Infinite products and zero–one laws in categorical probability

Eigil Fjeldgren Rischel[1] and Tobias Fritz[2]

[1]Deparment of Mathematical Sciences, University of Copenhagen, 2100 Denmark

[2]Perimeter Institute for Theoretical Physics, N2L 2Y5, Waterloo, Ontario, Canada

2020-07-09

Markov categories are a recent category-theoretic approach to the foundations of probability and statistics. Here we develop this approach further by treating infinite products and the Kolmogorov extension theorem. This is relevant for all aspects of probability theory in which infinitely many random variables appear at a time. These infinite tensor products $\bigotimes_{i \in J} X_i$ come in two versions: a weaker but more general one for families of objects $(X_i)_{i \in J}$ in semicartesian symmetric monoidal categories, and a stronger but more specific one for families of objects in Markov categories.

As a first application, we state and prove versions of the zero–one laws of Kolmogorov and Hewitt–Savage for Markov categories. This gives general versions of these results which can be instantiated not only in measure-theoretic probability, where they specialize to the standard ones in the setting of standard Borel spaces, but also in other contexts.

Contents

1 Introduction

Markov categories are an approach to the foundations of probability and statistics based on category theory, proposed first by Golubtsov [10], rediscovered recently independently by Cho and Jacobs [3], and developed extensively by the first-named author in [6]. The basic observation is that Markov kernels can be composed sequentially and in parallel, making them into a symmetric monoidal category Stoch. One can then find axioms which make a symmetric monoidal category behave similarly to the actual category of Markov kernels, and then state and prove results from probability and statistics in this general abstract context. Thus Markov categories are abstract versions of the category of Markov kernels, and hence their name.

Eigil Fjeldgren Rischel: ayegill@gmail.com

Tobias Fritz: tfritz@pitp.ca

This represents a *synthetic* approach to probability, in contrast with the usual measure-theoretic approach (which can be called *analytic*). Instead of defining *what* the words probability, distribution, random variable, etc, mean semantically, we instead state by which operations these objects can be combined and related to each other, and which properties they are supposed to satisfy with respect to each other. We can draw an analogy to the Euclidean approach to geometry: in this synthetic approach, points and lines are described only in terms of their relationships with one another. This is to be contrasted with the analytic approach often attributed to Descartes, in which geometric objects are defined concretely in terms of more primitive notions like sets and real numbers.

In this paper, we give a treatment of further fundamental and classical results of probability theory within the setting of Markov categories. This adds further support to the hypothesis that Markov categories are an adequate setting for the foundations of probability and statistics, adding to the existing treatments of Bayesian updating, almost sure equality, conditional independence, and theorems on sufficient statistics [3, 6]. Concretely, we develop a notion of infinite tensor products in Markov categories and then apply this notion to state and prove two of the standard zero–one laws of probability theory.

Summary. We now present a more detailed overview.

Section 2 presents some background by recalling the definition of Markov category, presenting some pertinent examples, and discussing those aspects of the existing theory of Markov categories that will be of relevance for the rest of the paper. With the exception of Example 2.8, there is nothing new in this section.

Section 3 then proceeds by introducing infinite tensor products in semicartesian symmetric monoidal categories, which are exactly those symmetric monoidal categories in which the monoidal unit object is terminal. We consider a number of examples of such categories and discuss whether infinite tensor products exist. This already includes examples relevant for probability theory, and in particular BorelStoch, the category of standard Borel spaces and measurable maps[1], for which we explain how countably infinite tensor products implement the Kolmogorov extension theorem (for countably many variables); the extension theorem states that the joint distributions of infinitely many random variables are in bijection with compatible families of joint distributions for each finite subset of variables. The universal property of our infinite tensor products makes the extension theorem into a definition, generalized from distributions to Markov kernels. An important subtlety here is the additional preservation condition in Definition 3.1, where infinite tensor products are introduced: this additional condition guarantees that infinite tensor products have nice compositionality properties, so that e.g. a tensor product of two infinite tensor products is itself an infinite tensor product (Lemma 3.9). This condition does not come up traditionally in the context of the Kolmogorov extension theorem, the reason being that it automatically holds in the case of BorelStoch (Example 3.6).

In the short Section 4, we then turn to infinite tensor products in Markov categories, which we call *Kolmogorov products* due to the connection with the Kolmogorov extension theorem. In addition to being an infinite tensor product, a Kolmogorov product also needs to respect the additional structure which makes a semicartesian symmetric monoidal category into a Markov category, namely the comultiplications $\mathsf{copy}_X : X \to X \otimes X$. This is implemented by postulating that the product projections should be deterministic morphisms, or equivalently that the infinite product should be an honest categorical product in the cartesian monoidal subcategory of deterministic morphisms (Proposition 4.3). Not surprisingly, the infinite tensor products in BorelStoch, which implement the Kolmogorov extension theorem, are indeed Kolmogorov products.

In Section 5, we state and prove the zero–one laws of Kolmogorov and Hewitt–Savage in their synthetic versions. Very roughly, the Kolmogorov zero–one law states that, given an infinite collection of independent random variables and an event which is in the σ-algebra generated by the variables but independent of any finite subset of them, the event has probability zero or one. This theorem goes back to Kolmogorov's foundational monograph [14], although the statement appearing there is significantly different from its modern form. For a more recent textbook treatment of this theorem, see e.g. [13, Theorem 2.37]

---

[1]Or equivalently Polish spaces and measurable maps.

The Hewitt–Savage zero–one law, due to Hewitt and Savage in [11, Theorem 11.3], is similar in flavour but technically quite different. Here, the infinitely many variables are assumed to be identically distributed in addition to being independent, and the event is assumed to be invariant under finite permutations of the variables. Again, under these assumptions, the event has probability zero or one.

Finally in Section 6, we apply our results to BorelStoch. We thereby rederive the standard zero–one laws for random variables in standard Borel spaces from our abstract theorems. We also instantiate one of our theorems in the Kleisli category of the lower Vietoris monad, which gives criteria for when a continuous map out of infinite product of topological spaces into a Hausdorff space is constant (Corollary 6.5).

Notation. Throughout, C is a symmetric monoidal category. To simplify notation, we generally omit mention of the structure isomorphisms by assuming that C is strict without loss of generality. We routinely make use of string diagram notation, and in doing so we omit object labels whenever these are obvious from the context. Our string diagrams are to be read from bottom to top.

In Section 3, C denotes more concretely a semicartesian strict symmetric monoidal category. In Section 4 and after, C denotes a Markov category in the sense of Definition 2.5. J is a set used for indexing products of objects; the definitions and results of this paper are nontrivial only when J is infinite. F denotes either an arbitrary finite set, or more concretely a finite subset of J.

2 Background on Markov categories

We assume familiarity with symmetric monoidal categories, up to and including string diagram notation and string diagram calculus for composite morphisms in symmetric monoidal categories.

A symmetric monoidal category C is *semicartesian* if the monoidal unit $I \in$ C is terminal. Equivalently [8, Theorem 3.5], C is semicartesian if it comes equipped with morphisms

$$X \otimes Y \longrightarrow X, \qquad X \otimes Y \longrightarrow Y, \tag{1}$$

which are natural in X and Y and coincide with the monoidal structure isomorphisms whenever $X = I$ or $Y = I$. In the probability context, these morphisms can be interpreted as marginalizations, or equivalently as discarding the value of the variable Y or X.

