Central Limits via Dilated Categories

112th Peripatetic Seminar on Sheaves and Logic
University of Nottingham

Oisín Flynn-Connolly (Leiden)
Joint work with H. Basold, C. Ford and H. Wang (all Leiden)

March 20, 2026

Categorical probability theory

There are many (really nice!) categorical approaches to probability theory

1. Probability monads (Giry '82)
2. Markov categories (Fritz '20)
    1. Very minimalistic.
    2. Captures Kleisli category of probability monads.
    3. Captures analogues of many structural theorems from probability. (De Finetti's theorem, ergodic decomposition theorem, entropy, Kolmogorov extension theorem...)
3. Effectus theory for quantum logic (Jacobs '15).
4. Many others (dagger categories, quasi-Borel spaces... Choose your favourite formalism!)

Many of these (that I'm aware of!) excel at capturing structural properties but tend to struggle a bit with analytic ones.

The central limit theorem

Let $\{X_1, X_2, \ldots X_n\}$ be a sequence of **independent, identically distributed** random variables.

Theorem (The Central Limit Theorem)

Suppose the variables have expectation 0 and variance σ. The sample average

$$\bar{X} := \frac{X_1 + \cdots + X_n}{\sqrt{n}}$$

converges in probability to $N(0, \sigma)$ as $n \to \infty$.

The suitably normalized sampling distribution approaches a normal distribution, regardless of the original distribution shape.

Ubiquitous in everyday life.

The central limit theorem

Let $\{X_1, X_2, \ldots X_n\}$ be a sequence of **independent, identically distributed** random variables.

Theorem (The Central Limit Theorem)

Suppose the variables have expectation 0 and variance σ. The sample average

$$\bar{X} := \frac{X_1 + \cdots + X_n}{\sqrt{n}}$$

converges in probability to $N(0, \sigma)$ as $n \to \infty$.

The suitably normalized sampling distribution approaches a normal distribution, regardless of the original distribution shape.

Ubiquitous in everyday life. (*at least if you are a professional poker player)

- Our arguments in this talk also recover the other big limiting theorem: the Law of Large Numbers.

Main idea

Slogan: Limiting theorems are best captured via enrichment.

Categorifying probability

Probability functors

Natural transformations & enriched fibres

Theorem (The Central Limit Theorem)

Suppose the variables/probability measures have expectation 0 *and variance* σ. *The sample average*

$$\bar{X} := \frac{X_1 + \cdots + X_n}{\sqrt{n}}$$

converges in probability to $N(0, \sigma)$ *as* $n \to \infty$.

Lax monoidal structure

Seminorm enrichment

Banach fixed point theorem

Dilated categories

Probabilistic setup I

1. There is a functor $\mathcal{P} : \mathbf{FinVect} \to \mathbf{CMet}$
2. It assigns to every vector space V the set of probability distributions $\mathcal{P}(V)$ on it.
3. This is equipped with an extended metric structure. (choice!)
4. The functor $\mathcal{P}$ is **lax monoidal**.
5. The coherence map $\mu : \mathcal{P}(V) \times \mathcal{P}(W) \to \mathcal{P}(V \times W)$ takes two probability measures on V and on W and constructs the **independent** joint probability distribution on $V \times W$.
6. Later, we shall see that all of this can be enriched.

Probabilistic setup II

1. There is a pointwise addition on vector spaces $+ : V \times V \to V$.
2. The map $\mathcal{P}(+)$ is precisely the **convolution:**

$$\mathcal{P}(+)(\mu_1, \mu_2)(U) := \int_{(x,y)\in V\times V} \chi_U(x+y)d\mu_1(x)d\mu_2(y).$$

3. The $\mathcal{P}$-**convolution operator**

$$\mathcal{P}(V) \xrightarrow{\Delta} \mathcal{P}(V) \times \mathcal{P}(V) \xrightarrow{\mu} \mathcal{P}(V \times V) \xrightarrow{\mathcal{P}(+)} \mathcal{P}(V)$$

takes a probability distribution, duplicates it, and then convolves two independent copies.

