the main objects of interest*.

  $\rightarrow$ Probability theory $\neq$ measure theory with measures of total mass 1
2. Random variables

Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $E$ be a Polish space, a **random variable** is a measurable map $X : \Omega \to E$.

- **bounded randomness**
2. Random variables

Fix a probability space \((\Omega, \mathcal{F}, \mathbb{P})\). Let \(E\) be a Polish space, a random variable is a measurable map \(X : \Omega \to E\).

- bounded randomness
- Two random variables \(X_1, X_2 : \Omega \to \mathbb{R}\), determine a random variable that describes their interaction:

\[(X_1, X_2) : \Omega \to \mathbb{R}^2\]
2. Random variables

Fix a probability space \((\Omega, \mathcal{F}, \mathbb{P})\). Let \(E\) be a Polish space, a random variable is a measurable map \(X : \Omega \to E\).

- bounded randomness
- Two random variables \(X_1, X_2 : \Omega \to \mathbb{R}\), determine a random variable that describes their interaction:

\[
(X_1, X_2) : \Omega \to \mathbb{R}^2
\]

- This does not look very categorical:
  - The domain and codomain seem of a different type.
  - What are morphisms between random variables? \(\to\) order, martingale relation?
2. Random variables

Fix a probability space \((\Omega, \mathcal{F}, \mathbb{P})\).

- A random variable \(X : \Omega \to E\) induces a probability measure on \(E\), namely \(\mathbb{P} \circ X^{-1}\).
2. Random variables

Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

- A random variable $X : \Omega \to E$ induces a probability measure on $E$, namely $\mathbb{P} \circ X^{-1}$. We are strictly losing information.

- A random variable $X : \Omega \to \mathbb{R}$ can be interpreted as a density function. It induces a measure on $\Omega$ by

$$A \mapsto \mathbb{E}[X 1_A].$$
2. Random variables

Fix a probability space \((\Omega, \mathcal{F}, P)\).

- A random variable \(X: \Omega \to E\) induces a probability measure on \(E\), namely \(P \circ X^{-1}\). We are strictly losing information.

- A random variable \(X: \Omega \to \mathbb{R}\) can be interpreted as a **density** function. It induces a measure on \(\Omega\) by

\[
A \mapsto E[X \mathbb{1}_A].
\]

This measure is absolutely continuous with respect to \(P\). Every such measure is of this form (**Radon-Nikodym theorem**).
3. Markov kernels

Let $\Omega_1$ and $\Omega_2$ be measurable spaces and let $\mathcal{G}\Omega_2$ be the space of probability measures on $\Omega_2$. A **Markov Kernel** is a measurable map

$$f : \Omega_1 \to \mathcal{G}\Omega_2.$$ 

For Markov kernels $f_1 : \Omega_1 \to \mathcal{G}\Omega_2$ and $f_2 : \Omega_2 \to \mathcal{G}\Omega_3$, there is a Markov kernel $f : \Omega_1 \to \mathcal{G}\Omega_3$ defined by

$$f(\omega)(A) = \int_{\Omega_2} f_2(\omega_2)(A) f_1(\omega_1)(d\omega_2),$$

for all $\omega_1 \in \Omega_1$ and measurable $A \subseteq \Omega_3$. 
3. Markov kernels

Let $\Omega_1$ and $\Omega_2$ be measurable spaces and let $\mathcal{G}\Omega_2$ be the space of probability measures on $\Omega_2$. A **Markov Kernel** is a measurable map

$$f : \Omega_1 \to \mathcal{G}\Omega_2.$$ 

For Markov kernels $f_1 : \Omega_1 \to \mathcal{G}\Omega_2$ and $f_2 : \Omega_2 \to \mathcal{G}\Omega_3$, there is a Markov kernel $f : \Omega_1 \to \mathcal{G}\Omega_3$ defined by

$$f(\omega)(A) = \int_{\Omega_2} f_2(\omega_2)(A)f_1(\omega_1)(d\omega_2),$$

for all $\omega_1 \in \Omega_1$ and measurable $A \subseteq \Omega_3$.

- Describes interactions between different stochastic events.
3. Markov kernels

For a Markov kernel $f : \Omega_1 \rightarrow \mathcal{G}\Omega_2$ and a probability measure $\mathbb{P}_1$ on $\Omega_1$, there is a probability measure $\mathbb{P}$ on $\Omega_1 \times \Omega_2$ such that

$$\mathbb{P}(A \times B) = \int_A f(\omega)(B)\mathbb{P}_1(d\omega),$$

for all measurable $A \subseteq \Omega_1$ and $B \subseteq \Omega_2$. 
3. Markov kernels

For a Markov kernel \( f : \Omega_1 \to \mathcal{G}\Omega_2 \) and a probability measure \( P_1 \) on \( \Omega_1 \), there is a probability measure \( P \) on \( \Omega_1 \times \Omega_2 \) such that

\[
P(A \times B) = \int_A f(\omega)(B)P_1(d\omega),
\]

for all measurable \( A \subseteq \Omega_1 \) and \( B \subseteq \Omega_2 \). *In general*, not every probability measure on \( \Omega_1 \times \Omega_2 \) is of this form (*regular conditional probabilities*).
For a Markov kernel \( f : \Omega_1 \rightarrow \mathcal{G}\Omega_2 \) and a probability measure \( P_1 \) on \( \Omega_1 \), there is a probability measure \( P \) on \( \Omega_1 \times \Omega_2 \) such that

\[
P(A \times B) = \int_A f(\omega)(B)P_1(d\omega),
\]

for all measurable \( A \subseteq \Omega_1 \) and \( B \subseteq \Omega_2 \). In general, not every probability measure on \( \Omega_1 \times \Omega_2 \) is of this form (regular conditional probabilities).