Example 2.1. Any Cartesian monoidal category is obviously semicartesian. In the probabilistic setting that we are considering, this corresponds to a situation in which there is no randomness (in such a category, all morphisms are deterministic, cf. Definition 2.6). Hence it is somewhat uninteresting from our point of view.

Here are the main examples of semicartesian symmetric monoidal categories that we will be considering.

Example 2.2. FinStoch is the category of finite sets with *stochastic matrices* as morphisms. This means that for finite sets X and Y, a morphism $f : X \to Y$ is a matrix $(f_{xy})_{x \in X, y \in Y}$ of nonnegative real numbers such that $\sum_y f_{xy} = 1$ for every x. Stochastic matrices compose by matrix multiplication, which in this context is also known as the *Chapman–Kolmogorov equation*. We consider FinStoch as a symmetric monoidal category with respect to the cartesian product of sets on objects and the tensor product (Kronecker product) of stochastic matrices on morphisms; the monoidal structure isomorphisms are the obvious ones coming from the embedding FinSet $\subseteq$ FinStoch.

Example 2.3. Let CRing$_+$ be the category of commutative rings with maps which are merely additive and unit-preserving, considered as a symmetric monoidal category in the obvious way with respect to the tensor product of rings. Then since CRing$_+$ has the monoidal unit $\mathbb{Z}$ as its initial object, we can conclude that CRing$_+^{\mathrm{op}}$ is semicartesian monoidal.

Example 2.4. Stoch is the category of measurable spaces (X, Σ_X) as objects and Markov kernels as morphisms. We sketch the definition here and refer to [6, Section 4] and references therein for the details. Given measurable spaces (X, Σ_X) and (Y, Σ_Y), a *Markov kernel* $(X, \Sigma_X) \to (Y, \Sigma_Y)$ is a map

$$f \,:\, \Sigma_Y \times X \longrightarrow [0,1], \quad (S, x) \longmapsto f(S|x)$$

assigning to every $x \in X$ a probability measure $f(-|x) : \Sigma_Y \to [0,1]$ in such a way that for every $S \in \Sigma_Y$, the function $x \mapsto f(S|x)$ is measurable. Intuitively, one can think of f as assigning to every $x \in X$ a random element of Y; or one can think of f as a statistical model, specifying one probability measure $f(-|x)$ for every parameter value x.

Composition of Markov kernels is again defined by a variant of the Chapman–Kolmogorov equation: for $f : (X, \Sigma_X) \to (Y, \Sigma_Y)$ and $g : (Y, \Sigma_Y) \to (Z, \Sigma_Z)$, the composite is given by, for all $x \in X$ and $T \in \Sigma_Z$,

$$(gf)(T|x) := \int_{y \in Y} g(T|y)\, f(dy|x),$$

where the relevant measurability condition in order for this to define a Markov kernel follows by a standard application of the π-λ-theorem; associativity of composition uses Fubini's theorem. In this way, we obtain a category Stoch which is, more or less by definition, isomorphic to the Kleisli category of the Giry monad [9].

The symmetric monoidal structure on Stoch is defined in terms of the usual product of measurable spaces, given by the Cartesian product of underlying sets equipped with the product σ-algebra. The tensor product of morphisms $f : (A, \Sigma_A) \to (X, \Sigma_X)$ and $g : (B, \Sigma_B) \to (Y, \Sigma_Y)$ is given by the unique Markov kernel between the product measurable spaces which satisfies

$$(f \otimes g)(S \times T|a, b) = f(S|a)\, g(T|b) \qquad \forall a \in A,\, b \in B,\, S \in \Sigma_X,\, T \in \Sigma_Y.$$

The monoidal unit I is given by any one-element set with its unique σ-algebra. As a special case, morphisms $I \to (X, \Sigma_X)$ can be identified with probability measures $\Sigma_X \to [0,1]$.

There is a well-known subclass of particularly well-behaved measurable spaces, the *standard Borel spaces*, or equivalently Polish spaces equipped with their Borel σ-algebras. We write BorelStoch $\subseteq$ Stoch for the full subcategory of standard Borel spaces with Markov kernels. Since finite products of standard Borel spaces are again standard Borel, BorelStoch is again semicartesian symmetric monoidal. In summary, BorelStoch is the category of standard Borel spaces as objects with measurable Markov kernels as morphisms.

Returning to the general theory, in terms of string diagrams the unique morphism $\mathsf{del}_X : X \to I$ for an object $X \in \mathsf{C}$ is denoted by

The projection maps $\mathsf{id} \otimes \mathsf{del}_Y : X \otimes Y \to X$ and $\mathsf{del}_X \otimes \mathsf{id} : X \otimes Y \to Y$, which are the ones from (1), are correspondingly written as

Our goal is to use semicartesian monoidal categories in order to develop aspects of probability theory in categorical terms, in such a way that instantiating this theory in Stoch or BorelStoch recovers the standard theory. As it turns out, doing so requires a bit more structure, in a form which has been axiomatized first by Cho and Jacobs [3] as *affine CD-categories*, although very similar definitions occur in earlier work of Golubtsov [10]. We here follow the more intuitive terminology of our own [6, Definition 2.1].

Definition 2.5. *A* Markov category C *is a semicartesian symmetric monoidal category where every object* $X \in \mathsf{C}$ *is equipped with a distinguished morphism*

$$\mathsf{copy}_X \quad = \quad \tag{2}$$

which, together with $\mathsf{del}_X : X \to I$, *makes X into a commutative comonoid, and such that*

$$\text{(diagram)} \qquad (3)$$

for all $X, Y \in \mathsf{C}$.

There is a strictification result [6, Theorem 10.17] which guarantees that C can be assumed to be strict monoidal without messing up the commutative comonoid structures. We assume from now on that C is strict throughout.

We think of the comultiplication (2) as a copying operation. The coassociativity and counitality conditions guarantee that copying with any number of output wires is well-defined, and we draw it likewise as a single black dot with any number of outgoing wires, like this:

Examples 2.2 to 2.4 are all Markov categories in a canonical way. In $\mathsf{FinStoch}$ and Stoch, the comultiplication morphisms are given indeed by copying, i.e. by those stochastic matrices or Markov kernels which map an element $x \in X$ to the Dirac delta measure at $(x, x) \in X \times X$. We refer to [6, Example 2.5 and Section 4] for the technical details. In $\mathsf{CRing}_+^{\mathrm{op}}$, we define the comultiplication of an object $R \in \mathsf{CRing}_+^{\mathrm{op}}$ to be represented by the multiplication map $R \otimes R \to R$. The multiplicativity condition (3) then amounts to the fact that the multiplication on a tensor product of commutative rings $R \otimes S$ is given by the defining equation

$$(r_1 \otimes s_1)(r_2 \otimes s_2) = r_1 r_2 \otimes s_1 s_2.$$

The following definition is among the central notions of the theory of Markov categories.