4. Consider the functor: $|-| :$ **FinVect** $\to$ **CMet**. All the $|-|$-convolution operator just multiplies by two.

Probabilistic setup II

1. There is a pointwise addition on vector spaces $+ : V \times V \to V$.
2. The map $\mathcal{P}(+)$ is precisely the **convolution:**

$$\mathcal{P}(+)(\mu_1, \mu_2)(U) := \int_{(x,y)\in V\times V} \chi_U(x+y)d\mu_1(x)d\mu_2(y).$$

3. The $\mathcal{P}$**-convolution operator**

$$\mathcal{P}(V) \xrightarrow{\Delta} \mathcal{P}(V) \times \mathcal{P}(V) \xrightarrow{\mu} \mathcal{P}(V \times V) \xrightarrow{\mathcal{P}(+)} \mathcal{P}(V)$$

   takes a probability distribution, duplicates it, and then convolves two independent copies.
4. Consider the functor: $|-| :$ **FinVect** $\to$ **CMet**. All the $|-|$-convolution operator just multiplies by two. (this will make sense later!)

Towards the Banach fixed point theorem

Quantales

To capture examples like quantum central limit theorems or Markov processes (future work!), we work with quantales.

A **quantale** $(\mathcal{V}, \otimes, e)$ is a complete lattice $\mathcal{V}$ equipped with the structure of a commutative monoid with **tensor** $\otimes$ and **unit** e such that for each $x \in \mathcal{V}$ the functor $x \otimes - : \mathcal{V} \to \mathcal{V}$ preserves joins ie.

$$r \otimes \left(\bigvee s_i \right) = \bigvee (r \otimes s_i).$$

We call $\mathcal{V}$ **contractive** if $q \otimes x < x$ when $q < e$ and $x \neq \top, \bot$.

Example

The set of (extended) postive real numbers $[0, \infty]$ is a contractive quantale with $\otimes$ given by extended multiplication and the unit e is 1.

Distance spaces over quantales

A $\mathcal{V}$-**space** (X, d) is a set X is a function $d : X \times X \to \mathcal{V}$ such that

$$d(x,x) = \bot \quad \text{and} \quad d(x,y) = d(y,x) \quad \text{for all } x, y \in X$$

and such that limit points of sequences are unique. These form a monoidal category with nonexpansive maps and the maximum metric.
A sequence $x : \omega \to X$ is **geometric** if for some $r < e$, $q < \top$ and for all $i, j \in \mathbb{N}$

$$d(x_i, x_{i+1}) \leq r^i \otimes q, \qquad d(x_i, x_j) < \top$$

We say X is **complete** if all geometric series with $r < e$ converge to a limit which satisfies some conditions.

Definition

The **Lipschitz constant** $|f|_{\mathrm{L}}$ of a nonexpansive map $f : X \to Y$ as the greastest lower bound (meet) of all $c \in \mathcal{V}$ such that, for all $x, y \in X$,

$$d(f(x), f(y)) \leq c \otimes d(x, y).$$

Banach fixed point theorem for $\mathcal{V}$-spaces

A $\mathcal{V}$-space X is said to be **metrically small** if $d(x,y) < \top$ for all $x, y \in X$.

Theorem

Let X be an inhabited, metrically small, complete $\mathcal{V}$-space and $f : X \to X$ a nonexpansive map such that $|f|_{\mathsf{L}} < e$. Then f has a unique fixed point $\mathrm{fix}(f)$, i.e. $f(\mathrm{fix}(f)) = \mathrm{fix}(f)$. Moreover, for any $x \in X$, $\mathrm{fix}(f)$ is characterised by

$$\mathrm{fix}(f) = \lim(f^n(x))_{n\in\mathbb{N}}.$$

Bases of enrichment

Definition

A **seminorm space** is a complete $\mathcal{V}$-space (A, d) equipped with map $|-|\colon A \to \mathcal{V}$ (the *seminorm*). A morphism of seminorm spaces is a nonexpansive map $f\colon X \to Y$ such that $|(f(x))|_Y \leq |x|_X$. We call this category **sNorm**.