Let \( P \) be a probability measure on \( \Omega_1 \times \Omega_2 \). The assignment

\[
A \mapsto \mathbb{E}[1_{\Omega_1 \times A} \mid \pi^{-1}(\mathcal{F}_1)]
\]

is \( P \circ \pi_1^{-1} \)-almost surely \( \sigma \)-additive.
Conclusion

Random variables \( \mathbb{P} \circ X^{-1} \)

Probability spaces \( \mathbb{E}[X1_\cdot] \)

Markov kernels \( \int - f(\omega)(-) \mathbb{P}(d\omega) \)

Markov chains

An overview of categorical probability theory

Ruben Van Belle

September 2022 10 / 40
An overview of categorical probability theory

Probability monads

Algebras: integration  Kleisli categories: Markov kernels

Markov categories
Where do probability monads come from?
Lawvere: probabilistic mappings (1962)

Discusses a category of Markov kernels.
Discusses a category of Markov kernels.

- Lawvere introduces the **category of probabilistic mappings** \( \mathcal{P} \):
  - objects: measurable spaces \((\Omega, \mathcal{F})\),
  - morphisms: Markov kernels (‘probabilistic mappings’) \( f : \Omega_1 \to \mathcal{G}\Omega_2 \),
  - composition: composition of Markov kernels
Lawvere: probabilistic mappings (1962)

Discusses a category of Markov kernels.

- Lawvere introduces the **category of probabilistic mappings** $\mathcal{P}$:
  - objects: measurable spaces $(\Omega, \mathcal{F})$,
  - morphisms: Markov kernels ('probabilistic mappings') $f : \Omega_1 \to \mathcal{G}\Omega_2$,
  - composition: composition of Markov kernels

- Define $\Phi : \mathcal{P}^\mathbb{N} \to \mathcal{P}^\mathbb{N}$ by
  \[
  \Phi((\Omega_n)_n)_m := \prod_{k < m} \Omega_k
  \]

A **discrete stochastic process** is an object $\Omega$ in $\mathcal{P}^\mathbb{N}$ together with a morphism $f : \Phi(\Omega) \to \Omega$ in $\mathcal{P}^\mathbb{N}$, i.e. a collection of Markov kernels

\[
\left( f_m : \prod_{k < m} \Omega_k \to \mathcal{G}\Omega_m \right)_m
\]
A morphism of discrete stochastic processes from \((\Omega^1, f^1)\) to \((\Omega^2, f^2)\) is a morphism \(g : \Omega^1 \to \Omega^2\) in \(\mathcal{P}^\mathbb{N}\) such that

\[ g \circ f = f' \circ \Phi(g). \]

Let \textbf{Stoch} be the category of stochastic processes.
A morphism of discrete stochastic processes from \((\Omega^1, f^1)\) to \((\Omega^2, f^2)\) is a morphism \(g : \Omega^1 \rightarrow \Omega^2\) in \(P^\mathbb{N}\) such that
\[
g \circ f = f' \circ \Phi(g).
\]

Let \textbf{Stoch} be the category of stochastic processes.

Let \(N\) be the monoid of natural numbers, considered as a category. A \textbf{discrete Markov process} is a functor \(N \rightarrow \mathcal{P}\), i.e. a measurable space \(\Omega\) together with a Markov kernel \(f : \Omega \rightarrow \mathcal{G}\Omega\). Let \textbf{Mark} be the category of discrete Markov processes, i.e. \([N, \mathcal{P}]\).
A morphism of discrete stochastic processes from \((\Omega^1, f^1)\) to \((\Omega^2, f^2)\) is a morphism \(g : \Omega^1 \to \Omega^2\) in \(\mathcal{P}^\mathbb{N}\) such that

\[ g \circ f = f' \circ \Phi(g). \]

Let \textbf{Stoch} be the category of stochastic processes.

Let \(\mathbb{N}\) be the monoid of natural numbers, considered as a category. A \textbf{discrete Markov process} is a functor \(\mathbb{N} \to \mathcal{P}\), i.e. a measurable space \(\Omega\) together with a Markov kernel \(f : \Omega \to \mathcal{G}\Omega\). Let \textbf{Mark} be the category of discrete Markov processes, i.e. \([\mathbb{N}, \mathcal{P}]\).

Question: Does the inclusion \textbf{Mark} \(\to\) \textbf{Stoch} have adjoints?
In this paper, Giry recognizes Lawvere’s category of probabilistic mappings as the Kleisli category of a certain monad (Giry monad).
In this paper, Giry recognizes Lawvere’s category of probabilistic mappings as the Kleisli category of a certain monad (Giry monad).

- For a measurable space $\Omega$. Let $\mathcal{G}\Omega$ be the set of probability measures on $\Omega$. 
In this paper, Giry recognizes Lawvere’s category of probabilistic mappings as the Kleisli category of a certain monad (Giry monad).

- For a measurable space $\Omega$. Let $\mathcal{G}\Omega$ be the set of probability measures on $\Omega$. For a measurable subset $A \subseteq \Omega$, we have an evaluation map

  $$\text{ev}_A : \mathcal{G}\Omega \to [0, 1].$$

$\mathcal{G}\Omega$ becomes a measurable spaces by endowing it with the $\sigma$-algebra generated by the evaluation maps.
In this paper, Giry recognizes Lawvere’s category of probabilistic mappings as the Kleisli category of a certain monad (Giry monad).

- For a measurable space $\Omega$. Let $\mathcal{G}\Omega$ be the set of probability measures on $\Omega$. For a measurable subset $A \subseteq \Omega$, we have an evaluation map

$$\text{ev}_A : \mathcal{G}\Omega \to [0, 1].$$

$\mathcal{G}\Omega$ becomes a measurable spaces by endowing it with the $\sigma$-algebra generated by the evaluation maps.

For a measurable map $f : \Omega_1 \to \Omega_2$, pushing forward along $f$ defines a measurable map $\mathcal{G}f : \mathcal{G}\Omega_1 \to \mathcal{G}\Omega_2$. 
In this paper, Giry recognizes Lawvere’s category of probabilistic mappings as the Kleisli category of a certain monad (Giry monad).

