The Formal Theory of Random Variables

Paolo Perrone
University of Oxford

Probability theory is, to a large extent, a quantitative formalism. *Numbers*, as opposed to points of an arbitrary space, play a particularly prominent role in probability, because:

- First of all, *probability itself* is modeled as a number, between 0 and 1;
- The most common and traditional applications of probability theory are about predicting random numbers, such as prices or outcomes of noisy measurements.

When we deal with random numbers, as opposed to random points of an arbitrary space, we can *average them over our uncertainty.* This is one of the most distinctive, and most productive, operations in probability theory: *expectation.* In traditional probability, given a probability space $(X, \mathcal{A}, p)$ and an integrable function (random variable) $f : X \to \mathbb{R}$, the *expectation value* of f is the Lebesgue integral

$$\mathbb{E}_p[f] \;\coloneqq\; \int_X f(x)\, p(dx).$$

Now, how do we talk about this categorically? To answer that, it helps to distinguish, as above, between *when the numbers are probabilities* and *when they are arbitrary quantities*:

- *Averaging probabilities* has often an interpretation of "combining" or "composing" two sources of randomness in a canonical way. This idea leads, categorically, to probability monads, and to the (Kleisli) category of Markov kernels;
- More generally, *averaging arbitrary quantities* is a helpful procedure to have a "rough estimate" of random number, or a "placeholder for the entire distribution". This idea leads, categorically, to *algebras* of a probability monad.

In categorical probability, we can achieve further abstraction by moving from monads and Markov kernels to the more general setting of *Markov categories* [3], where the idea of "averaging probability" becomes a simple composition of morphisms.

We can extend that idea: as we show, we can capture these notions of expectation for general quantities, and not just for probabilities, in terms of *presheaves* on a Markov category. The main result is the following.

Theorem 1. *Let* C *be a Markov category, and denote by* $\mathrm{C}_{\mathrm{det}}$ *the wide subcategory of deterministic morphisms. For every object* R*, there is a bijection between*

1. *Notions of expectation for* R*-valued deterministic morphisms (random variables);*
2. *Extensions of the presheaf* $\mathrm{C}_{\mathrm{det}}(-, R) : \mathrm{C}_{\mathrm{det}}^{\mathrm{op}} \to \mathrm{Set}$ *to the whole of* C*.*

Moreover, using Street's formal theory of monads [6], when C is the Kleisli category of a probability monad, through the theorem above we recover exactly its algebras.

References

[1] Paolo Perrone, *How to Represent Non-Representable Functors.* Submitted (2024), available at arXiv:2410.08019.

[2] Paolo Perrone and Ruben Van Belle, *Convergence of martingales via enriched dagger categories.* Submitted (2024), available at arXiv:2404.15191.

[3] Tobias Fritz, *A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics.* Advances in Mathematics 370 (2020), available at arXiv:1908.07021.

[4] Tobias Fritz and Paolo Perrone, *Monads, partial evaluations, and rewriting.* Proceedings of MFPS (2020), available at arXiv:1810.06037.

[5] Paolo Perrone, *Categorical probability and stochastic dominance in metric spaces.* PhD thesis, University of Leipzig (2018), available at paoloperrone.org/phdthesis.pdf.

[6] Ross Street, *The formal theory of monads.* Journal of Pure and Applied Algebra 2,2 (1972).