The **unit interval monoid** of a quantale $\mathcal{V}$ is the submonoid of $(\mathcal{V}, \otimes, e)$ on

$$\mathcal{V}_{\leq e} = \{x \in \mathcal{V} \mid x \leq e\}.$$

Given an appropriate choice of $\mathcal{V}$-space structure $d_{\mathcal{V}_{\leq e}}$ on it, $\mathcal{V}_{\leq e} \wedge -$ is a monad on $\mathbf{sNorm}_{\bullet}$. The category of EM-algebras is called **dNorm**.

Theorem

sNorm *and* **dNorm** *are cosmoi; ie. complete, cocomplete and monoidal closed.*

Enrichment

We build the Lipschitz property into the definition of the enrichment. For suitably non-trivial objects:

$$d(f_* g_1, f_* g_2) \leq |f| \cdot d(g_1, g_2) \quad \text{resp.} \quad d(f^* g_1, f^* g_2) \leq |f| \cdot d(g_1, g_2).$$

We say categories enriched over **sNorm** are seminorm categories and categories enriched over **dNorm** are dilated categories.

Theorem

Any **Ban***-enriched category* $\underline{\mathscr{C}}$ *is a dilated category. So* **Hilb**, **Ban**, **FinVect** *are dilated categories.*

Theorem (Main examples!)

The category $\underline{\mathbf{CMet}}$ *is a seminorm category and the category of algebras* $\underline{\mathbb{R}\mathbf{CMet}}$ *over the monad* $[0,1] \wedge (-)$ *on* $\underline{\mathbf{CMet}_\bullet}$ *is a dilated category.*

Categorical Banach fixed point theorem

Seminorm categories have a lot of metric structure! Seminorms of identity maps of sufficently nontrivial objects can be computed to be e for example. Below, $\mathbf{1}_{\underline{\mathscr{C}}}$ is the terminal object.

Theorem

A category is Lipschitz if $|\mathrm{id}_{\mathbf{1}_{\underline{\mathscr{C}}}}| = \bot$. Let $X \in \underline{\mathscr{C}}$ be a metrically small object in a Lipschitz seminorm category such that $\mathscr{C}$ is well-pointed. Then for all $Y \in \underline{\mathscr{C}}$, a morphism $f : X \to Y$ is constant if and only if it has seminorm $\bot$.

Theorem

Let $J \in \underline{\mathscr{C}}$ be metrically small object in a seminorm category such that $\underline{\mathscr{C}}(\mathbf{1}_{\underline{\mathscr{C}}}, J)$ is nonempty. Consider the left composition operator $f_* : \underline{\mathscr{C}}(\mathbf{1}_{\underline{\mathscr{C}}}, J) \to \underline{\mathscr{C}}(\mathbf{1}_{\underline{\mathscr{C}}}, J)$, where $f \in \underline{\mathscr{C}}(J, J)$ is such that $|f|_{\mathsf{L}} < e$. Then f_* has a unique fixed point $\mathrm{fix}(f) \in \underline{\mathscr{C}}(\mathbf{1}_{\underline{\mathscr{C}}}, J)$. Moreover, for any $x \in \underline{\mathscr{C}}(\mathbf{1}_{\underline{\mathscr{C}}}, J)$, $\mathrm{fix}(f)$ is characterised by

$$\mathrm{fix}(f) = \lim(f^n(x))_{n \in \mathbb{N}}.$$

Back to probability

Gradings of functors

1. Two enriched lax monoidal functor: $\underline{F}$ (probability) and $\underline{G}$ (variance)
2. A natural transformation $p\colon \underline{F} \to \underline{G}$ (pointwise epi).
3. Commutes with pointwise addition and lax monoidal maps:

