Probability Functor Models

John Harby
4/23/2025

Abstract

We introduce the probability functor model, a formal framework for representing deterministic and probabilistic computation within a unified structure. The model defines a mapping

$$F: S \to \mathrm{Prob}(Y) \times \mathbb{R}_{\geq 0},$$

where outputs consist of both a probability distribution and an explicitly represented uncertainty measure.

Deterministic computation is embedded as a special case via Dirac distributions with zero uncertainty, while probabilistic behavior is confined to a complementary region in which uncertainty is intrinsic and explicitly modeled. A routing mechanism determines whether evaluation proceeds through deterministic or probabilistic components.

Within this framework, error admits a precise decomposition: it is zero on the deterministic domain and coincides with intrinsic probabilistic uncertainty elsewhere. This yields an uncertainty localization property, ensuring that uncertainty is confined to explicitly defined regions and does not propagate implicitly across the system. Furthermore, conditional dependencies arise only through explicitly modeled probabilistic structure.

The resulting formulation provides a mathematically consistent foundation for hybrid systems combining exact computation and probabilistic reasoning, with explicit control over uncertainty, error attribution, and model behavior.

Deterministic and Probability Functor Models

Let S be the input space and let Y be the output space. Let

$$D \subset S$$

denote the region in which an exact governed mapping is available, and define

$$U := S \setminus D$$

as the complementary region in which uncertainty is explicitly modeled.

Let $\mathrm{Prob}(Y)$ denote the space of probability distributions on Y.

Definition (Deterministic functor model). A deterministic functor model is a map

$$g: D \to Y.$$

We extend it to a probabilistic representation

$$F_{\text{det}}: D \to \text{Prob}(Y) \times \mathbb{R}_{\geq 0}$$

defined by

$$F_{\text{det}}(x) := (\delta_{g(x)}, 0),$$

where $\delta_{g(x)}$ denotes the Dirac distribution concentrated at $g(x)$.

Definition (Probability functor model). A probability functor model is a map

$$F_{\text{prob}}: U \to \text{Prob}(Y) \times \mathbb{R}_{\geq 0},$$

which admits a decomposition

$$F_{\text{prob}}(x) := \big(f_{\text{prob}}(x), \varepsilon_{\text{prob}}(x)\big),$$

where

$$f_{\text{prob}}: U \to \text{Prob}(Y), \quad \varepsilon_{\text{prob}}: U \to \mathbb{R}_{\geq 0}.$$

Thus, uncertainty is explicitly internalized as part of the model output.

Definition (Routing functions). An *external routing function* is a map

$$r_{\text{ext}}: S \times \Lambda \to \{\text{det}, \text{prob}\}.$$

For fixed $\lambda \in \Lambda$, define

$$r_\lambda(x) := r_{\text{ext}}(x, \lambda).$$

An *internal routing function* is a map

$$r_{\text{int}}: S \to \{\text{det}, \text{prob}\}.$$

Definition (Hybrid functor system). Let r denote either r_{int} or r_λ.

Assume the routing constraint

$$r(x) = \text{det} \quad \Rightarrow \quad x \in D.$$

Define the hybrid functor

$$F: S \to \text{Prob}(Y) \times \mathbb{R}_{\geq 0}$$

by

$$F(x) := \begin{cases} F_{\text{det}}(x), & r(x) = \text{det}, \\ F_{\text{prob}}(x), & r(x) = \text{prob}. \end{cases}$$

We define the projection maps

$$\pi_1: \text{Prob}(Y) \times \mathbb{R}_{\geq 0} \to \text{Prob}(Y), \quad \pi_2: \text{Prob}(Y) \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0},$$

and write

$$F(x) = (\mu_x, \varepsilon(x)),$$

where

$$\mu_x := \pi_1(F(x)), \quad \varepsilon(x) := \pi_2(F(x)).$$

Error Decomposition Principle

Let $F: S \to \mathrm{Prob}(Y) \times \mathbb{R}_{\geq 0}$ be a hybrid functor system.

Principle (Error Decomposition). The total error functional is given by

$$\varepsilon(x) := \pi_2(F(x)).$$

Then necessarily:

$$\varepsilon(x) = \begin{cases} 0, & x \in D, \\ \varepsilon_{\mathrm{prob}}(x), & x \in U. \end{cases}$$

Interpretation. All nonzero error is intrinsic to the probabilistic component of the model and arises solely from the uncertainty defined on $U = S \setminus D$.

Remark. The identification of all nonzero error as intrinsic uncertainty is an architectural assumption of the model.

Uncertainty Localization Principle

Let F be a hybrid functor system with routing function r.

