Probability Sheaves and Sweedler Runners

Constant arities, random states, and distribution-sensitive observers

Companion note to Part XA

11 October 2026

Abstract

We use Simpson's probability-sheaf topos as a possible ambient category for Sweedler runners. Its Giry-type monad M is accessible and preserves small coproducts. This leads to a selection theorem: every natural transformation from the internal polynomial monad of finite formal lotteries to M sends each *global* lottery to one of its tickets, rather than to its probability mixture. Nevertheless, a sampling theory with response object $B = \mathrm{RV}(\{0,1\})$ has a natural monad interpretation in M. We calculate fresh and reused random coins, their conditional risks, and a stochastic state transformer that stores the last sampled bit. Finally we give an explicit distribution-sensitive observation unavailable for the ordinary pointwise Boolean object of QBS. The comparison with Sweedler measuring comonoids remains a separate structural problem.

1 Which topos, and which Booleans?

Part XA uses quasi-Borel spaces (QBS) as an ambient setting for probabilistic effects. Simpson's *probability sheaves* provide a different semantics: variation of sample space, almost-sure equality, and probabilistic independence become categorical structure [1, 2, 3]. Simpson also distinguishes the random topos and random probability sheaves [4]; this note studies only his first topos.

Let **SBP** be a small presentation of the category of Polish probability spaces and measurable, measure-preserving maps. Maps in this site are actual functions, not almost-sure equivalence classes. Simpson proves that the right Ore condition supplies the atomic topology and that

$$\mathcal{P} = \mathrm{Sh}_{\mathrm{at}}(\mathbf{SBP})$$

is an atomic Grothendieck topos [1, Secs. 2–3]. A map $q : \Omega' \to \Omega$ describes an extension of sample space; independent pullbacks combine two such extensions. For a Polish space A, set

$$\underline{A} = \mathrm{RV}(A), \qquad \underline{A}(\Omega) = \{X : \Omega \to A \text{ measurable}\}/ =_{\text{a.s.}} .$$

The functor $\mathrm{RV} : \mathbf{Pol}_{\mathrm{um}} \to \mathcal{P}$ is faithful but not full.

There are two relevant Boolean objects:

$$D = 1 + 1, \qquad B = \mathrm{RV}(\{0,1\}).$$

For the fair two-point probability space $C = (\{0,1\}, \nu)$,

$$D(C) = \{0,1\}, \qquad B(C) = \{0, 1, \mathrm{id}_C, 1 - \mathrm{id}_C\}.$$

D has only constant labels at each stage, whereas B contains nonconstant random variables. We call D a *constant finite* object. It would be misleading to contrast these using 'decidability': $\mathcal{P}$ is Boolean, so that distinction cannot mean that B is non-decidable in the internal logical sense.

2 The probabilistic monad and its coproducts

Simpson defines a functor $\widetilde{M}$ on presheaves by the coend

$$(\widetilde{M}F)(\Omega) = \int^{\Omega'} \mathbf{SBP}(\Omega', \Omega) \times F(\Omega'). \tag{1}$$

An element is represented by $[q, z]$, where $q : \Omega' \to \Omega$ is a sample-space extension and $z \in F(\Omega')$. The colimit implicit in (1) is *not filtered*: for example, the parallel maps $\mathrm{id}_C, \neg : C \rightrightarrows C$ over 1 have no measure-preserving equaliser. Indeed, the coend identifies $[C \to 1, \mathrm{id}_C]$ and $[C \to 1, \neg]$.

The coend first gives a covariant assignment in Ω; its contravariant restriction maps are *not* obtained simply from the coend. Simpson constructs them with independent pullbacks, proves that the resulting presheaf is already a sheaf, and obtains a strong monad M on $\mathcal{P}$ [1, Thm. 7]. His comparison theorem supplies

$$M(\underline{A}) \cong \underline{\mathrm{Prob}(A)}, \tag{2}$$

natural in the conventional universally measurable maps and compatible with the strong monad structures [1, Thm. 8]. The fair coin $\mathsf{c} : 1 \to M(B)$ corresponds to $\nu = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$ and is represented by $[C \to 1, \mathrm{id}_C]$.