- For a measurable space $\Omega$. Let $G\Omega$ be the set of probability measures on $\Omega$. For a measurable subset $A \subseteq \Omega$, we have an evaluation map

$$ev_A : G\Omega \to [0, 1].$$

$G\Omega$ becomes a measurable spaces by endowing it with the $\sigma$-algebra generated by the evaluation maps.

For a measurable map $f : \Omega_1 \to \Omega_2$, pushing forward along $f$ defines a measurable map $Gf : G\Omega_1 \to G\Omega_2$.

This gives an endofunctor $G : \text{Mble} \to \text{Mble}$.
There is a measurable map $\mu_\Omega : gG \Omega \to G \Omega$:

$$\mu_\Omega(P)(A) := \int_{\lambda \in G \Omega} \lambda(A)P(d\lambda),$$

for all $P \in gG \Omega$ and measurable subsets $A \subseteq \Omega$. 

Giry also introduces a monad on $Pol$, the category of Polish spaces and continuous functions.
There is a measurable map $\mu_\Omega : \mathcal{G}\mathcal{G}\Omega \to \mathcal{G}\Omega$:

$$\mu_\Omega(P)(A) := \int_{\lambda \in \mathcal{G}\Omega} \lambda(A)P(d\lambda),$$

for all $P \in \mathcal{G}\mathcal{G}\Omega$ and measurable subsets $A \subseteq \Omega$.

We have a map $\eta_\Omega : \Omega \to \mathcal{G}\Omega$:

$$\eta_\Omega(\omega) := \delta_\omega,$$

for all $\omega \in \Omega$. 

These form natural transformation $\mu : \mathcal{G}\mathcal{G} \to \mathcal{G}$ and $\eta : 1 \to \mathcal{G}$ and $$(\mathcal{G}, \mu, \eta)$$ forms a monad, the Giry monad.

Lawvere's category of probabilistic mappings is the Kleisli category of the Giry monad.

Giry also introduces a monad on $\mathcal{Pol}$, the category of Polish spaces and continuous functions.
There is a measurable map $\mu_\Omega : \mathcal{G}\mathcal{G}\Omega \to \mathcal{G}\Omega$:

$$\mu_\Omega(P)(A) := \int_{\lambda \in \mathcal{G}\Omega} \lambda(A)P(d\lambda),$$

for all $P \in \mathcal{G}\mathcal{G}\Omega$ and measurable subsets $A \subseteq \Omega$.

We have a map $\eta_\Omega : \Omega \to \mathcal{G}\Omega$:

$$\eta_\Omega(\omega) := \delta_\omega,$$

for all $\omega \in \Omega$.

These form natural transformation $\mu : \mathcal{G}\mathcal{G} \to \mathcal{G}$ and $\eta : 1_{\text{Mble}} \to \mathcal{G}$ and $(\mathcal{G}, \mu, \eta)$ forms a monad, the Giry monad.
There is a measurable map $\mu_\Omega : \mathcal{G}\mathcal{G}\Omega \to \mathcal{G}\Omega$:

$$\mu_\Omega(P)(A) := \int_{\lambda \in \mathcal{G}\Omega} \lambda(A)P(d\lambda),$$

for all $P \in \mathcal{G}\mathcal{G}\Omega$ and measurable subsets $A \subseteq \Omega$.

We have a map $\eta_\Omega : \Omega \to \mathcal{G}\Omega$:

$$\eta_\Omega(\omega) := \delta_\omega,$$

for all $\omega \in \Omega$.

These form natural transformation $\mu : \mathcal{G}\mathcal{G} \to \mathcal{G}$ and $\eta : 1_{\text{Mble}} \to \mathcal{G}$ and $(\mathcal{G}, \mu, \eta)$ forms a monad, the Giry monad.

Lawvere’s category of probabilistic mappings is the Kleisli category of the Giry monad.
There is a measurable map \( \mu_\Omega : \mathcal{G}\mathcal{G}\Omega \to \mathcal{G}\Omega : \)

\[
\mu_\Omega(P)(A) := \int_{\lambda \in \mathcal{G}\Omega} \lambda(A)P(d\lambda),
\]

for all \( P \in \mathcal{G}\mathcal{G}\Omega \) and measurable subsets \( A \subseteq \Omega \).

We have a map \( \eta_\Omega : \Omega \to \mathcal{G}\Omega : \)

\[
\eta_\Omega(\omega) := \delta_\omega,
\]

for all \( \omega \in \Omega \).

These form natural transformation \( \mu : \mathcal{G}\mathcal{G} \to \mathcal{G} \) and \( \eta : 1_{\text{Mble}} \to \mathcal{G} \) and \((\mathcal{G}, \mu, \eta)\) forms a monad, the Giry monad.

Lawvere’s category of probabilistic mappings is the Kleisli category of the Giry monad.

Giry also introduces a monad on \( \text{Pol} \), the category of Polish spaces and continuous functions.
**The Kolmogorov extension problem:** Let \((\Omega_i)_{i \in I}\) be a collection of measurable spaces. Consider a probability measure \(P_J\) on \(\prod_{j \in J} \Omega_j\) for every finite set \(J \subseteq I\). Suppose that this collection is *consistent*. Does there exist a probability measure \(P\) on \(\prod_{i \in I} \Omega_i\) such that

\[
P \circ \pi^{-1}_J = P_J,
\]

for all finite \(J \subseteq I\)?

**Answer:** sometimes. (*Kolmorogov extension problem*)
The Kolmogorov extension problem: Let \((\Omega_i)_{i \in I}\) be a collection of measurable spaces. Consider a probability measure \(P_J\) on \(\prod_{j \in J} \Omega_j\) for every finite set \(J \subseteq I\). Suppose that this collection is consistent. Does there exist a probability measure \(P\) on \(\prod_{i \in I} \Omega_i\) such that

\[ P \circ \pi^{-1}_J = P_J, \]

for all finite \(J \subseteq I\)?

