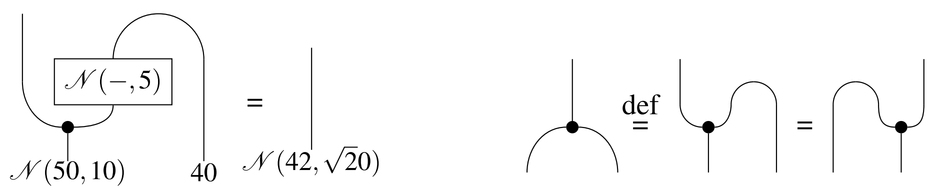
A Hypergraph Category for Exact Gaussian Inference

Dario Stein
Radboud University Nijmegen
dario.stein@ru.nl

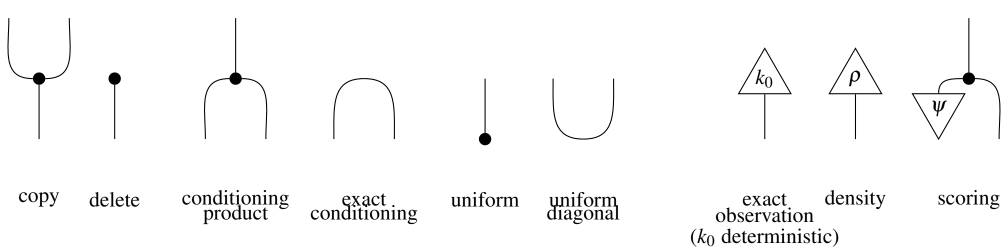
1 Uninformative Priors and Duality for Conditioning

We recall *exact conditioning* of random variables as a primitive for Bayesian inference, and argue that the presence of *uninformative priors* gives rise to a powerful self-duality via Frobenius structures. We apply this approach to Gaussian probability, which we have to synthetically extend with an idealized uniform distribution over the real line to obtain that kind of duality. We achieve this through a novel construction combining linear relations with probability.

In [18], we have proposed exact conditioning as a primitive for Bayesian inference and probabilistic programming. We demonstrate this using a noisy measurement example: Let some quantity X be normally distributed with mean $\mu = 50$ and standard deviation $\sigma = 10$, written $\mathcal{N}(50,10)$. We only have access to a noisy measurement Y, which itself has standard deviation 5, and observe a value of $Y = 40$. Conditioned on that observation, the posterior distribution over X is now $\mathcal{N}(42,\sqrt{20})$. In a categorical setting, the exact conditioning operation can be rendered as an effect $(=\!:\!=) : X \times X \to I$ in an appropriate symmetric monoidal category of 'open inference problems' (see below). We draw $(=\!:\!=)$ as a cap, and the comultiplication is copying.



We can define a *conditioning product* $\bullet : X \otimes X \to X$ in terms of $(=\!:\!=)$ as above, and proved in [16] that it satisfies the laws of a special commutative Frobenius multiplication. The last puzzle piece is whether $\bullet$ admits a unit $u_X : I \to X$, which must act as an uninformative prior over X, a generalized uniform distribution, conditioning on which gives no information. With such units in place, every object has the structure of a Frobenius algebra, making the surrounding category a hypergraph category [6] enjoying a rich duality theory: For example, effects can be seen as synthetic analogues of density functions in probability, and the study of open inference problems $X \to Y$ can be reduced to states $I \to X \otimes Y$. We obtain a convenient graphical formalism for conditioning and arrive at the following dictionary



Submitted to:
ACT 2022

Treating Bayesian inference using Frobenius structures and closed categories is a classic idea due to [5], and [15] identifies hypergraph categories as a natural setting for celebrated algorithms such as message-passing inference. We overcome **two challenges** that have hitherto prevented this approach from applying to situations with continuous random variables such as Gaussians: Firstly, exact conditioning has to be defined in a general setting possibly involving probability zero observations. We address this using the notion of open inference problem in Markov categories. The second challenge, which is the focus of this abstract, is that uniform distributions over spaces such as the real line do not exist measure-theoretically. This is often addressed informally using the method of *improper priors*, where the Lebesgue measure is treated as a probability distribution despite being unnormalized [11]. Instead, we propose a categorical construction to extend Gaussian probability with uninformative priors using a refined linear relations model.

2 Background: Compositional Semantics for Exact Conditioning

Markov and CD categories [7, 4] are receiving increased attention as a convenient formalism for categorical probability theory [10, 8, 13, 9] and a foundation for the semantics of probabilistic programming languages [16]. Morphisms in a Markov category $\mathbb{C}$ are thought of as generalized stochastic maps, and admit abstract characterizations of notions such as (conditional) independence, almost-sure equality, support and conditional probability. In [18], we introduce a synthetic notion of inference problem and define a CD category $\mathsf{Cond}(\mathbb{C})$ which faithfully includes $\mathbb{C}$ while internalizing exact conditioning. Morphisms in $\mathsf{Cond}(\mathbb{C})$ are *open inference problems* modulo contextual equivalence $\sim$, in an optic-like construction

$$\mathsf{Cond}(\mathbb{C})(X,Y) = \left(\sum_{K \in \mathsf{obj}(\mathbb{C})} \mathbb{C}(X, Y \otimes K) \times \mathbb{C}_{\mathsf{det}}(I,K) \right) / \sim$$

This approach to conditioning purely relies on the compositional structure of $\mathbb{C}$ and avoids all mention of measure theory, densities or limits. Naively conditioning on probability zero events easily leads to paradoxes such as Borel's paradox (e.g. [14]), which the categorical approach can help sidestep.