Definition 2.6 ([6, Definition 10.1])**.** *Let C be a Markov category. A morphism $f : X \to Y$ in C is* deterministic *if it is a comonoid homomorphism,*
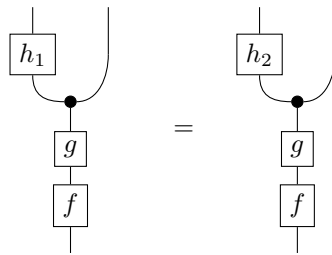
As per [6, Remark 10.13], the deterministic morphisms form a symmetric monoidal subcategory $\mathsf{C}_{\det} \subseteq \mathsf{C}$ which contains all structure morphisms of C, including the comultiplications themselves. The fact that all morphisms are comonoid homomorphisms implies that $\mathsf{C}_{\det}$ is actually cartesian monoidal.
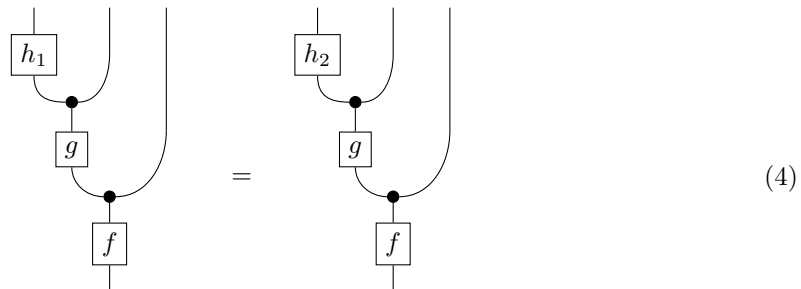
In $\mathsf{FinStoch}$, the deterministic morphisms $X \to Y$ are exactly the $\{0, 1\}$-valued stochastic matrices, which are indeed those which are deterministic in the sense that they do not involve any randomness. Since these stochastic matrices are in obvious bijection with honest functions $X \to Y$, we conclude that $\mathsf{FinStoch}_{\det}$ is equivalent to FinSet. In Stoch, the deterministic morphisms $f : X \to Y$ are exactly those Markov kernels for which $f(S|x) \in \{0, 1\}$ for every $S \in \Sigma_Y$ and $x \in X$. In other words, as the term suggests, the deterministic morphisms are again those which do not involve any randomness; in $\mathsf{BorelStoch}$, the deterministic morphisms can be identified with the measurable functions. In $\mathsf{CRing}_+^{\mathrm{op}}$, the deterministic morphisms $R \to S$ are precisely those morphisms which are represented by additive unital maps $S \to R$ which are in addition multiplicative, i.e. the ring homomorphisms. Thus the subcategory of deterministic morphisms is exactly the opposite of the usual category of commutative rings.

We recall some more definitions that we will be using later on.

Definition 2.7 ([6, Definition 11.31])**.** *A Markov category* C *is* causal *if whenever*

holds for morphisms as indicated, then also

$$(4)$$

Intuitively, this condition states that if the choice between h_1 and h_2 is irrelevant for what happens "in the future of g", then that choice is likewise irrelevant for what happened "in the past of g". It is known that Stoch is causal [6, Example 11.35], and therefore so are the subcategories BorelStoch and FinStoch.

Example 2.8. $\mathsf{CRing}_+^{\mathrm{op}}$ is *not* causal. To see this, consider the ring $\mathbb{Z}[t]$, where the following counterexample similar to [6, Example 11.33] can be formulated. Consider the additive unital maps given as the additive extensions of

$$f(t^n) = \begin{cases} t^{n-1} & n \geq 1 \\ 1 & n = 0 \end{cases}$$

$$g(t^n) = \begin{cases} t & n \geq 1 \\ 1 & n = 0 \end{cases}$$

$$h_1(t^n) = t^n$$

$$h_2(t^n) = 1$$

On the basis monomials t^n, we have $(fg)(t^n) = 1$, and therefore
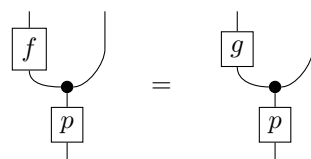
$$(fg)(h_1(t^n)t^m) = 1 = (fg)(h_2(t^n)t^m)$$

for all $n, m \in \mathbb{N}$, so that the hypothesis holds. If $\mathsf{CRing}_+^{\mathrm{op}}$ were causal, then we would also have

$$f(g(h_1(t^n)t^m)t^\ell) = f(g(h_2(t^n)t^m)t^\ell)$$

for all $n, m, \ell \in \mathbb{N}$. But this clearly fails for $n = \ell = 1$ and $m = 0$.

The following definition goes back to Cho and Jacobs [3, Definition 5.1]. A more detailed investigation of its properties can be found in [6, Section 13].

Definition 2.9. *Given morphisms* $p : A \to X$ *and* $f, g : X \to Y$ *in a Markov category, we say that* f *and* g *are* p-a.s. equal *(or* p-almost surely equal) *if*

and we also write this more concisely as $f =_{p\text{-a.s.}} g$.

This term comes from measure theory, and should be interpreted as meaning that for all those values of X which can occur as outputs of p, feeding these values to f and g as inputs is guaranteed to result in the same distribution; at the same time, the behavior of f and g may be different on those input values which do not occur as outputs of p.

We finally introduce a notion of conditional independence, which is a straightforward generalization of [6, Defnition 12.12] from the binary case to the n-ary case.

Definition 2.10. *Given a morphism* $p : A \to X_1 \otimes \ldots \otimes X_n$ *in a Markov category, we say that* p *displays the conditional independence* $X_1 \perp \ldots \perp X_n \mathbin{\|} A$ *if the equation*

holds.

Conditional independence does not depend on the order of the tensor factors: if p displays the conditional independence $X_1 \perp \ldots \perp X_n \mathbin{\|} A$, then composing p with any permutation σ of the tensor factors $X_1, \ldots, X_n$ gives a morphism $A \to X_{\sigma(1)} \otimes \ldots \otimes X_{\sigma(n)}$ which displays the conditional independence $X_{\sigma(1)} \perp \ldots \perp X_{\sigma(n)} \mathbin{\|} A$. Therefore if $\bigotimes_{j \in F} X_j$ is a finite product without any particular order on the factors[2], then for $p : A \to \bigotimes_{j \in F} X_j$ there is no ambiguity about whether p displays conditional independence or not, and we also write $\perp_{j \in F} X_j \mathbin{\|} A$ in case that it does.

3 Infinite tensor products in semicartesian symmetric monoidal categories

Many theorems in probability theory involve infinitely many random variables at a time, together with a joint distribution for them. In the setting of Markov categories, this means that we need to consider morphisms whose codomain is an "infinite tensor product" in a suitable sense. In this section, we start by considering infinite tensor products in the more general context of semicartesian symmetric monoidal categories.