Answer: sometimes. (Kolmorogov extension problem)

Giry translates this problem as follows:

'Does the functor \(Mble \rightarrow Mble_G\) preserve cofiltered limits?'.

Several (technical) conditions are given for which this is the case. As a corollary the Ionescu-Tulcea theorem is discussed.
Lawvere expressed *discrete* Markov processes using $\mathcal{P}(\equiv \text{Mble}_G)$. In probability theory, we also want to talk about *continuous* stochastic processes.
Lawvere expressed *discrete* Markov processes using $\mathcal{P}(=\text{Mble}_G)$. In probability theory, we also want to talk about *continuous* stochastic processes. For this, Giry introduced **random topological actions**. Let $C$ be a **category internal to Pol**. A random topological action is functor $C \to \text{Pol}_G$ satisfying certain (continuity) conditions.
Probability monads

Algebras: integration

Kleisli categories: Markov kernels

Markov categories
Swirszcz: Monadic functors and convexity (1973)

Studies which monads have the category of convex spaces as algebras.
Studies which monads have the category of convex spaces as algebras.

- The category $\text{Conv}$:
  - objects: convex subsets of vector spaces,
  - morphisms: affine maps
Studies which monads have the category of convex spaces as algebras.

- The category $\textbf{Conv}$:
  - objects: convex subsets of vector spaces,
  - morphisms: affine maps

- The forgetful functor $\textbf{Conv} \to \textbf{Set}$ is not monadic.
The category **CompConv**:

- objects: compact convex subsets of locally convex topological vector spaces,
- morphisms: affine maps

The forgetful functor \( \text{CompConv} \to \text{Comp} \) is monadic. The corresponding monad is the \( \text{Radon monad} \), that sends a compact Hausdorff space \( X \) to the space \( R^X \) of Radon probability measures on \( X \).

The algebra structure on a compact convex space \( X \) is given by the barycenter map \( b: R^X \to X \), which sends a Radon probability measure \( P \) to the unique element \( x \in X \) such that \( f(x) = \int f \, dP \) for all continuous affine maps \( f: X \to \mathbb{R} \).

The results uses a monadicity theorem (Linton).
The category **CompConv**:
- objects: compact convex subsets of locally convex topological vector spaces,
- morphisms: affine maps

The forgetful functor **CompConv** → **Comp** is monadic.

The results uses a monadicity theorem (Linton).
The category \textbf{CompConv}:
- objects: compact convex subsets of locally convex topological vector spaces,
- morphisms: affine maps

The forgetful functor \textbf{CompConv} → \textbf{Comp} is monadic.

The corresponding monad is the \textit{Radon monad}, that sends a compact Hausdorff space \( X \) to the space \( \mathcal{R}X \) of Radon probability measures on \( X \).
The category **CompConv**:
- objects: compact convex subsets of locally convex topological vector spaces,
- morphisms: affine maps

The forgetful functor **CompConv** → **Comp** is monadic.

The corresponding monad is the *Radon monad*, that sends a compact Hausdorff space $X$ to the space $\mathcal{R}X$ of Radon probability measures on $X$.

The algebra structure on a compact convex space $X$ is given by the *barycenter map* $b : \mathcal{R}X \to X$, which sends a Radon probability measure $P$ to the unique element $x \in X$ such that $f(x) = \int f \, dP$ for all continuous affine maps $f : X \to \mathbb{R}$.
The category **CompConv**:
- objects: compact convex subsets of locally convex topological vector spaces,
- morphisms: affine maps

The forgetful functor **CompConv** → **Comp** is monadic.

The corresponding monad is the Radon monad, that sends a compact Hausdorff space \( X \) to the space \( \mathcal{R}X \) of Radon probability measures on \( X \).

The algebra structure on a compact convex space \( X \) is given by the barycenter map \( b : \mathcal{R}X \to X \), which sends a Radon probability measure \( \mathbb{P} \) to the unique element \( x \in X \) such that

\[
f(x) = \int f \, d\mathbb{P}
\]

for all continuous affine maps \( f : X \to \mathbb{R} \).
The category **CompConv**:
- objects: compact convex subsets of locally convex topological vector spaces,
- morphisms: affine maps

The forgetful functor $\text{CompConv} \to \text{Comp}$ is monadic.

The corresponding monad is the *Radon monad*, that sends a compact Hausdorff space $X$ to the space $\mathcal{R}X$ of Radon probability measures on $X$.

The algebra structure on a compact convex space $X$ is given by the *barycenter map* $b : \mathcal{R}X \to X$, which sends a Radon probability measure $\mathbb{P}$ to the unique element $x \in X$ such that

$$f(x) = \int f \, d\mathbb{P}$$

for all continuous affine maps $f : X \to \mathbb{R}$.

The results uses a monadicity theorem (Linton).
The category **CompConv:**
- objects: compact convex subsets of locally convex topological vector spaces,
- morphisms: affine maps
The category \textbf{CompConv}:

- objects: compact convex subsets of locally convex topological vector spaces,
- morphisms: affine maps

The forgetful functor \textbf{CompConv} \rightarrow \textbf{Set} is monadic.
The category \textbf{CompConv}:
- objects: compact convex subsets of locally convex topological vector spaces,
- morphisms: affine maps

The forgetful functor \textbf{CompConv} \to \textbf{Set} is monadic.
The corresponding monad on \textbf{Set} sends a set $X$ to the set of Radon measures on the Čech-Stone compactification of $X$. 
Probability monads

Algebras: integration

Kleisli categories: Markov kernels

Markov categories
Since then, many variations on the Giry monad have been studied and are referred to as *probability monads*.

- **Distribution monad**: A probability monad on $\mathbf{Set}$, sending every set $X$ to $\{\text{finitely supported probability measures on } X\}$.
Since then, many variations on the Giry monad have been studied and are referred to as \textit{probability monads}.

- \textit{Distribution monad}: A probability monad on \textbf{Set}, sending every set $X$ to

  $\{\text{finitely supported probability measures on } X\}$. \\

- Monads of valuations on topological spaces and locales (Fritz, Perrone, Vickers)
Probability monads

Since then, many variations on the Giry monad have been studied and are referred to as *probability monads*.