Proposition 2.1 (Accessibility and coproducts)**.** *The monad $M : \mathcal{P} \to \mathcal{P}$ is accessible and preserves small coproducts. For each finite cardinal n, the unit $\eta_n : n \to M(n)$ is an isomorphism.*

Proof. Let $i : \mathcal{P} \hookrightarrow [\mathbf{SBP}^{\mathrm{op}}, \mathbf{Set}]$ denote sheaf inclusion. Both categories are locally presentable and i is a right adjoint, hence accessible. For fixed Ω, (1) is a colimit of evaluations of F, and therefore preserves all small colimits of presheaves. By Simpson's Theorem 7, $\widetilde{M}$ lands in sheaves without further sheafification; even were sheafification needed, it would be a cocontinuous left adjoint. Thus $\widetilde{M}$ is accessible and $M = \widetilde{M}i$ is accessible.

Coproducts of probability sheaves are computed stagewise: for a singleton cover q, any matching section of a pointwise disjoint sum belongs to one summand, and descends there. Hence i preserves coproducts, and the colimit formula implies that M does as well. The extension category over Ω is nonempty and connected: any two extensions admit a common refinement by an independent product. Consequently the coend of the constant presheaf 1 is 1, and its unit is invertible. Preservation of coproducts gives $M(n) \cong n$; the isomorphism is induced by η_n. □

Thus Simpson's monad is eligible as an object of $[\mathcal{P}, \mathcal{P}]_{\mathrm{acc}}$. Its coproduct-preservation has a stronger conceptual consequence: M does not mix separate global coproduct summands, even though it generates nontrivial random variables in objects such as B.

3 The internal polynomial lottery monad

Let Δ_n be the ordinary probability simplex and put $\mathbf{\Delta}_n = \mathrm{RV}(\Delta_n)$. Define

$$L(X) = \coprod_{n \geq 1} \mathbf{\Delta}_n \times X^n. \tag{3}$$

This is the monad of *formal*, ordered finite lotteries. Its unit is $x \mapsto ((1); x)$; substitution sends nested lotteries $((p_i); ((q_{ij}); x_{ij})_i)$ to $((p_i q_{ij})_{i,j}; (x_{ij})_{i,j})$. The simplex identity and monad axioms hold pointwise. Filtered colimits commute with finite products in a Grothendieck topos, so L is finitary. No quotient by permutation, zero weights, or repeated tickets is imposed.

At Ω, a formal lottery includes $p_i \in \mathrm{RV}([0,1])(\Omega)$, so weights may be random and correlated with the tickets. For every conventional $\underline{A}$ there is an evaluation map

$$e_A : L(\underline{A}) \to M(\underline{A}), \qquad ((p_i); (X_i)) \longmapsto \left[\omega \mapsto \sum_i p_i(\omega)\delta_{X_i(\omega)}\right]. \tag{4}$$

It is natural for maps induced by universally measurable functions $A \to A'$, and finite substitution has the expected measure-valued arithmetic on conventional random-variable sheaves. But the family (4) is *not* natural for arbitrary sheaf morphisms. The obstruction can be stated without introducing a distribution-sensitive morphism.

4 A selection theorem for all natural $L \Rightarrow M$

Proposition 4.1 (Global finite-lottery selection)**.** *Let $\tau : L \Rightarrow M$ be* any *natural transformation, not necessarily a monad morphism. For every $p \in \Delta_n = \mathbf{\Delta}_n(1)$ there is an index $k_\tau(p) \in \{1, \dots, n\}$ such that, for every object X of $\mathcal{P}$ and all global elements $x_i : 1 \to X$,*

$$\tau_X(p; x_1, \dots, x_n) = \eta_X(x_{k_\tau(p)}).$$

In particular, no natural $L \Rightarrow M$ sends the global binary half-and-half lottery on $0, 1 \in B(1)$ to c.