Recall that the Kolmogorov's extension theorem states that joint distributions of a given (infinite) family of random variables are in bijection with compatible families of distributions of finite subsets of these variables. Here, a joint distribution of all of them determines such a compatible family by considering all its marginals on finite subsets. This motivates our general definition of infinite tensor products, which we now introduce. Suppose that $(X_i)_{i \in J}$ is any family of objects, where the indexing set J is typically infinite. For any finite subset $F \subseteq J$, we also write $X_F := \bigotimes_{i \in F} X_i$ for simplicity of notation. If $F \subseteq F' \subseteq J$ are two finite subsets, then the fact that C is semicartesian monoidal gives us *marginalization morphisms*

$$\pi_{F',F} : X_{F'} \longrightarrow X_F.$$

Via these maps, the finite tensor products $(X_F)_{F \subseteq J}$ make up a cofiltered diagram in the form of a functor from the poset of finite subsets of J, ordered by reverse inclusion, to C.

Definition 3.1. *Let* $(X_i)_{i \in J}$ *be a family of objects in* C. *An* infinite tensor product

$$X_J := \bigotimes_{i \in J} X_i$$

---

[2]See e.g. [4, Proposition II.1.5] for how to make sense of tensor products in symmetric monoidal categories without a prescribed order of the factors.

is the limit of the diagram $F \mapsto X_F$, indexed by the poset of finite subsets $F \subseteq J$ ordered by reverse inclusion, if this limit exists and is preserved by the functor $- \otimes Y$ for every object $Y \in \mathsf{C}$.

We will refer to the structure maps $\pi_F : X_J \to X_F$ as *finite marginalizations*. The extra preservation means more concretely that $X_J \otimes Y$ is likewise the cofiltered limit of the $X_F \otimes Y$ with respect to the maps

$$\pi_F \otimes \mathsf{id}_Y \;:\; X_J \otimes Y \longrightarrow X_F \otimes Y.$$

We will motivate this preservation condition in Remark 3.3.

Of course, an infinite tensor product may or may not exist in C. Example 3.7 will present a simple example in which infinite tensor products only exist for *finite J*.

Remark 3.2. If J is finite, then the infinite tensor product always exists and coincides with the tensor product $X_J = \bigotimes_{i \in J} X_i$ specified by the monoidal structure: then the finite product itself is initial in the defining diagram of the limit, and the preservation condition holds automatically for the same reason. Thus infinite tensor products $\bigotimes_{i \in J} X_i$ are of interest only for infinite J, as the terminology suggests. And for finite J, the notation is such that $\bigotimes_{i \in J} X_i$ can be interpreted either as specified directly by the monoidal structure, or as an infinite tensor product in the sense of Definition 3.1. Since the result is the same (up to unique isomorphism), there is no ambiguity and the notation is consistent.

Remark 3.3. It is natural to require different infinite tensor products to interact well with another. If X_J is an infinite tensor product of a family $(X_i)_{i \in J}$ and we are given an additional object X_* for $* \notin J$, then there is a canonical comparison isomorphism

$$X_J \otimes X_* \longrightarrow X_{J \sqcup \{*\}} \tag{5}$$

induced via the universal property of $X_{J \sqcup \{*\}}$, which is the infinite tensor product of the original family with X_* thrown in. In order for infinite tensor products to be coherent, one will want this comparison isomorphism to be an isomorphism. We now show that this is indeed the case thanks to the preservation condition in Definition 3.1, even if $X_{J \sqcup \{*\}}$ does not exist a priori.

More precisely, we exhibit $X_J \otimes X_*$ as the defining cofiltered limit of the infinite tensor product $X_{J \sqcup \{*\}}$ as follows. The object $X_J \otimes X_*$ can be equipped with finite marginalization morphisms with respect to $F \subseteq J \sqcup \{*\}$ given by

$$\pi_{F \setminus \{*\}} \otimes \mathsf{id} \;:\; X_J \otimes X_* \longrightarrow X_{F \setminus \{*\}} \otimes X_*$$

whenever $* \in F$, and just by

$$\pi_F \otimes \mathsf{del}_{X_*} : X_J \otimes X_* \longrightarrow X_F$$

in case that $* \notin F$. It is straightforward to check that these morphisms indeed make $X_J \otimes X_*$ carry the universal property of the infinite tensor product $X_{J \sqcup \{*\}}$ if and only if the functor $- \otimes X_*$ preserves the defining cofiltered limit of the infinite tensor product X_J. It is also easy to see that this new cofiltered limit is again preserved by every functor $- \otimes Y$, based on the assumption that the defining cofiltered limit of the original infinite tensor product is preserved by $- \otimes (X_* \otimes Y)$ as well.

In summary, the preservation condition in this definition amounts to the requirement that the canonical comparison morphism (5) must be an isomorphism, as one would expect intuitively from a good notion of infinite tensor product.

Example 3.4. The dual definition of infinite tensor products of algebraic structures as *filtered colimits* of finite tensor products is well-known in the literature, e.g. in the case of C*-algebras [2, p. 315], although we do not know whether the preservation condition has been made explicit before. For example in the case where our category is $\mathsf{CRing}_+^{\mathrm{op}}$, we recover the usual folklore definition of an infinite tensor product of rings $\bigotimes_{i \in J} R_i$ in terms of formal sums of elementary tensors, where an elementary tensor is a family of elements $(r_i)_{i \in J}$ such that all but finitely many are equal to the respective unit. The necessary preservation condition is easily seen to hold. Note that we do not yet consider the multiplication on $\bigotimes_{i \in J} R_i$, since Definition 3.1 is not yet concerned with the comonoid structures on the objects. We will get to this in the next section.

Intuitively, our definition matches up with the known algebraic ones under the categorical duality of algebra and geometry, where our Definition 3.1 is on the geometrical side of the duality.

Remark 3.5. In those semicartesian monoidal categories C that are of interest to us, morphisms $I \to X$ play the role of *probability measures* on X. Thus applying the defining universal property of an infinite tensor product with respect to maps out of I implements the *Kolmogorov extension theorem* in C: probability measures on an infinite product X_J are in bijection with consistent families of probability measures on the finite products X_F for $F \subseteq J$. The spirit of Definition 3.1 is thus such that it turns the Kolmogorov extension theorem into a definition.

Example 3.6. We now consider infinite tensor products in Stoch, where we would like infinite tensor products to be given by the corresponding infinite products of measurable spaces in the usual sense. Since the Kolmogorov extension theorem does not hold for general (even merely countable) products of measurable spaces [1], this is not the case without further additional assumptions on the measurable spaces involved. Thus we do not know whether Stoch has infinite tensor products, although we suspect that it does not; but even in the unlikely case that it does, they are in general not the ones that one will care about for the purposes of probability theory.

However, the situation is much better for countable products in $\mathsf{BorelStoch}$, for which the Kolmogorov extension theorem indeed holds [13, Theorem 14.35]. In other words, if $((X_i, \Sigma_i))_{i \in \mathbb{N}}$ is a sequence of standard Borel spaces, then the cartesian product $X_{\mathbb{N}} = \prod_{i \in \mathbb{N}} X_i$, when equipped with the product σ-algebra $\Sigma_{\mathbb{N}}$, satisfies the universal property of an infinite tensor product with respect to maps out of I, since the marginalization maps implement a bijection between probability measures on $X_{\mathbb{N}} = \prod_{i \in \mathbb{N}} X_i$ and compatible families of probability measures on the finite subproducts $X_F = \prod_{i \in F} X_i$, where the compatibility is with respect to marginalization to smaller subproducts specified by $F' \subseteq F$.