3 Decorated Linear Relations for Gaussian Probability

Our introductory example takes place in $\mathsf{Cond}(\mathsf{Gauss})$, where Gauss is the Markov category of finite-dimensional vector spaces and affine maps with Gaussian noise [7]. Conditioning in $\mathsf{Cond}(\mathsf{Gauss})$ does not admit a unit, because every Gaussian distribution is biased towards its mean and never fully uninformative. The novel contribution of this work is our construction of a Markov category $\mathsf{GaussEx}$ of *extended Gaussians* which extends Gaussian distributions by 'nondeterministic noise' along a vector subspace D [17]. Roughly, an extended Gaussian on a vector space X is an ordinary Gaussian on a quotient X/D. More generally, we reanalyze linear relations and define for every functor $S : \mathsf{Vec} \to \mathsf{CMon}$ two Markov categories Lin_S of *decorated linear maps* and LinRel_S of *decorated linear relations* as

$$\mathsf{Lin}_S(X,Y) = \mathsf{Vec}(X,Y) \times S(Y) \qquad \mathsf{LinRel}_S(X,Y) = \sum_{D \subseteq X} \mathsf{Vec}(X, Y/D) \times S(Y/D)$$

of which Gauss and $\mathsf{GaussEx}$ are special cases for appropriate S. Linear relations have been considered in [2, 1, 3]; their seamless probabilistic extension is surprising, because probability and nondeterminism do not combine well in general [19, 12]. Conditioning now has a unit, namely the nondeterministic distribution on X, which gives $\mathsf{Cond}(\mathsf{GaussEx})$ the desired structure of a hypergraph category.

References

[1] John C. Baez, Brandon Coya & Franciscus Rebro (2018): *Props in Network Theory*. arXiv:1707.08321.

[2] John C. Baez & Jason Erbele (2015): *Categories in Control*. Theory Appl. Categ. 30, pp. 836–881.

[3] Filippo Bonchi, Pawel Sobocinski & Fabio Zanasi (2017): *The calculus of signal flow diagrams I: linear relations on streams*. Inform. Comput. 252.

[4] Kenta Cho & B. Jacobs (2019): *Disintegration and Bayesian inversion via string diagrams*. Mathematical Structures in Computer Science 29, pp. 938 – 971.

[5] Bob Coecke & Robert W. Spekkens (2011): *Picturing classical and quantum Bayesian inference*. Synthese 186, pp. 651–696.

[6] Brendan Fong & David I. Spivak (2019): *Hypergraph Categories*. ArXiv abs/1806.08304.

[7] T. Fritz (2020): *A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics*. Adv. Math. 370.

[8] T. Fritz, Tomáš Gonda, P. Perrone & Eigil Fjeldgren Rischel (2020): *Representable Markov Categories and Comparison of Statistical Experiments in Categorical Probability*. Available at `https://arxiv.org/abs/2010.07416`.

[9] Tobias Fritz, Tomáš Gonda & Paolo Perrone (2021): *De Finetti's Theorem in Categorical Probability*. arXiv.

[10] Tobias Fritz & Eigil Fjeldgren Rischel (2020): *Infinite products and zero-one laws in categorical probability*. Compositionality 2, doi:10.32408/compositionality-2-3. Available at `https://doi.org/10.32408/compositionality-2-3`.

[11] Andrew Gelman, John B. Carlin, Hal S. Stern & Donald B. Rubin (2004): *Bayesian Data Analysis*, 2nd ed. edition. Chapman and Hall/CRC.

[12] Alexandre Goy & Daniela Petrişan (2020): *Combining Probabilistic and Non-Deterministic Choice via Weak Distributive Laws*. In: Proceedings of the 35th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS '20, Association for Computing Machinery, New York, NY, USA, doi:10.1145/3373718.3394795. Available at `https://doi.org/10.1145/3373718.3394795`.

[13] Bart Jacobs (2021): *Multinomial and Hypergeometric Distributions in Markov Categories*. Electronic Proceedings in Theoretical Computer Science 351, pp. 98–115, doi:10.4204/eptcs.351.7. Available at `https://doi.org/10.4204%2Feptcs.351.7`.

[14] Jules Jacobs (2021): *Paradoxes of Probabilistic Programming*. In: Proc. POPL 2021.

[15] Jason Morton (2014): *Belief propagation in monoidal categories*. Electronic Proceedings in Theoretical Computer Science 172, pp. 262–269, doi:10.4204/EPTCS.172.18.

[16] Dario Stein (2021): *Structural Foundations for Probabilistic Programming Languages*. Ph.D. thesis, University of Oxford.

[17] Dario Stein (2022): *Decorated Linear Relations: Extending Gaussian Probability with Uninformative Priors*, doi:10.48550/ARXIV.2204.14024. Available at `https://arxiv.org/abs/2204.14024`.

[18] Dario Stein & Sam Staton (2021): *Compositional Semantics for Probabilistic Programs with Exact Conditioning*. In: Proceedings of Thirty-Sixth Annual ACM/IEEE Conference on Logic in Computer Science (LICS 2021).

[19] Maaike Zwart & Dan Marsden (2019): *No-Go Theorems for Distributive Laws*. In: Proceedings of the 34th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS '19, IEEE Press.