- **Distribution monad**: A probability monad on \( \text{Set} \), sending every set \( X \) to
  \[
  \{ \text{finitely supported probability measures on } X \}.
  \]

- Monads of valuations on topological spaces and locales (Fritz, Perrone, Vickers)
- Monads of subprobabilities/stochastic relations (Panangaden)
• *Kantorovich monad* (van Breugel, Fritz, Perrone): A probabilty monad on $\text{CMet}_1$, the category of complete metric spaces and 1-Lipschitz maps. A complete metric space is send to its *Kantorovich space*
Kantorovich monad (van Breugel, Fritz, Perrone): A probability monad on $\text{CMet}_1$, the category of complete metric spaces and 1-Lipschitz maps. A complete metric space is sent to its Kantorovich space.

A Radon probability measure $\mathbb{P}$ has finite moment if

$$\int d(x, y)\mathbb{P}(dx)\mathbb{P}(dy) < \infty.$$
Kantorovich monad (van Breugel, Fritz, Perrone): A probability monad on $\text{CMet}_1$, the category of complete metric spaces and 1-Lipschitz maps. A complete metric space is send to its Kantorovich space.

A Radon probability measure $\mathbb{P}$ has finite moment if

$$\int d(x, y)\mathbb{P}(dx)\mathbb{P}(dy) < \infty.$$ 

The Kantorovich space of a complete metric space $X$ is the complete metric space of all Radon probability measure that have finite moment. The metric is given by the Wasserstein distance:

$$d_W(\mathbb{P}_1, \mathbb{P}_2) = \sup \left\{ \int f d\mathbb{P}_1 - \int f d\mathbb{P}_2 \mid f : X \to \mathbb{R} \text{ 1-Lipschitz} \right\}$$

for all Radon probability measure of finite moment $\mathbb{P}_1$ and $\mathbb{P}_2$. 
Kantorovich spaces and Wasserstein distances are important in transport theory and have many useful properties.
Kantorovich spaces and Wasserstein distances are important in transport theory and have many useful properties. E.g. the finitely supported probability measures are dense in the Kantorovich space.
Kantorovich spaces and Wasserstein distances are important in *transport theory* and have many useful properties. E.g. the finitely supported probability measures are dense in the Kantorovich space.

- Many *other metric variations*: ordered metric spaces, compact metric spaces, general Lipschitz maps, . . .
Kantorovich spaces and Wasserstein distances are important in transport theory and have many useful properties. E.g. the finitely supported probability measures are dense in the Kantorovich space.

- Many other metric variations: ordered metric spaces, compact metric spaces, general Lipschitz maps, . . . .
- Many probability monads are codensity monads of functor of probability measures on countable spaces.
What are the algebras of probability monads?
From Swirszcz, we already know that algebras of probability monads should have a convex structure and the structure map should give a barycenter or centre of mass.

**Example:** Define a map $\alpha : G[0, \infty] \to [0, \infty]$ by

$$\mathbb{P} \mapsto \int_0^\infty x\mathbb{P}(dx).$$
The algebras of probability monads

From Swirszcz, we already know that algebras of probability monads should have a \textit{convex structure} and the structure map should give a \textit{barycenter} or \textit{centre of mass}.

\textbf{Example:} Define a map \( \alpha : \mathcal{G}[0, \infty] \rightarrow [0, \infty] \) by

\[
\mathbb{P} \mapsto \int_{0}^{\infty} x \mathbb{P}(dx).
\]

\[\alpha(\delta_{x_0}) = \int x \delta_{x_0}(dx) = x_0 \text{ for all } x_0 \in [0, \infty],\]
The algebras of probability monads

From Swirszcz, we already know that algebras of probability monads should have a **convex structure** and the structure map should give a *barycenter* or *centre of mass*.

**Example:** Define a map \( \alpha : \mathcal{G}[0, \infty] \rightarrow [0, \infty] \) by

\[
\mathbb{P} \mapsto \int_0^\infty x \mathbb{P}(dx).
\]

\[\alpha(\delta_{x_0}) = \int x \delta_{x_0}(dx) = x_0 \text{ for all } x_0 \in [0, \infty],\]

\[\text{For } \mathbb{P} \in \mathcal{G}[0, \infty],\]

\[
\alpha(\mu(\mathbb{P})) = \int_0^\infty x \mu(\mathbb{P})(dx) = \int_\lambda \int_0^\infty x \lambda(dx) \mathbb{P}(d\lambda) = \int_\lambda \alpha(\lambda) \mathbb{P}(d) = \alpha(\mathbb{P} \circ \alpha^{-1})
\]
Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(X : \Omega \to [0, \infty]\) be a random variable.

\[ \mathbb{E}[X] = \alpha(\mathbb{P} \circ X^{-1}). \]
Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(X : \Omega \to [0, \infty]\) be a random variable.

\[
\mathbb{E}[X] = \alpha(\mathbb{P} \circ X^{-1}).
\]

→ Random variables should take values in *algebras of probability monads.*
Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(X : \Omega \rightarrow [0, \infty]\) be a random variable.

\[
\mathbb{E}[X] = \alpha(\mathbb{P} \circ X^{-1}).
\]

Random variables should take values in \textit{algebras of probability monads}.

\textit{Distribution monad}: The algebras are convex spaces (in the sense of Stone).

\textit{Radon monad}: The algebras are compact convex subsets of locally convex topological vector spaces. (Swirszcz)

\textit{Kantorovich monad}: The algebras are closed convex subsets of Banach spaces. (Perrone, Fritz)

\textit{Giry monad}: more difficult! (Dobberkat, Sturtz)

Ruben Van Belle
An overview of categorical probability theory
September 2022 29 / 40
Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(X : \Omega \to [0, \infty]\) be a random variable.

\[
\mathbb{E}[X] = \alpha(\mathbb{P} \circ X^{-1}).
\]

Random variables should take values in **algebras of probability monads**.