We now prove that this implies the universal property in general: if A and all the X_i are standard Borel spaces, then Markov kernels $A \to X_J$ are in bijection with compatible families of Markov kernels $A \to X_F$. While we just saw that this is the case for $A = I$, we now show that it holds for arbitrary A. Thus suppose that

$$\Big(g_F : (A, \Sigma_A) \longrightarrow (X_F, \Sigma_F)\Big)_{F \subseteq J \text{ finite}}$$

is a family of Markov kernels satisfying the compatibility condition $\pi_{F',F} \circ g_{F'} = g_F$ for all finite $F \subseteq F' \subseteq J$. Then for every $a \in A$, the probability measures $g_F(-|a) : \Sigma_F \to [0,1]$ are a compatible family to which the Kolmogorov extension theorem in the form of [13, Theorem 14.35] applies, and we obtain a unique probability measure $g_J(-|a) : \Sigma_J \to [0,1]$ which has the original $g_F(-|a)$ as its finite marginals. It remains to be shown that for every $S \in \Sigma_J$, the map $a \mapsto g_J(S|a)$ is measurable. Since limits of pointwise convergent sequences of measurable real-valued functions are again measurable, the set of S for which this measurability holds is closed under countable disjoint union, and it is clearly closed under complements. We therefore have a λ-system. Since the map is measurable by assumption whenever S is a measurable cylinder set, and the cylinder sets form a π-system, the π-λ-theorem implies that $a \mapsto g_J(S|a)$ is measurable for all S in the σ-algebra generated by the measurable cylinder sets, which is exactly the product σ-algebra Σ_J. Hence the existence part of the universal property indeed holds; and the uniqueness part is obvious by considering each $a \in A$ separately and applying the Kolmogorov extension theorem.

Since the resulting comparison morphism (5) is an isomorphism by construction—both sides are given by the corresponding products of measurable spaces—it follows that the required preservation condition of infinite tensor products holds as well.

In conclusion, $\mathsf{BorelStoch}$ indeed has countable tensor products. It is plausible that there is another subcategory of Stoch strictly larger than $\mathsf{BorelStoch}$ which has all infinite tensor products; for example, one can try to construct such a subcategory by imposing compactness or perfectness conditions [5, §451] on the Markov kernels, since under such assumption one again has a version of the Kolmogorov extension theorem [5, Corollary 454G]. However, we have so far not been able to find a definition of "compact Markov kernel" or "perfect Markov kernel" which would guarantee closure under composition, due to problems with perfect measures not being stable under mixtures [15]. Hence the problem of finding a subcategory of Stoch with all infinite tensor products remains open.

Example 3.7. Infinite tensor products never exist in $\mathsf{FinStoch}$, in the following sense: if the family $(X_i)_{i \in J}$ is such that no X_i is empty and infinitely many of them contain at least two elements,

then $\bigotimes_i X_i$ does not exist. One way to see this is to use the fact that the hom-sets of FinStoch are convex subsets of the vector space of matrices, and that composition distributes over these convex combinations, making FinStoch into a category enriched in convex sets.[3] Since these hom-sets are finite-dimensional, for every Y and Z there is $n \in \mathbb{N}$ such that among every n morphisms $Y \to Z$, one of them can be written as a convex combination of the others. Now suppose that the product $\bigotimes_i X_i$ existed. Then by choosing varying elements of each X_j, we could construct uncountably many morphisms $1 \to \bigotimes_i X_i$ whose marginalizations $1 \to X_F$ are all deterministic. Per the above, one of these hypothetical morphisms can be written as a convex combination of finitely many others. By choosing $F \subseteq J$ suitably, we can achieve that the finite marginalizations $1 \to X_F$ of the finitely many morphisms involved in this convex combination are all distinct. Since these morphisms $1 \to X_F$ are deterministic by construction, and no deterministic morphism in FinStoch can be written as a convex combination of other deterministic morphisms, we have arrived at a contradiction. Thus $\bigotimes_i X_i$ does not exist.

Example 3.8 (Paolo Perrone, personal communication)**.** Suppose that C is a partially ordered set, considered as a category with $X \geq Y$ if and only if the morphism $X \to Y$ exists.[4] Then the monoidal structure makes C into a preordered commutative monoid, which we write additively. Here, the functoriality of the monoidal structure amounts to the assumption that addition is monotone,

$$X \geq Y \quad \Longrightarrow \quad X + Z \geq Y + Z.$$

Furthermore, the monoidal structure is semicartesian if and only if $X \geq 0$ for every X. One can then apply Lemma 3.10 to show that C has infinite tensor products if and only if C is a *finitary complete monoid* in the sense of Goldstern and Karner [12, Section 2.2], and these infinite tensor products then correspond to infinite sums $\sum_{i \in J} X_i$ in the monoid.

We return to the general theory.

Lemma 3.9. *Let $J = J_1 \sqcup J_2$ be a disjoint union and $(X_i)_{i \in J}$ a family of objects in C. Suppose that the infinite tensor products X_{J_1} and X_{J_2} exist. Then the object*

$$X_{J_1} \otimes X_{J_2}$$

is an infinite tensor product X_J with respect to the finite marginalizations morphisms given by, for every finite $F \subseteq J$,

$$\rho_F \ : \ X_{J_1} \otimes X_{J_2} \xrightarrow{\pi_{F \cap J_1} \otimes \pi_{F \cap J_2}} X_{F \cap J_1} \otimes X_{F \cap J_2}.$$

Proof. For fixed finite $F_1 \subseteq J_1$, the morphisms

$$\mathsf{id} \otimes \pi_{F_2} \ : \ X_{F_1} \otimes X_{J_2} \longrightarrow X_{F_1} \otimes X_{F_2}$$

exhibit $X_{F_1} \otimes X_{J_2}$ as the cofiltered limit of the $X_{F_1} \otimes X_{F_2}$ by the preservation assumption. Similarly for varying F_1, the morphisms

$$\pi_{F_1} \otimes \mathsf{id} \ : \ X_{J_1} \otimes X_{J_2} \longrightarrow X_{F_1} \otimes X_{J_2}$$

exhibit $X_{J_1} \otimes X_{J_2}$ as the cofiltered limit of the $X_{F_1} \otimes X_{J_2}$. The claimed universal property follows since a limit of limits is a limit, and it is easy to see that the diagram shapes match up: the poset of finite subsets of $J_1 \sqcup J_2$ is the product of the posets of finite subsets of J_1 and J_2. The preservation property with respect to applying a tensoring functor $- \otimes Y$ carries along both of the above limits individually. $\qquad\square$

More generally, we may consider infinite tensor products of infinite tensor products, $\bigotimes_{k \in K} \bigotimes_{j \in J_k} X_{k,j}$, where now $(J_k)_{k \in K}$ is a family of sets and $(X_{k,j})$ a doubly indexed family of objects. One may

---

[4]This direction of the ordering is opposite to the usual convention, but necessary for the following to work without reversing direction.