Proof. Write $n = 1 + \cdots + 1$, with global points $e_i : 1 \to n$, and let $t_0 = (p; e_1, \dots, e_n) \in L(n)(1)$. By Proposition 2.1, $\eta_n : n \to M(n)$ is an isomorphism. Therefore $\tau_n(t_0) = \eta_n(e_k)$ for a unique index k. For any tuple (x_i), its copairing $f : n \to X$ satisfies $(p; x_1, \dots, x_n) = L(f)(t_0)$. Naturality of τ gives

$$\tau_X(p; x_1, \dots, x_n) = M(f)(\tau_n(t_0)) = \eta_X(f(e_k)) = \eta_X(x_k).$$

For $X = B$ and tickets $0, 1$, $\eta_B(0)$ and $\eta_B(1)$ are Dirac measures under (2), whereas c is fair and non-Dirac. □

There *are* natural transformations $L \Rightarrow M$, for example, 'take the first ticket' followed by η. The result says that *all* of them behave as selectors on global finite lotteries. It makes no corresponding claim for arbitrary local sections with random coefficients.

The obstruction comes from constant finite arities n, not from an absence of probability operations in $\mathcal{P}$ and not primarily from statistical observers. This is specific to the probability-sheaf semantics. The standard finite coproduct $1 + 1$ in QBS does admit nonconstant random elements, and its usual probability monad can naturally evaluate finite weighted lotteries. For an operational sampler in $\mathcal{P}$, the response object must instead be the random-variable object B, not the constant two-point object D.

5 An operational theory of sampling

Let $H(X) = X^B$, and let $T_B = \mathrm{Free}(H)$. Since exponentiation by B is a right adjoint between locally presentable categories, it is accessible; hence the free monad exists. The strength of M and $\mathsf{c} : 1 \to M(B)$ define

$$\sigma_X : X^B \longrightarrow X^B \times M(B) \xrightarrow{\mathrm{st}} M(X^B \times B) \xrightarrow{M(\mathrm{ev})} M(X).$$

This family is natural for *all* probability-sheaf morphisms, so the free-monad property yields a genuine monad morphism

$$\alpha : T_B \longrightarrow M. \tag{5}$$

In contrast to L, this is not an interpretation of branching over a constant finite arity; the sampler returns an element of B.

There are plausible further equations for an *ideal fair-coin* theory. Writing $\mathsf{sample}(k)$ for the command with continuation $k : B \to T_B(X)$, they take the forms

$$\mathsf{sample}(\lambda b.\, t) = t, \qquad \mathsf{sample}(k) = \mathsf{sample}(k \circ \neg_B),$$
$$\mathsf{sample}(\lambda b.\, \mathsf{sample}(\lambda b'.\, u(b, b'))) = \mathsf{sample}(\lambda b'.\, \mathsf{sample}(\lambda b.\, u(b, b'))).$$

They express discard, symmetry, and exchange of independent sampling commands. The fair-coin interpretation satisfies the corresponding probability-law equalities. Identifying the exact quotient theory and proving faithfulness of its map to M are separate questions; we do *not* assert that these schemata present an internal dyadic-distribution monad on all of $\mathcal{P}$.

6 Fresh coins, reuse, and shared parameters

In the Kleisli category of M, form

$$\begin{aligned} \mathsf{fresh}: &\quad x \leftarrow \mathsf{c};\ y \leftarrow \mathsf{c}; \mathrm{return}(x, y), \\ \mathsf{reuse}: &\quad x \leftarrow \mathsf{c}; \mathrm{return}(x, x). \end{aligned}$$

Under $B \times B \cong \mathrm{RV}(\{0,1\}^2)$, the first is represented by $C \times C \to 1$ with independent product measure and outcome (π_1, π_2); the second is represented by $C \to 1$ and $(\mathrm{id}_C, \mathrm{id}_C)$. Their laws are

$$\begin{aligned} \mu_{\mathrm{f}} &= \tfrac{1}{4}(\delta_{00} + \delta_{01} + \delta_{10} + \delta_{11}), \\ \mu_{\mathrm{r}} &= \tfrac{1}{2}(\delta_{00} + \delta_{11}). \end{aligned}$$

For the unsafe event $U = \{(1,1)\}$ the risks are $1/4$ and $1/2$, respectively. This calculation is correct but is *not* a reason to prefer probability sheaves over QBS: any suitable commutative probability monad can express it.