$$\begin{array}{ccc} \underline{F}_{lin}(X) \times \underline{F}_{lin}(Y) & \xrightarrow{p_X \times p_Y} & \underline{G}_{lin}(X) \times \underline{G}_{lin}(Y) \\ \downarrow{\scriptstyle \mu(F)_{X,Y}} & & \downarrow{\scriptstyle \mu(G)_{X,Y}} \\ \underline{F}_{lin}(X \times Y) & \xrightarrow{p_{X \times Y}} & \underline{G}_{lin}(X \times Y) \end{array}$$

$$\begin{array}{ccc} \underline{F}_{lin}(X \times X) & \xrightarrow{p_{X \times X}} & \underline{G}_{lin}(X \times X) \\ \downarrow{\scriptstyle \underline{F}(+)} & & \downarrow{\scriptstyle \underline{G}(+)} \\ \underline{F}_{lin}(X) & \xrightarrow{p_X} & \underline{G}_{lin}(X). \end{array}$$

Example

The functor $\underline{\mathcal{P}^0}$ is graded with respect the functor $\underline{\Sigma^+}$ that assigns to every vector space the set of positive definite matrices on it. The grading is given by the map $\mathrm{Var}_V : \mathcal{P}(V) \to \underline{\Sigma^+}(V)$ which assigns its variance matrix to every probability measure in $\mathcal{P}(V)$. These functors are **enriched lax monoidal functors**.

CLT-systems

Recall the convolution operator:

$$F(V) \xrightarrow{\Delta} F(V) \times F(V) \xrightarrow{\mu} F(V \times V) \xrightarrow{F(+)} F(V)$$

The convolution operator associated to $\underline{\Sigma^+}$ is very nice. It multiplies the identity map by $\sqrt{2}$. We call this property **perfectly rescalability.**

Definition

A *(Cartesian) pre-CLT system* $(\underline{F}, \underline{G}, p)$ is a grading where $\underline{G}$ is perfectly rescalable.

Structural CLT I

Theorem (Structural CLT I)

Given a pre-CLT system $(\underline{F}, \underline{G}, p)$, the diagram below commutes

$$\begin{array}{ccc} \underline{F}X & \xrightarrow{[c,e]\star\vartheta^{F}_{X}} & \underline{F}X \\ & {}_{p_X}\searrow \quad \swarrow_{p_X} & \\ & \underline{G}X & \end{array}$$

Structural CLT II

We can look at the fibre:

$$\begin{array}{ccc} \underline{F}X_p & \xrightarrow{i_p} & \underline{F}X \\ \downarrow{\scriptstyle t} & & \downarrow{\scriptstyle p_X} \\ \mathbf{1}_{\underline{\mathscr{D}}} & \xrightarrow{p} & \underline{G}X \end{array}$$

Theorem (Structural CLT II)

The morphism $[c, e] \star \vartheta_X^F$ *induces an endomorphism* $\vartheta_p \in \mathfrak{S}_*\underline{\mathscr{D}}\,(\underline{F}(X)_p, \underline{F}(X)_p)$.

Structural CLT II

We can look at the fibre:

$$\begin{array}{ccc} \underline{F}X_p & \xrightarrow{i_p} & \underline{F}X \\ \downarrow t & & \downarrow p_X \\ \mathbf{1}_{\underline{\mathscr{D}}} & \xrightarrow{p} & \underline{G}X \end{array}$$

Theorem (Structural CLT II)

The morphism $[c, e] \star \vartheta_X^F$ *induces an endomorphism* $\vartheta_p \in \mathfrak{S}_* \underline{\mathscr{D}}(\underline{F}(X)_p, \underline{F}(X)_p)$.

Remind you of anything?