think that this doubly infinite tensor product should be isomorphic to the single-step infinite tensor product $\bigotimes_{k \in K, \, j \in J_k} X_{k,j}$. And indeed, if for every finite $F \subseteq \prod_{k \in K} J_k$ we choose any finite $G \subseteq K$ such that $(k, j) \in F$ implies $k \in G$, then we have morphisms

$$\rho_F \; : \; \bigotimes_{k \in K} \bigotimes_{j \in J_k} X_{k,j} \xrightarrow{\;\pi_G\;} \bigotimes_{k \in G} \bigotimes_{j \in J_k} X_{k,j} \xrightarrow{\;\bigotimes_{k \in G} \pi_{F \cap J_k}\;} \bigotimes_{k \in G, \, j \in F \cap J_k} X_{k,j} \qquad (6)$$

where the object on the right can also be written as $\bigotimes_{(k,j) \in F} X_{k,j}$. It is straightforward to see that this ρ_F does not depend on the particular choice of G (up to monoidal structure isomorphisms). We now show that these ρ_F's are finite marginalization morphisms which make the doubly infinite tensor product $\bigotimes_{k \in K} \bigotimes_{j \in J_k} X_{k,j}$ into the single infinite tensor product $\bigotimes_{k \in K, \, j \in J_k} X_{k,j}$.

Lemma 3.10. *Suppose that* $(J_k)_{k \in K}$ *is a family of sets and* $(X_{k,j})_{j \in J_k, k \in K}$ *a doubly indexed family of objects. If all of the infinite tensor products* $\bigotimes_{j \in J_k} X_{j,k}$ *and* $\bigotimes_{k \in K} \bigotimes_{j \in J_k} X_{k,j}$ *exist, then the morphisms* (6) *exhibit* $\bigotimes_{k \in K} \bigotimes_{j \in J_k} X_{k,j}$ *as the infinite tensor product* $\bigotimes_{k \in K, \, j \in J_k} X_{k,j}$.

Proof. By repeated application of Lemma 3.9, the second half of (6) exhibits $\bigotimes_{k \in G} \bigotimes_{j \in J_k} X_{k,j}$ as the infinite tensor product of the $X_{k,j}$ for $k \in G$ and $j \in J_k$. The relevant universal property as well as its preservation by $- \otimes Y$ then follow from another straightforward diagram chase. $\qquad \square$

4 Infinite tensor products in Markov categories

If a Markov category C has infinite tensor products, then one should expect a compatibility condition between these infinite tensor products and the comonoid structures on the objects. This is imposed as follows.

Definition 4.1. *Let* C *be a Markov category and* $(X_i)_{i \in J}$ *a family of objects. We say that an infinite tensor product* $X_J = \bigotimes_{i \in J} X_i$ *is a* Kolmogorov product *if the finite marginalization morphisms* $\pi_F : X_J \to X_F$ *are deterministic.*

In $\mathsf{BorelStoch}$, the infinite tensor products constructed in Example 3.6 are Kolmogorov products. In $\mathsf{CRing}_+^{\mathrm{op}}$, the infinite tensor products from Example 3.4 are Kolmogorov products as well, provided that one equips $\bigotimes_i R_i$ with the tensor product ring structure, since this one is the only one which makes the canonical inclusions $R_j \to \bigotimes_i R_i$ into ring homomorphisms. These examples illustrate the following important point. In some Markov categories, such as $\mathsf{CRing}_+^{\mathrm{op}}$, not all isomorphisms are deterministic, and there may even be isomorphic objects which have no deterministic isomorphism [6, Remark 10.10]. In these categories, whether a given infinite tensor product is a Kolmogorov product or not generally depends on the specific choice of that object. In particular, there may be infinite tensor products which are not Kolmogorov products. For example in $\mathsf{CRing}_+^{\mathrm{op}}$, we may take the tensor product of abelian groups $\bigotimes_i R_i$ and equip it with a different multiplication than the tensor product of rings one (as long as it has the same unit), and one will then have an infinite tensor product in $\mathsf{CRing}_+^{\mathrm{op}}$ which is not a Kolmogorov product.
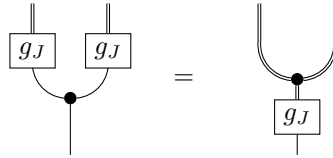
Remark 4.2. If the Markov category C is positive in the sense of [6, Definition 11.22], then a morphism $A \to B \otimes C$ is deterministic if and only if both of its marginals $A \to B$ and $A \to C$ are deterministic [6, Corollary 12.15]. In this case, it follows that the finite marginalizations π_F are deterministic if and only if all the single-factor marginalizations $\pi_{\{i\}} : X_J \to X_i$ are. Thus for an infinite tensor product to be a Kolmogorov product, it is then enough to verify that the $\pi_{\{i\}}$ are deterministic.

Using the multiplicativity of the comonoid structure of (3), we see that Lemmas 3.9 and 3.10 hold also for Kolmogorov products, in the sense that tensoring together Kolmogorov products, or taking Kolmogorov products of Kolmogorov products, results again in a Kolmogorov product.
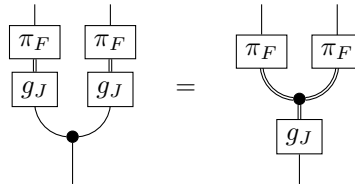
Proposition 4.3. *Every Kolmogorov product is a categorical product in the cartesian monoidal subcategory of deterministic morphisms* $\mathsf{C}_{\mathrm{det}}$.

Proof. For $(X_i)_{i \in J}$ a family of objects with Kolmogorov product X_J and finite marginalizations $\pi_F : X_J \to X_F$, we first show that if we have a compatible family of deterministic morphisms

$g_F : A \to X_F$, then also the induced $g_J : A \to X_J$ is deterministic. Drawing the Kolmogorov product X_J as a double wire, we need to prove the determinism equation

By Lemma 3.9, the codomain object $X_J \otimes X_J$ is itself an infinite tensor product, and it is therefore enough to prove that for every finite $F \subseteq J$,

where we have assumed without loss of generality that the two finite marginalization maps are the same, which we can by replacing the corresponding finite subsets by their union. Using the determinism assumption for π_F on the right, as well as the assumption that $g_F = \pi_F \circ g_J$ is deterministic, implies the claim.

Now since every finite tensor product X_F is a categorical product in $\mathsf{C}_{\det}$, we can use the fact that in every category,

$$\lim_{F \subseteq J \text{ finite}} \prod_{j \in F} X_j \cong \prod_{j \in J} X_j,$$

and the claim follows. $\qquad\square$

In particular, this proves that Definition 4.1 determines the comonoid structure on a Kolmogorov product uniquely: $\mathsf{copy}_{\bigotimes_i X_i}$ has to be the equal to the corresponding diagonal morphism in $\mathsf{C}_{\det}$, and Proposition 4.3 determines $\bigotimes_i X_i$ up to unique *deterministic* isomorphism.

