A category of probability space and a conditional expectation functor

Yoshihiro Ryu (Ritsumeikan University)
joint work with
Takanori Adachi (Ritsumeikan University)

An advantage of using category theory is that it can visualize relations between different mathematical fields. Further, when we find a relation between different mathematical fields, it sometimes helps for developings a theory in a new direction. This fact motivates us to use category theory for studying probability theory.

One of the most prominent trials of applying category theory to probability theory so far is Lawvere and Giry's approach of formulating transition probabilities in a monad framework ([Lawvere, 1962], [Giry, 1982]). However, their approach is based on two categories, the category of measurable spaces and the category of measurable spaces of a Polish space , not a category of probability spaces. Further, there are few trials of making categories consisting of all probability spaces due to a difficulty of finding an appropriate condition of their arrows.

Our approach is one of this simple-minded trials. We introduce a category **Prob** of all probability spaces in order to see a possible generalization of some classical tools in probability theory including conditional expectations. Actually, [Adachi, 2014] provides a simple category for formulating conditional expectations, but its objects and arrows are so limited that we cannot use it as a foundation of categorical probability theory.

Definition 1. A category **Prob** is the category whose objects are all probability spaces and the set of arrows between them are defined by

$$\begin{aligned}&\mathbf{Prob}((X,\Sigma_X,\mathbb{P}_X),(Y,\Sigma_Y,\mathbb{P}_Y))\\ &:= \{f^- \mid f:(Y,\Sigma_Y,\mathbb{P}_Y)\to(X,\Sigma_X,\mathbb{P}_X) : \text{measurable with } \mathbb{P}_Y\circ f^{-1}\ll\mathbb{P}_X\},\end{aligned}$$

where f^- is a symbol corresponding uniquely to a measurable function f.

We write $(X, \Sigma_X, \mathbb{P}_X) \xrightarrow{f^-} (Y, \Sigma_Y, \mathbb{P}_Y)$ in **Prob**, however, note that the arrow f^- has an opposite direction of the function f.

From now on, $f^- : (X, \Sigma_X, \mathbb{P}_X) \to (Y, \Sigma_Y, \mathbb{P}_Y)$ is an arbitrary arrow in **Prob**. For any $v \in \mathcal{L}^1(Y, \Sigma_Y, \mathbb{P}_Y)$, thanks to Radon-Nikodym theorem, we can find $E^{f^-}(v) \in \mathcal{L}^1(X, \Sigma_X, \mathbb{P}_X)$, a (version of) conditional expectation of v along f^-, satisfying

$$\int_A E^{f^-}(v)\, d\mathbb{P}_X = \int_{f^{-1}(A)} v\, d\mathbb{P}_Y$$

for all $A \in \Sigma_X$. This is a generalization of conditional expectation. Because if $f = id_\Omega : (\Omega, \mathcal{F}, \mathbb{P}) \to (\Omega, \mathcal{G}, \mathbb{P})$ and $\mathcal{G} \subset \mathcal{F}$, then $E^{id_\Omega^-}(v)$ becomes a usual conditional expectation $\mathbb{E}(v|\mathcal{G})$. Since the arrow $(\Omega, \mathcal{F}, \mathbb{P}) \xrightarrow{id_\Omega} (\Omega, \mathcal{G}, \mathbb{P})$ is identified as a sub σ-algebra $\mathcal{G}$ of $\mathcal{F}$, we can think of an arrow f^- in **Prob** as a σ-algebra.

Additionally, one can show the well-definedness of $[v]_{\sim_{\mathbb{P}_Y}} \mapsto [E^{f^-}(v)]_{\sim_{\mathbb{P}_X}}$, here $\sim_{\mathbb{P}}$ is $\mathbb{P}$-a.s. equivalence relation. So we have the first theorem:

Theorem 2. *There exists a contravariant functor $\mathcal{E}$ from* **Prob** *to* **Set** *(the category of all sets and all functions) as following:*

$$\begin{tikzcd}
X & (X, \Sigma_X, \mathbb{P}_X) \arrow[r, mapsto, "\mathcal{E}"] \arrow[d, "f^-"'] & \mathcal{E}(X, \Sigma_X, \mathbb{P}_X) & := & L^1(X, \Sigma_X, \mathbb{P}_X) & \ni & {[E^{f^-}(v)]_{\sim_{\mathbb{P}_X}}} \\
Y \arrow[u, "f"] & (Y, \Sigma_Y, \mathbb{P}_Y) \arrow[r, mapsto, "\mathcal{E}"] & \mathcal{E}(Y, \Sigma_Y, \mathbb{P}_Y) \arrow[u, "\mathcal{E}f^-"'] & := & L^1(Y, \Sigma_Y, \mathbb{P}_Y) & \ni & {[v]_{\sim_{\mathbb{P}_Y}}} . \arrow[u, mapsto, "\mathcal{E}f^-"']
\end{tikzcd}$$

Continually, we define a concept of measurability.

Definition 3. A random variable $v \in \mathcal{L}^\infty(Y, \Sigma_Y, \mathbb{P}_Y)$ is called f^--measurable if there exists $w \in \mathcal{L}^\infty(X, \Sigma_X, \mathbb{P}_X)$ such that $v \sim_{\mathbb{P}_Y} w \circ f$.

It seems natural because f^- is a "σ-algebra". Due to this definition, our second theorem is obtained.

Theorem 4. *Let u be an element of $\mathcal{L}^1(Y, \Sigma_Y, \mathbb{P}_Y)$ and v be a random variable in $\mathcal{L}^\infty(Y, \Sigma_Y, \mathbb{P}_Y)$, and assume that v is f^--measurable. Then we have*

$$E^{f^-}(v \cdot u) \sim_{\mathbb{P}_X} w \cdot E^{f^-}(u),$$

where $w \in \mathcal{L}^\infty(X, \Sigma_X, \mathbb{P}_X)$ is a random variable satisfying $v \sim_{\mathbb{P}_Y} w \circ f$.

This theorem shows that our "conditional expectation" still has a similar property about measurability.

Next definition is a modification of [Franz, 2003].

Definition 5. We say $v \in \mathcal{L}^1(Y, \Sigma_Y, \mathbb{P}_Y)$ is independent of f^- if there exists a measure preserving map q which makes the following diagram commute:

$$\begin{tikzcd}
 & (Y, \Sigma_Y, \mathbb{P}_Y) \arrow[dl, "v"'] \arrow[d, "q"] \arrow[dr, "f"] & \\
(\mathbf{R}, \mathcal{B}, \mathbb{P}_Y \circ v^{-1}) & (\mathbf{R} \times X, \mathcal{B} \otimes \Sigma_X, (\mathbb{P}_Y \circ v^{-1}) \otimes (\mathbb{P}_Y \circ f^{-1})) \arrow[l, "\pi_1"] \arrow[r, "\pi_2"'] & (X, \Sigma_X, \mathbb{P}_X).
\end{tikzcd}$$

By a straightforward calculation, we see that this definition means usual independence in the case of two σ-algebras.

Finally, we encounter our last theorem.

Theorem 6. *Let $v \in \mathcal{L}^1(Y, \Sigma_Y, \mathbb{P}_Y)$ be a random variable that is independent of f^-. Then we have,*

$$E^{f^-}(v) \sim_{\mathbb{P}_X} \mathbb{E}^{\mathbb{P}_Y}[v] E^{f^-}(1_Y).$$

When f is measure preserving, $E^{f^-}(1_Y) \sim_{\mathbb{P}_X} 1$, then the above formula turns well known formula of conditional expectation with independence since $E^{f^-}(1_Y)$ is the Radon-Nikodym derivative $d(\mathbb{P}_Y \circ f^{-1})/d\mathbb{P}_X$.

References

[Adachi, 2014] Adachi, T. (2014). Toward categorical risk measure theory. *Theory and Applications of Categories*, 29(14):389–405.

[Franz, 2003] Franz, U. (2003). What is stochastic independence? In *Quantum probability and White Noise Analysis, Non-commutativity, Infinite-dimensionality, and Probability at the Crossroads*, pages 254–274. World Sci. Publishing.

[Giry, 1982] Giry, M. (1982). A categorical approach to probability theory. In Banaschewski, B., editor, *Categorical Aspects of Topology and Analysis*, volume 915 of *Lecture Notes in Mathematics*, pages 68–85. Springer-Verlag.

[Lawvere, 1962] Lawvere, F. W. (1962). The category of probabilistic mappings. Preprint.