A more informative relative version fixes $p \in \mathrm{RV}([0,1])(\Omega)$ and samples twice independently *conditional on the original sample space* Ω. Then

$$\Pr(U \mid \Omega)_{\mathrm{fresh}} = p^2, \qquad \Pr(U \mid \Omega)_{\mathrm{reuse}} = p.$$

The unconditional risks are $\mathbb{E}[p^2]$ and $\mathbb{E}[p]$. In particular

$$\mathbb{E}[p^2] = (\mathbb{E}p)^2 + \mathrm{Var}(p).$$

Unless p is almost surely constant, the two fresh samples are *marginally dependent*, even though they are conditionally independent over Ω. This relative-independence calculation is an appropriate input for a runner certificate, but QBS can also represent it.

7 A stochastic one-bit memory

We should distinguish the state-outside-effect expression $T(S \times -) \to S \times M(-)$ from the standard stochastic state-transformer monad

$$\mathrm{St}_S\, M(X) = \big(M(S \times X)\big)^S. \tag{6}$$

A monad morphism $T \to \mathrm{St}_S\, M$ threads state *inside* the residual probability effect. This formulation is familiar from stateful and effectful runners [6, 7]. It is the appropriate immediate test for probabilistic state updates; we do *not* identify it without proof with a coalgebra for the Sweedler comonoid $\{T, M\}$ associated with a previously chosen duoidal structure. For fixed S, (6) is accessible, since $S \times -$, M, and $(-)^S$ are accessible.

For a state meant to remember the sampled Boolean, choose

$$S = B = \mathrm{RV}(\{0,1\}),$$

not $D = 1 + 1$. The co-operation 'sample and store' is the map

$$\rho_{\mathrm{sample}} : B \xrightarrow{!} 1 \xrightarrow{\mathsf{c}} M(B) \xrightarrow{M(\Delta_B)} M(B \times B), \tag{7}$$

where the first component is the new state and the second is the returned bit. A deterministic 'read' command is implemented by

$$\rho_{\text{read}} : B \xrightarrow{\Delta_B} B \times B \xrightarrow{\eta} M(B \times B).$$

These co-operations induce a monad morphism from the free theory of sampling and reading into $\mathrm{St}_B\, M$. In particular, the sample can overwrite the state with a *random* Boolean and reading returns that very same random Boolean; it does not generate a fresh independent sample.

Why not use D for the state? By Proposition 2.1, $M(D) \cong D$, and more generally

$$M(D \times X) \cong M(X) + M(X).$$

A global probabilistic result in $M(D \times X)$ must choose a single constant summand. It therefore cannot encode a new D-state correlated with the outcome of an independent fair coin. This is a second operational consequence of the D/B distinction: putting the state *inside* the effect is necessary but is not sufficient if the state object itself is chosen incorrectly.

8 A specifically sheaf-theoretic observer

Simpson exhibits a morphism $h : B \to B$ given at a probability space Ω by

$$h_\Omega(X) = \begin{cases} 0, & X \text{ is almost surely constant,} \\ 1, & \text{otherwise.} \end{cases}$$

It is natural because measure-preserving change of sample space preserves the distribution of X [1, Sec. 4]. It cannot be induced by a pointwise Boolean function $\{0,1\} \to \{0,1\}$: at the terminal sample space it sends both constant Booleans to 0, but at C it sends id_C to 1.

Consequently, in $M(B)$,

$$M(h)(\mathsf{c}) = \eta_B(1), \qquad h(0) = h(1) = 0. \tag{8}$$

Applying h to a *sampled random-variable section* is thus different from applying h to each constant ticket *before* sampling. The stronger selection theorem in Proposition 4.1 did not require this fact; it is a separate phenomenon about observational power.

It also gives an experiment with the state B of §7. After 'sample and store', let the observer read $h(s)$ rather than s. The observed result is constantly 1; if the state is a fixed constant 0 or 1, the result is constantly 0. This probe is *not* definable as a pointwise Boolean observer in QBS. Its presence might strengthen certificates by allowing statistical properties to be tested; conversely, it may make observers non-local in a way inappropriate for some operational interpretations.

This is a demonstrable extra morphism of $\mathcal{P}$, not yet a theorem that two runners equivalent in QBS become inequivalent in $\mathcal{P}$. Such a statement requires a specified comparison of machines and an observational-equivalence relation on each side. Establishing (or refuting) that strict separation is our next *P-specific* test.