Example 4.4. Let $\mathsf{SetMulti}$ be the category with objects ordinary sets, morphisms $f : X \to Y$ given by *multivalued functions*, which are maps $f : X \to \mathcal{P}(Y) \setminus \emptyset$, with composition of morphisms $f : X \to Y$ and $g : Y \to Z$ given by

$$(g \circ f)(x) := \bigcup_{y \in f(x)} g(y). \tag{7}$$

$\mathsf{SetMulti}$ could also be described as the Kleisli category of the *nonempty powerset monad*, which sends a set X to $\mathcal{P}(X) \setminus \emptyset$. In particular, $\mathsf{FinSetMulti}$ is closely related to Rel, the category of sets and relations, since the latter can equivalently be regarded as the Kleisli category of the powerset monad (where the empty set is allowed). Intuitively, $\mathsf{SetMulti}$ is the closest relative of Rel to which our formalism can be applied. See also [6, Example 2.6] for a treatment of the full subcategory $\mathsf{FinSetMulti}$ on finite sets.

The Cartesian product of sets equips $\mathsf{SetMulti}$ with a symmetric monoidal structure, if for $f : A \to X$ and $g : B \to Y$ we set

$$(f \otimes g)(a, b) := f(a) \times g(b) \subseteq X \times Y.$$

The monoidal unit is any one-element set I. With the obvious structure isomorphisms inherited from Set, this turns $\mathsf{SetMulti}$ into a semicartesian symmetric monoidal category, and we can give it a Markov category structure by setting $\mathsf{copy}_X(x) := \{(x, x)\} \subset X \times X$. Then the deterministic morphisms are precisely those sending each point into a singleton set, resulting in an isomorphism of categories $\mathsf{SetMulti}_{\det} \cong \mathsf{Set}$.

Of course, Set has all infinite products. But perhaps surprisingly, these products are not Kolmogorov products, or even infinite tensor products in $\mathsf{SetMulti}$. To see this, consider for example

the family $(X_n)_{n\in\mathbb{N}}$ with $X_n = \{0,1\}$ for all n. A map $I \to \prod_{n\in\mathbb{N}}\{0,1\}$ is just a nonempty subset of this infinite product. Given such a subset U, the finite marginals $I \to \prod_{n\in F}\{0,1\}$ for finite $F \subset \mathbb{N}$ are the images of U under the projection map. But these images do not determine U. For instance, consider the subset U of those sequences containing at least one 1. Its image in every $\prod_{n\in F}\{0,1\}$ is the whole set, since any finite sequence of 0s and 1s can be extended to one containing a 1, but U itself is a proper subset. In particular, U has the same compatible family of finite marginalizations as e.g. the set of those sequences containing at least one 0, and hence the uniqueness part of the universal property fails.

It is clear that this counterexample is quite generic and works for any infinite family of nonempty sets, infinitely many of which contain at least two elements, as in Example 3.7. Thus SetMulti is very far from having Kolmogorov products. A good way to remedy this deficiency is to consider topological spaces instead of sets: as we will see in Section 6, the Kleisli category of the lower Vietoris monad on Top does have Kolmogorov products, and it can be thought of as the topological counterpart of SetMulti.

The concept of Kolmogorov product describes a notion of infinite collection of random variables whose joint distribution is encoded in its finite marginals. It is natural to postulate the same for when such an infinite collection is considered to be independent.
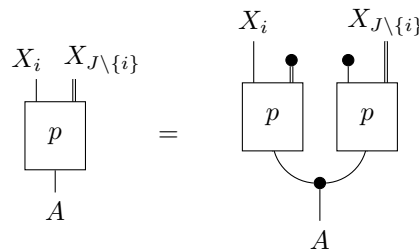
Definition 4.5. *Let $X_J = \bigotimes_{i\in J} X_i$ be a Kolmogorov product in a Markov category. We say that a morphism $p : A \to X_J$ displays the conditional independence $\perp_{i\in J} X_i \parallel A$ if each finite marginalization $p_F : A \to X_F$ displays the conditional independence $\perp_{i\in F} X_i \parallel A$.*

5 The zero–one laws

We now make use of the theory developed in the previous sections in order to state and prove synthetic versions of the classical zero–one laws of Kolmogorov and Hewitt–Savage. Throughout this section, $(X_i)_{i\in J}$ is a family of objects in a Markov category C with Kolmogorov product X_J and finite marginalizations $\pi_F : X_J \to X_F$.

Lemma 5.1 (The infinite independence lemma). *Suppose $p : A \to X_J$ displays the conditional independence $\perp_i X_i \parallel A$. Then for every $i \in J$, it also displays the independence $X_i \perp X_{J\setminus\{i\}} \parallel A$.*

Proof. This is nontrivial because the assumption by definition only involves conditional independence of the finite marginalizations. What we need to prove is that

As the codomain is exactly the Kolmogorov product X_J, it is enough to prove this equation upon composing the second output with the finite marginalizations $\pi_F : X_{J\setminus\{i\}} \to X_F$. But this is a consequence of the assumption. $\square$

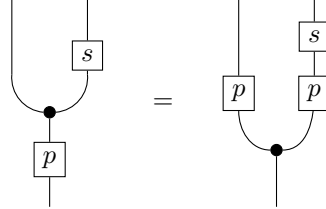**Lemma 5.2** (The determinism lemma). *Suppose that we have $p : A \to X$ and deterministic $s : X \to T$. If the joint*
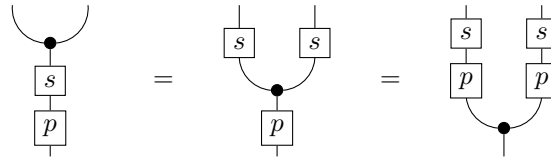

(8)

displays the conditional independence $X \perp T \parallel A$, then the composite $sp : A \to T$ is deterministic.

This result can be thought of as a partial converse to the condition which characterizes positivity of a Markov category [6, Definition 11.22]. Here and in the following, our notation "s" indicates that we think of this morphism as a statistic [6, Definition 14.2].

Proof. Since the marginals of (8) are exactly p and sp, the conditional independence assumption is

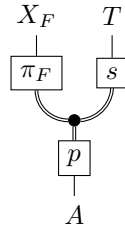Using this equation together with the assumption that s is deterministic, we then compute

which is what was to be shown. □

Here is our abstract version of Kolmogorov's zero–one law.

Theorem 5.3 (Abstract Kolmogorov zero–one law). *Suppose that C is a Markov category, and that $(X_i)_{i \in J}$ is a family of objects with Kolmogorov product X_J and finite marginalizations $\pi_F : X_J \to X_F$. Suppose that morphisms $p : A \to X_J$ and deterministic $s : X_J \to T$ satisfy the following:*

1. p displays the conditional independence $\perp_i X_i \parallel A$.

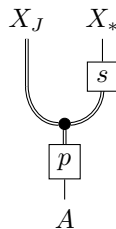*2. For every finite $F \subseteq J$, the joint*

displays the conditional independence $X_F \perp T \parallel A$.

Then the composite $sp : A \to T$ is deterministic as well.

Note that this result applies even in the case where J is finite, in which case it amounts to Lemma 5.2.