*Distribution monad:* The algebras are convex spaces (in the sense of Stone).

- Given a convex space \(X\), we can define \(\alpha \left( \sum_{n=1}^{N} \alpha_n \delta_{x_n} \right) = \sum_{n=1} \alpha_n x_n\).

*Radon monad:* The algebras are compact convex subsets of locally convex topological vector spaces. (Swirszcz)

*Kantorovich monad:* The algebras are closed convex subsets of Banach spaces. (Perrone, Fritz)

*Giry monad:* more difficult! (Dobberkat, Sturtz)
Let \((\Omega, \mathcal{F}, P)\) be a probability space and let \(X : \Omega \to [0, \infty]\) be a random variable.

\[
\mathbb{E}[X] = \alpha(P \circ X^{-1}).
\]

→ Random variables should take values in \textit{algebras of probability monads}.

\textit{Distribution monad}: The algebras are convex spaces (in the sense of Stone).

- Given a convex space \(X\), we can define \(\alpha(P^N n=1 \alpha_n \delta_{x_n}) = \sum_{n=1}^N \alpha_n x_n\).
- Give an algebra \(A\), we define a convex structure by \(\lambda a + (1 - \lambda)b := \alpha(\lambda \delta_a + (1 - \lambda)\delta_b)\).
Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(X : \Omega \to [0, \infty]\) be a random variable.

\[
\mathbb{E}[X] = \alpha(\mathbb{P} \circ X^{-1}).
\]

→ Random variables should take values in \textit{algebras of probability monads}.

- **Distribution monad**: The algebras are convex spaces (in the sense of Stone).
  - Given a convex space \(X\), we can define \(\alpha \left( \sum_{n=1}^{N} \alpha_n \delta_{x_n} \right) = \sum_{n=1} \alpha_n x_n\).
  - Give an algebra \(A\), we define a convex structure by \(\lambda a + (1 - \lambda) b := \alpha(\lambda \delta_a + (1 - \lambda) \delta_b)\).

- **Radon monad**: The algebras are compact convex subsets of locally convex topological vector spaces. (Swirszcz)
Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(X : \Omega \rightarrow [0, \infty]\) be a random variable.

\[ \mathbb{E}[X] = \alpha(\mathbb{P} \circ X^{-1}). \]

Random variables should take values in \textit{algebras of probability monads}.

- **Distribution monad**: The algebras are convex spaces (in the sense of Stone).
  - Given a convex space \(X\), we can define \(\alpha\left(\sum_{n=1}^{N} \alpha_n \delta_{x_n}\right) = \sum_{n=1}^{N} \alpha_n x_n\).
  - Give an algebra \(A\), we define a convex structure by \(\lambda a + (1 - \lambda)b := \alpha(\lambda \delta_a + (1 - \lambda)\delta_b)\).

- **Radon monad**: The algebras are compact convex subsets of locally convex topological vector spaces. (Swirszcz)

- **Kantorovich monad**: The algebras are closed convex subsets of Banach spaces. (Perrone, Fritz)

Ruben Van Belle
An overview of categorical probability theory
September 2022 29 / 40
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $X : \Omega \to [0, \infty]$ be a random variable.

$$\mathbb{E}[X] = \alpha(\mathbb{P} \circ X^{-1}).$$

→ Random variables should take values in *algebras of probability monads*.

**Distribution monad:** The algebras are convex spaces (in the sense of Stone).

- Given a convex space $X$, we can define $\alpha \left( \sum_{n=1}^{N} \alpha_n \delta_{x_n} \right) = \sum_{n=1} \alpha_n x_n$.
- Given an algebra $A$, we define a convex structure by $\lambda a + (1 - \lambda) b := \alpha(\lambda \delta_a + (1 - \lambda) \delta_b)$.

**Radon monad:** The algebras are compact convex subsets of locally convex topological vector spaces. (Swirszcz)

**Kantorovich monad:** The algebras are closed convex subsets of Banach spaces. (Perrone, Fritz)

**Giry monad:** more difficult! (Dobberkat, Sturtz)
**Conditional expectation:** Let $X : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}$ be an integrable random variable and let $\mathcal{G} \subseteq \mathcal{F}$ be a sub-$\sigma$-algebra.
**Conditional expectation:** Let $X : (\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ be an integrable random variable and let $\mathcal{G} \subseteq \mathcal{F}$ be a sub-$\sigma$-algebra.

$$\mathbb{E}[X \mid \mathcal{G}] : (\Omega, \mathcal{G}, \mathbb{P} \mid \mathcal{G}) \to \mathbb{R}$$

is the $\mathbb{P}$-almost surely unique random variable such that

$$\mathbb{E}[\mathbb{E}[X \mid \mathcal{G}]1_A] = \mathbb{E}[X1_A],$$

for all $A \in \mathcal{G}$. 

---

**Perrone: partial evaluations**

- Conditional expectation: Let $X : (\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ be an integrable random variable and let $\mathcal{G} \subseteq \mathcal{F}$ be a sub-$\sigma$-algebra.

$$\mathbb{E}[X \mid \mathcal{G}] : (\Omega, \mathcal{G}, \mathbb{P} \mid \mathcal{G}) \to \mathbb{R}$$

is the $\mathbb{P}$-almost surely unique random variable such that

$$\mathbb{E}[\mathbb{E}[X \mid \mathcal{G}]1_A] = \mathbb{E}[X1_A],$$

for all $A \in \mathcal{G}$. 

Conditional expectation: Let $X : (\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ be an integrable random variable and let $\mathcal{G} \subseteq \mathcal{F}$ be a sub-$\sigma$-algebra.

$$\mathbb{E}[X \mid \mathcal{G}] : (\Omega, \mathcal{G}, \mathbb{P} \mid _{\mathcal{G}}) \to \mathbb{R}$$

is the $\mathbb{P}$-almost surely unique random variable such that

$$\mathbb{E}[\mathbb{E}[X \mid \mathcal{G}]1_A] = \mathbb{E}[X1_A],$$

for all $A \in \mathcal{G}$.