9 Duoidal and measuring structure: orientation matters

The category $[\mathcal{P}, \mathcal{P}]_{\text{acc}}$ is locally presentable; composition and pointwise Cartesian product are accessible. There is an evident interchange

$$(F \times F') \circ (G \times G') \longrightarrow (F \circ G) \times (F' \circ G'),$$

obtained using the two projections from $G(X) \times G'(X)$. *Its direction matters.* In the conventions of López Franco–Vasilakopoulou, the interchange is $(a \star b) \diamond (c \star d) \to (a \diamond c) \star (b \diamond d)$, where

the measuring theorem concerns $\star$-monoids [5, Thm. 20]. The displayed elementary interchange takes $\star = \times$ and $\diamond = \circ$, so its $\star$-monoids are pointwise-product monoids, *not* monads under composition. It is therefore insufficient by itself to construct the desired $\{T_B, M\}$.

The relevant research problem is to furnish a duoidal structure with *composition as the monad tensor* and a suitable second biclosed tensor (for example an appropriately defined convolution on an arity-controlled endofunctor universe). The LFV theorem then applies if the underlying category is locally presentable, the composition tensor is accessible, and the second tensor is biclosed, with the requisite interchange. No P-specific obstacle to local presentability is apparent; the actual verification of this correctly oriented duoidal structure is still outstanding. A proposed cofree description of $\{\mathrm{Free}(H), M\}$ must wait for this check.

10 Conclusions and subsequent tests

There are now four distinct levels of claim:

1. *Established:* Simpson's M is accessible and preserves coproducts; constant finite objects admit no probabilistic mixing by M.
2. *Established:* every $L \Rightarrow M$ is a deterministic selector on global finite lotteries; the sampling theory T_B nevertheless maps naturally to M.
3. *Established:* independent samples, correlated reuse, conditional risk, stochastic memory using $S = B$, and the distribution-sensitive observer h can be calculated explicitly.
4. *Open:* the correctly oriented Sweedler measuring structure, a comparison of its runner coalgebras with stochastic state transformers, and a strict P-versus-QBS separation theorem for runner observations.

The first P-specific runner experiment should combine the sample-and-store map (7) with the h-observer (8). We should then specify two implementations and a common comparison of observations before claiming one ambient setting identifies them and the other separates them.

Further ambient candidates can be mapped without developing them here: Simpson's random topos and random probability sheaves [4]; the related probabilistic nominal/sheaf construction of Li–Aytac–Johnson-Freyd–Ahmed–Holtzen [8]; ω-QBS when recursion matters; and Markov categories, whose copy–discard tensor is non-Cartesian in the relevant sense and would require a different duoidal analysis.

References

[1] A. Simpson, "Probability Sheaves and the Giry Monad," *CALCO 2017*, LIPIcs 72, Article 1 (2017). https://doi.org/10.4230/LIPIcs.CALCO.2017.1.

[2] A. Simpson, "Equivalence and Conditional Independence in Atomic Sheaf Logic," *LICS 2024*, pp. 70:1–70:14 (Distinguished Paper). https://arxiv.org/abs/2405.11073.

[3] A. Simpson, "Equivalence and Conditional Independence in Atomic Sheaf Logic," *Journal of the ACM* **73**(3), Article 17 (June 2026). https://doi.org/10.1145/3809163.

[4] A. Simpson, "Three Toposes for Probability and Randomness," colloquium slides (2024). https://alexksimpson.github.io/Talks/ThreeToposes.pdf.

[5] I. López Franco and C. Vasilakopoulou, "Duoidal Categories, Measuring Comonoids and Enrichment" (2020), Theorem 20. https://arxiv.org/abs/2005.01340.

[6] T. Uustalu, "Stateful Runners of Effectful Computations," *Electronic Notes in Theoretical Computer Science* **319**, 403–421 (2015). https://doi.org/10.1016/j.entcs.2015.12.024.

[7] D. Ahman and A. Bauer, "Runners in Action," *ESOP 2020*, LNCS 12075, pp. 29–55. https://doi.org/10.1007/978-3-030-44914-8_2.

[8] J. M. Li, J. Aytac, P. Johnson-Freyd, A. Ahmed, and S. Holtzen, "A Nominal Approach to Probabilistic Separation Logic" (2024). https://arxiv.org/abs/2405.06826.