Towards Minimal Axioms for Probabilistic Programming

Eigil Fjeldgren Rischel

May 7, 2026

The study the semantics of programs with stochastic behavior—probabilistic programming—is well established. Although equality between probability measures is undecidable (for the trivial reason that this would imply deciding equality between real numbers), it is still interesting to consider collections of axioms that allow you to deduce all the equations between distributions (or at least some large class of such equations). Probabilistic programs is usually described in terms of a monad on the category of "ordinary" types and functions—in these terms, such an axiomatization amounts to a presentation of this monad by generators and relations.

There are have been many previous attempts to give such probability monads a universal property of this type (I will not attempt to give a full review of the literature here). Mardare–Panangaden–Plotkin [2] gave a presentation in terms of what they termed *quantitative equational reasoning*. Roughly, the set of *finitely supported* measures on a set can be described as the free algebra for the monad of abstract convex combinations. The p-Wasserstein metric on this space is exactly the initial one obeying certain bounds on the distances between convex combinations (an algebra obeying these is dubbed a *barycentric algebra*), and the full set of distributions is the completion of this metric space. Ghani, Mardare, and the author [1] later showed how to axiomatize this metric completion again in terms of quantitative algebras, but this approach to probabilistic programming requires one to first implement the entire theory of metric spaces, and so is not exactly minimal. Piedeleu et al [3] recently gave an axiomatization of *discrete* probabilistic programs (those taking values in finite types), although similarly this assumes an existing theory of real numbers.

In recent work [4] I have exhibited a very different axiomatization of the monad of probability distributions, albeit in the context of measurable spaces, rather than something that looks more like functional programming. In simple terms, on the category of standard Borel measurable spaces, the Giry monad P is initial among strong commutative affine monads which

(i) Contain a binary operation $\mathrm{cf} : P(2)$, satisfying symmetry and interchange equations, and

(ii) preserve pullbacks along monomorphisms and sequential inverse limits.

The binary operation of (i) corresponds to a fair coinflip. The fact that P preserves pullbacks along monomorphisms amounts to the statement that, given $f : A \to B$ and $V \subseteq B$, a distribution on the preimage $f^{-1}(V)$ is equivalently a distribution on A such that the image happens to be concentrated on V. In measure theory this is true whenever V is a measurable subset of B.[1] Finally the inverse limit property is a consequence of *Kolmogorov's extension theorem*, which states that a distributions on an infinite product of spaces are in bijection with compatible families of distributions on the finite subproducts.

An interesting aspect of this result is that there appear to be several important properties which are not provable from the axioms, but nevertheless hold in the initial model (that is, they are admissible). For example, at several key steps in the proof we need to make an inference of the following type: let $\mu_i^{(m)}$ be a sequence of sequences of measures, so that for each m, $\sum_i \mu_i^{(m)}/2^i = \nu$ for some fixed ν. Suppose moreover each of the sequences $m \mapsto \mu_i^{(m)}$ is eventually constant, call this limit $\mu_i^{(\infty)}$. Then also $\sum_i \mu_i^{(\infty)} = \nu$. This is a trivial consequence of the topological properties of probability, but appears not to be provable from the axioms. However, as a consequence of our result it is true for anything that can be constructed using the axioms (i.e every probability measure).

Thus, with these fairly austere axioms, every probability measure can be presented (this is very straightforward) and every equation between these can be proved (this is the hard part). In addition to outlining the background and proof of the above result, I'll present some work in progress on strengthening it to categories that look more like the syntax of functional programming. In particular, I conjecture that in any elementary topos with NNO carrying a monad as above—satisfying instead an *internal* version of the filtered limits axiom—there is a well-defined map from the n-simplex to the space of distributions on n points.

References

[1] Radu Mardare, Neil Ghani, and Eigil Rischel. "Metric Equational Theories". In: *EPTCS* 428 (), pp. 144–160. (Visited on 11/10/2025).

[2] Radu Mardare, Prakash Panangaden, and Gordon D. Plotkin. "Quantitative Algebraic Reasoning". In: *Proceedings of the Thirty first Annual IEEE Symposium on Logic in Computer Science (LICS 2016)*. New York City: IEEE Computer Society Press, July 2016, pp. 700–709.

[3] Robin Piedeleu et al. "A Complete Axiomatisation of Equivalence for Discrete Probabilistic Programming". In: May 2025, pp. 202–229. ISBN: 978-3-031-91120-0. DOI: 10.1007/978-3-031-91121-7_9.

[1] But not in general—but standard Borel spaces have the nice property that every monomorphism between them is in fact the inclusion of a measurable subset. In this they are somewhat analogous to compact Hausdorff spaces.

[4] Eigil Fjeldgren Rischel. *The Universal Property of Measure-Theoretic Probability*. 2025. arXiv: 2512.15485 [math.CT]. URL: https://arxiv.org/abs/2512.15485.