Let $(A, \alpha)$ be an algebra of a probability monad $T$ and let $\mathbb{P}_1$ and $\mathbb{P}_2$ be in $TA$. Then $\mathbb{P}_2$ is a partial evaluation of $\mathbb{P}_1$ if there exists $\mathbb{P} \in TTA$ such that

$$\mathbb{P} \circ \alpha^{-1} = \mathbb{P}_1 \quad \text{and} \quad \mu_A(\mathbb{P}) = \mathbb{P}_2.$$
Example: $P_{|G} \circ E[X \mid G]^{-1}$ is a partial evaluation of $P \circ X^{-1}$.
Example: $\mathbb{P} |_G \circ \mathbb{E}[X | G]^{-1}$ is a partial evaluation of $\mathbb{P} \circ X^{-1}$.

Moreover, every $\mathbb{P}_1, \mathbb{P}_2$ in $TA$ such that $\mathbb{P}_2$ is a partial evaluation of $\mathbb{P}_1$ are of this form for some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, sub-$\sigma$-algebra $G \subseteq \mathcal{F}$ and random variable $X : (\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$. (Perrone)
Kleisli categories of probability monads and Markov categories
Markov categories are monoidal categories, similar to Kleisli categories of probability monads.
Markov categories are monoidal categories, similar to Kleisli categories of probability monads. For measurable spaces $\Omega_1$ and $\Omega_2$, we have morphisms

$$\mathcal{G}\Omega_1 \times \mathcal{G}\Omega_2 \to \mathcal{G}(\Omega_1 \times \Omega_2)$$

by sending $(\mathbb{P}_1, \mathbb{P}_2)$ to $\mathbb{P}_1 \otimes \mathbb{P}_2$. 

This makes the Giry monad into a commutative monad. Therefore the Kleisli category of the probability monad inherits the symmetric (semicartesian) monoidal structure on $\text{Mble}$. There are similar constructions for other probability monads.
Markov categories are monoidal categories, similar to Kleisli categories of probability monads. For measurable spaces $\Omega_1$ and $\Omega_2$, we have morphisms

$$G\Omega_1 \times G\Omega_2 \to G(\Omega_1 \times \Omega_2)$$

by sending $(P_1, P_2)$ to $P_1 \otimes P_2$. This makes the Giry monad into a commutative monad.

Therefore the Kleisli category of the probability monad inherits the symmetric (semicartesian) monoidal structure on $\text{Mble}$. 
Markov categories are monoidal categories, similar to Kleisli categories of probability monads. For measurable spaces $\Omega_1$ and $\Omega_2$, we have morphisms

$$G\Omega_1 \times G\Omega_2 \to G(\Omega_1 \times \Omega_2)$$

by sending $(P_1, P_2)$ to $P_1 \otimes P_2$. This makes the Giry monad into a commutative monad. Therefore the Kleisli category of the probability monad inherits the symmetric (semicartesian) monoidal structure on $\text{Mble}$. There are similar constructions for other probability monads.
Let $\Omega$ be a measurable space. Every object has a canonical *commutative comonoid* structure:

- The *comultiplication* $\Omega \to \mathcal{G}(\Omega \times \Omega)$ is defined by
  $$\omega \mapsto \delta_{(\omega,\omega)}.$$

- The *conuit* is the unique map $X \to \mathcal{G}1$. 

Ruben Van Belle
An overview of categorical probability theory
September 2022 34 / 40
Let $\Omega$ be a measurable space. Every objects has a canonical \textit{commutative comonoid} structure:

- The \textit{comultiplication} $\Omega \to G(\Omega \times \Omega)$ is defined by
  $$\omega \mapsto \delta_{(\omega,\omega)}.$$
- The \textit{conuit} is the unique map $X \to G1$.

A \textbf{Markov category} is a symmetric monoidal category with a comultiplication structure on every object,
Let $\Omega$ be a measurable space. Every objects has a canonical *commutative comonoid* structure:

- The *comultiplication* $\Omega \to \mathcal{G}(\Omega \times \Omega)$ is defined by
  \[
  \omega \mapsto \delta_{(\omega,\omega)}.
  \]

- The *conuit* is the unique map $X \to \mathcal{G}1$.

A *Markov category* is a symmetric monoidal category with a comultiplication structure on every object, such that the counits form a natural transformation.
A lot has been written about Markov categories, including versions of the
• de Finetti theorem,
• certain 0-1 laws,
• the ergodic decomposition theorem
in Markov categories (Fritz, Perrone,Moss, . . ).
A lot has been written about Markov categories, including versions of the
- de Finetti theorem,
- certain 0-1 laws,
- the ergodic decomposition theorem
in Markov categories (Fritz, Perrone, Moss, . . .).
We will look at some important definitions in concepts in Markov categories:
- deterministic morphism,
- almost surely equal morphisms
- conditionals
- Kolmogorov products
Let $C$ be a Markov category:

- A morphism $f : A \to X$ is **deterministic** if

\[
\begin{align*}
X & \xrightarrow{f} X \\
A & \xleftarrow{f} A
\end{align*}
\]

\[
= \\
\begin{align*}
X & \xrightarrow{f} X \\
A & \xleftarrow{f} A
\end{align*}
\]
Let $C$ be a Markov category:

- A morphism $f : A \to X$ is **deterministic** if

  $$f(a)(A) f(a)(B) = f(a)(A \cap B)$$

  for all $a \in A$ and measurable $A, B \subseteq X$. 

For $\text{Mble}_G$, this means

$$f(a)(A) f(a)(B) = f(a)(A \cap B)$$

for all $a \in A$ and measurable $A, B \subseteq X$. 
Let $\mathcal{C}$ be a Markov category:

- A morphism $f : A \to X$ is **deterministic** if

$$f(a)(A)f(a)(B) = f(a)(A \cap B)$$

for all $a \in A$ and measurable $A, B \subseteq X$. In particular,

$$f(a)(A)^2 = f(a)(A)$$

and therefore $f(a)(A) \in \{0, 1\}$.