CLT-systems

Definition

A *pre-CLT system* $(\underline{F}, \underline{G}, p)$ is a CLT system if

1. Each fibre is metrically small.
2. $|\vartheta_p| < e$.

Theorem (Categorical Central Limit Theorem)

Suppose $(\underline{F}, \underline{G}, p)$ *is a CLT system. Then for every generalised point* $p\colon 1 \to \underline{G}(X)$ *in* $\mathfrak{S}_*\underline{\mathscr{D}}$, *the convolution operator* ϑ_p *on the fibre* $\underline{F}(X)_p$ *has a unique fixed point* $\mathcal{N}_p \in \underline{\mathscr{D}}(1, \underline{F}(X)_p)$ *which we call the* **central limit**. *For any initial point* $x_0 \in \underline{\mathscr{D}}(1, \underline{F}(X)_p)$, *this fixed point is given by the limit:*

$$\mathcal{N}_p = \lim_{n\to\infty} (\vartheta_p)^{\circ n}(x_0).$$

Limiting theorems

Theorem (Probabilistic limiting theorems)

1. *Let μ be a probability measure on a finite vector space V with $\int \|x\|^\lambda d\mu < \infty$ for some $\lambda > 1$. Then $\frac{1}{2^n}\mu^{*2^n}$ converges as $n \to \infty$ and the limit depends only on the expectation of μ.*
2. *Let μ be a probability measure on a finite vector space V with expected value 0 and $\int \|x\|^\lambda d\mu < \infty$ for some $\lambda > 2$. Then $\frac{1}{\sqrt{2^n}}\mu^{*2^n}$ converges as $n \to \infty$ and the limit depends uniquely on the variance matrix of μ.*

Extensions

Theorem (Functoriality of the central limit)

In a Lipschitz category $\underline{\mathscr{D}}$, the map

$$\eta_X \colon \mathscr{D}_{lin}\left(\mathbf{1}_{\underline{\mathscr{D}}}, \underline{G}(X)\right) \to \mathscr{D}_{lin}\left(\mathbf{1}_{\underline{\mathscr{D}}}, \underline{F}(X)\right) \quad p \mapsto i_p \circ \mathcal{N}_p$$

where $i_p \colon \underline{F}(X)_p \to \underline{F}(X)$ is the canonical inclusion, defines a natural transformation in **Set**

$$\eta \colon \mathscr{D}_{lin}\left(\mathbf{1}_{\underline{\mathscr{D}}}, \underline{G}(-)\right) \to \mathscr{D}_{lin}\left(\mathbf{1}_{\underline{\mathscr{D}}}, \underline{F}(-)\right)$$

which we call the **central limit natural transformation in Set**.

Example

The image of the Gaussian distribution $\mathcal{N}(0, M)$ along $\mathcal{P}(f)$ will be the Gaussian distribution $\mathcal{N}(0, fMf^T)$.

Applications: the CLT for observables

1. A **classical Hamiltonian system** is a triple (M, λ_ω, H), where:
   1. M is a compact symplectic manifold.
   2. its Liouville measure λ_ω
   3. $H\colon M \to \mathbb{R}$ is the Hamiltonian (energy measure!)
2. Consider an *ensemble* of 2^N non-interacting, identical systems. The normalised energy of the ensemble is $\frac{1}{2^{n/2}} H_{tot}(x_1, \ldots, x_{2^N}) = \sum_{i=1}^{2^N} H(x_i)$.
3. We assume the pushforward of the initial energy distribution has mean zero and finite variance.
4. A **higher order** version of our theorem shows that the normalisation of the total energy converges to the fixed point of the CLT system as $N \to \infty$. (Not terribly surprising!)
5. The general theorem says that functions into a category with a CLT also satisfy a CLT.

Categorifying probability

Probability functors

Natural transformations & enriched fibres

Theorem (The Central Limit Theorem)

Suppose the probability measures have expectation 0 *and variance* σ. *The sample average*

$$\bar{X} := \frac{X_1 + \cdots + X_n}{\sqrt{n}}$$

converges in probability to $N(0, \sigma)$ *as* $n \to \infty$.

Lax monoidal structure

Seminorm enrichment

Banach fixed point theorem

Dilated categories