Principle (Uncertainty Localization). The system satisfies the uncertainty localization property if:

$$\varepsilon(x) = 0 \quad \forall x \in D,$$

and

$$\varepsilon(x) = \varepsilon_{\mathrm{prob}}(x) \quad \forall x \in U,$$

with the additional requirement that uncertainty on U does not propagate into evaluations on D.

In particular, uncertainty is explicitly confined to U and does not affect deterministic computation on D.

Controlled Conditional Uncertainty

Let $(\Omega, \mathcal{F}, P)$ be a probability space induced by the probabilistic functor components, and let $E_1, E_2 \in \mathcal{F}$ be events.

Principle (Controlled Conditional Uncertainty). A conditional dependency

$$P(E_1 \mid E_2)$$

is admissible only if:

- the dependency between E_1 and E_2 is explicitly represented within the probabilistic functor structure, and
- the associated uncertainty is explicitly accounted for within the model.

Otherwise, the model treats E_1 and E_2 as independent, i.e.

$$P(E_1 \mid E_2) = P(E_1),$$

as a modeling convention.

Thus, probabilistic dependence arises only through explicitly declared structure and does not propagate implicitly across model components.

Governance and Transparency via Uncertainty Centralization

In conventional machine learning systems, uncertainty is typically distributed implicitly across model parameters and computations. Consequently:

- uncertainty compounds without explicit control,
- dependencies emerge implicitly,
- conditional relationships are not explicitly represented, and
- the origin of error is difficult to trace.

In the present framework, uncertainty is centralized and explicitly modeled.

Principle (Uncertainty Centralization). All probabilistic behavior is confined to explicitly defined probability functors, and all uncertainty is represented as part of the model output.

Consequences.

- Uncertainty is localized to specific components of the system.
- Conditional dependencies arise only through explicitly modeled interactions.
- Error can be attributed to intrinsic probabilistic uncertainty within U.
- The system admits auditability and traceability of probabilistic behavior.

Thus, the architecture provides a foundation for governed and transparent systems in which uncertainty is a controlled and explicitly observable object rather than an emergent side effect.

Illustrative Example of Hybrid Functor Evaluation

We provide a simple example to illustrate the behavior of the hybrid functor system and the associated error decomposition.

Let $S = \mathbb{R}$ and $Y = \mathbb{R}$. Define the deterministic region

$$D := \mathbb{Z} \subset \mathbb{R},$$

and its complement

$$U := \mathbb{R} \backslash \mathbb{Z}.$$

Deterministic component. Let $g: D \to Y$ be defined by

$$g(n) := n^2, \quad n \in \mathbb{Z}.$$

Then the deterministic functor is

$$F_{\text{det}}(n) = (\delta_{n^2}, 0).$$

Probabilistic component. For $x \in U$, define the probabilistic functor by

$$F_{\text{prob}}(x) := (\mu_x, \varepsilon_{\text{prob}}(x)),$$

where $\mu_x \in \text{Prob}(Y)$ is a distribution centered near x^2, for example a Gaussian

$$\mu_x = \mathcal{N}(x^2, \sigma^2),$$

and $\varepsilon_{\text{prob}}(x) := \sigma$ represents intrinsic uncertainty.

Routing. Define an internal routing function

$$r(x) = \begin{cases} \text{det}, & x \in \mathbb{Z}, \\ \text{prob}, & x \in \mathbb{R} \backslash \mathbb{Z}. \end{cases}$$

Hybrid evaluation. The hybrid functor $F: S \to \text{Prob}(Y) \times \mathbb{R}_{\geq 0}$ is given by

$$F(x) = \begin{cases} (\delta_{x^2}, 0), & x \in \mathbb{Z}, \\ (\mu_x, \varepsilon_{\text{prob}}(x)), & x \in \mathbb{R} \backslash \mathbb{Z}. \end{cases}$$

Error behavior. By construction,

$$\varepsilon(x) = \begin{cases} 0, & x \in D, \\ \varepsilon_{\text{prob}}(x), & x \in U. \end{cases}$$

Thus, all nonzero error is confined to U and arises solely from intrinsic uncertainty in the probabilistic component.

Interpretation. This example demonstrates:

- exact deterministic behavior on D,

- explicit probabilistic modeling on U,
- localization of uncertainty to the complement of the governed region, and
- consistency with the Error Decomposition and Uncertainty Localization principles.

Acknowledgements

The author acknowledges the use of ChatGPT-5 (OpenAI, Inc.) in the preparation of this manuscript. The system was used to assist with drafting, refinement, and formalization of portions of the text. All concepts, definitions, and final content decisions were determined by the author.