- A **representable Markov category** is a Markov category that is the Kleisli category of some monad on $\mathcal{C}_{\text{det}}$, the category with the same objects as $\mathcal{C}$ and **deterministic** morphisms.
Let $m : I \to X$ and $f, g : A \to X$ in $C$, then $f$ and $g$ are $m$-almost surely equal if
Let $m : I \to X$ and $f, g : A \to X$ in $C$, then $f$ and $g$ are $m$-almost surely equal if

$$
\int_B f(a)(A)m(da) = \int_B g(a)(A)m(da)
$$

for all $a \in A$ and measurable $A, B \subseteq X$. Therefore,

$$
f(\cdot)(A) = g(\cdot)(A) \quad m - \text{almost surely.}
$$
Conditionals in Markov categories

A Markov category **has conditionals** if for every \( f : A \to X \otimes Y \), there exists \( f \mid_X : X \otimes A \to Y \) such that

\[
\begin{array}{c}
X \otimes Y \\
\begin{array}{c}
\text{f} \\
\text{A}
\end{array}
\end{array} =
\begin{array}{c}
X \\
\begin{array}{c}
\text{f} \\
\text{f}_X
\end{array}
\end{array}
\begin{array}{c}
Y \\
\begin{array}{c}
\text{A}
\end{array}
\end{array}
\]

Ruben Van Belle
An overview of categorical probability theory
September 2022 38 / 40
A Markov category has conditionals if for every \( f : A \to X \otimes Y \), there exists \( f |_X : X \otimes A \to Y \) such that

\[
\begin{array}{c}
\text{Mble}_G \text{ does not have conditionals, since not every } \mathbb{P} \text{ on } \Omega_1 \times \Omega_2 \text{ is of the form }\\
\quad A \times B \mapsto \int_B f(\omega)(B)\mathbb{P}(d\omega)
\end{array}
\]

for some Markov kernel \( f : \Omega_1 \to G\Omega_2 \).
Conditionals in Markov categories

A Markov category **has conditionals** if for every \( f : A \to X \otimes Y \), there exists \( f_{|X} : X \otimes A \to Y \) such that

\[
\begin{array}{c}
X \\
\downarrow f \\
A
\end{array}
= 
\begin{array}{c}
X \\
\downarrow f_{|X} \\
Y \\
\downarrow f \\
A
\end{array}
\]

\( \text{Mble}_G \) does **not** have conditionals, since not every \( \mathbb{P} \) on \( \Omega_1 \times \Omega_2 \) is of the form

\[
A \times B \mapsto \int_B f(\omega)(B)\mathbb{P}(d\omega)
\]

for some Markov kernel \( f : \Omega_1 \to \mathcal{G}\Omega_2 \).

**Standard Borel spaces** do have conditionals.
Let $C$ be a Markov category and let $(X_i)_{i \in I}$ be a collection of objects.
Let $\mathcal{C}$ be a Markov category and let $(X_i)_{i \in I}$ be a collection of objects. Suppose that

- The limit $\lim_{J \subseteq I} \bigotimes_{j \in J} X_j$ over all finite subsets $J$ of $I$ exists,

Then $\lim_{J \subseteq I} \bigotimes_{j \in J} X_j$ is called the Kolmogorov product of $(X_i)_{i \in I}$. In $\textit{Mble} \mathcal{G}$, Kolmogorov products don't exist in general, but countable Kolmogorov products of Standard Borel spaces do exist (Kolmogorov extension theorem).
Kolmogorov products in Markov categories

Let $C$ be a Markov category and let $(X_i)_{i \in I}$ be a collection of objects. Suppose that

- The limit $\lim_{J \subseteq I} \bigotimes_{j \in J} X_j$ over all finite subsets $J$ of $I$ exists,
- and that it is preserved by $-Y$ for all objects $Y$,

is called the Kolmogorov product of $(X_i)_{i \in I}$.

In $\mathcal{Mbl}$ categories, Kolmogorov products don't exist in general, but countable Kolmogorov products of Standard Borel spaces do exist (Kolmogorov extension theorem).

Kolmogorov products in Markov categories

Let $C$ be a Markov category and let $(X_i)_{i \in I}$ be a collection of objects. Suppose that

- The limit $\lim_{J \subseteq I} \bigotimes_{j \in J} X_j$ over all finite subsets $J$ of $I$ exists,
- and that it is preserved by $-Y$ for all objects $Y$,
- and that

$$\pi_{J' \subseteq J} : \bigotimes_{j \in J} X_j \to \bigotimes_{j \in J'} X_j$$

is deterministic for all finite sets $J' \subseteq J$ of $I$.

Then $\lim_{J \subseteq I} \bigotimes_{j \in J} X_j$ is called the **Kolmogorov product** of $(X_i)_{i \in I}$. 
Kolmogorov products in Markov categories

Let $C$ be a Markov category and let $(X_i)_{i \in I}$ be a collection of objects. Suppose that

- The limit $\lim_{J \subseteq I} \bigotimes_{j \in J} X_j$ over all finite subsets $J$ of $I$ exists,
- and that it is preserved by $-Y$ for all objects $Y$,
- and that

$$\pi_{J' \subseteq J} : \bigotimes_{j \in J} X_j \to \bigotimes_{j \in J'} X_j$$

is deterministic for all finite sets $J' \subseteq J$ of $I$.

Then $\lim_{J \subseteq I} \bigotimes_{j \in J} X_j$ is called the **Kolmogorov product** of $(X_i)_{i \in I}$. In $\mathbf{Mble}_g$ Kolmogorov products don’t exist in general, but countable Kolmogorov products of Standard Borel spaces do exist (**Kolmogorov extension theorem**).
Probability monads

Algebras: integration

Kleisli categories: Markov kernels

Markov categories