Proof. Assuming $* \notin J$, let us write $X_* := T$, so that $X_J \otimes X_*$ is the Kolmogorov product of the family $(X_i)_{i \in J \sqcup \{*\}}$. It is clear from the definitions and the assumption that the morphism

displays the conditional independence $\perp_{i \in J \sqcup \{*\}} X_i \parallel A$. By the infinite independence lemma, Lemma 5.1, this means that it also displays the conditional independence $X_* \perp X_J \parallel A$. Now the determinism lemma implies exactly what we want, namely that the composite $sp : A \to T$ is deterministic. □

We now move on to our treatment of the Hewitt–Savage zero–one law. It applies to Kolmogorov products where all factors are copies of the same object. So for an object X and the constant family $(X)_{i \in J}$, let us say that the corresponding Kolmogorov product X_J is the *Kolmogorov power* of X by J, whenever it exists.

A *finite permutation* of the set J is a bijection $\sigma : J \to J$ which fixes all but finitely many elements. Suppose that $\sigma : J \to J$ is a finite permutation or more generally any injection. Then an application of the universal property of the Kolmogorov power shows that σ induces a morphism $\hat{\sigma} : X_J \to X_J$ by applying the universal property of X_J to the compatible family of composite morphisms

$$X_J \xrightarrow{\pi_{\sigma(F)}} X_{\sigma(F)} \xrightarrow{\cong} X_F,$$

where the second morphism matches up the factors of the two finite products involved as prescribed by the bijection $\sigma|_F : F \xrightarrow{\cong} \sigma(F)$. The relevant compatibility condition is straightforward to check. Moreover, $\hat{\sigma}$ is deterministic by Proposition 4.3.

Theorem 5.4 (Abstract Hewitt–Savage zero–one law). *Suppose that* C *is a causal Markov category (Definition 2.7). Let* X_J *be a Kolmogorov power of an object* $X \in \mathsf{C}$ *with respect to an infinite set* J. *Suppose that morphisms* $p : A \to X_J$ *and deterministic* $s : X_J \to T$ *satisfy the following:*

1. *p displays the conditional independence $\perp_{i \in J} X_i \parallel A$.*

2. *For every finite permutation $\sigma : J \to J$, we have $\hat{\sigma}p = p$ and $s\hat{\sigma} = s$.*

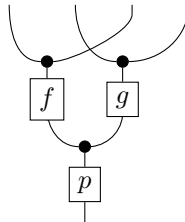*Then the composite* $sp : A \to T$ *is deterministic.*

In contrast to Theorem 5.3, we now have two relevant assumptions in addition to those on p and s, namely that C is causal and J is infinite. Before getting to the proof, we first make a useful auxiliary observation.

Lemma 5.5. *Let* C *be a causal Markov category. Let* $p : A \to X$ *and* $f, g : X \to Y$ *be morphisms in* C *with* f *deterministic and such that*

Then $f =_{p\text{-a.s.}} g$ *(Definition 2.9).*

Proof. Consider the morphism

Applying the assumed equation together with determinism of f shows that marginalizing over the first output gives the same morphism $A \to Y \otimes Y \otimes Y$ as marginalizing over the second output. Therefore, applying the causality property of Definition 2.7 with h_1 and h_2 the two marginalizations, we can conclude the equality

Marginalizing this further over the second and third outputs and truncating the redundant pieces proves the desired equation. □

Proof of Theorem 5.4. Letting a finite permutation $\sigma : J \to J$ act on the X_J part of the joint of X_J and T gives

$$(9)$$

so that the joint given by the right-hand side is invariant. Here, the first and last equation hold by the invariance assumption, and the second because $\hat{\sigma}$ is deterministic.

We next argue that this equation still holds if σ is merely an injection. By the universal property of $X_J \otimes T$ as the infinite tensor product of the family $(X_i)_{i \in J}$ together with T, it is enough to prove (9) upon marginalization by $\pi_F \otimes \mathrm{id}$ for all finite $F \subseteq J$. Since we can find a finite permutation $\sigma' : J \to J$ such that $\sigma'|_F = \sigma|_F$ for any given F, the claim follows from the finite permutation case.

Consider the morphism given by composing (9) with s on the first output. Since these are the same for all injections σ, Lemma 5.5 tells us that the morphisms $s\hat{\sigma}$ are all p-a.s. equal to s, for every injection $\sigma : J \to J$.

Now let $J = J_1 \sqcup J_2$ be a decomposition of J into two disjoint subsets having the same cardinality as J; this is the step which relies on J being infinite. Let $\tau_1, \tau_2 : J \to J$ be injections with images J_1 and J_2, respectively. Then we claim that

$$(10)$$

By Lemma 3.9, it is enough to prove this upon postcomposing with two finite marginalizations π_{F_1} and π_{F_2}. But then using the fact that τ_1 and τ_2 have disjoint image, this follows from the independence assumption, which has

as a special case.

Putting things together, we can now compute

which is what was to be shown. □

6 Examples

We now show how our results specialize to the standard Kolmogorov and Hewitt–Savage zero–one laws in the case of standard Borel spaces. Further down we will give another application to continuous maps between topological spaces.

Corollary 6.1 (Kolmogorov zero–one law)**.** *Suppose Ω is a standard Borel space with a probability measure P and that $(f_i : \Omega \to X_i)_{i \in \mathbb{N}}$ is a sequence of independent random variables taking values in standard Borel spaces X_i. Suppose that $T \subseteq \Omega$ is in the coarsest σ-algebra which makes the f_i measurable, and independent of any finite subset of the f_i. Then $P(T) \in \{0,1\}$.*

Proof. We want to apply Theorem 5.3 with $\mathsf{C} = \mathsf{BorelStoch}$ and $J = \mathbb{N}$ and $A = I$, the one-element measurable space. Then $P : I \to \Omega$. We write $f_{\mathbb{N}} : \Omega \to X_{\mathbb{N}}$ for the induced map to the product measurable space. In Ω, the coarsest σ-algebra which makes the f_i measurable coincides with the σ-algebra consisting of the sets of the form $f_{\mathbb{N}}^{-1}(S)$ for measurable $S \subseteq X_J$. Hence the assumption implies that there is measurable $S \subseteq X_J$ such that $T = f_{\mathbb{N}}^{-1}(S)$. Then we apply Theorem 5.3 with $s : X_J \to \{0,1\}$ the indicator function of S, which is a measurable map and therefore can also be regarded as a morphism in $\mathsf{BorelStoch}$, and with $p = f_{\mathbb{N}}P$. We therefore obtain that $sp : I \to \{0,1\}$ is deterministic. But since sp corresponds to the probability measure on $\{0,1\}$ with weight $P(f_{\mathbb{N}}^{-1}(S)) = P(T)$ on $\{1\}$ and the complementary weight $1 - P(T)$ on 0, we conclude that indeed $P(T) \in \{0,1\}$. □

We note that the standard Kolmogorov zero–one law holds for any measurable spaces, not just for standard Borel spaces, and hence this is a strictly less general statement. As discussed in Example 3.6, it seems unlikely that the general statement can be proven by applying our machinery to the category Stoch of measurable spaces and Markov kernels. This leads to the following problem:

Problem 6.2. *Is there a Markov category C with countable Kolmogorov products such that Theorem 5.3 specializes to the standard Kolmogorov zero–one law, in its general form applicable to all measurable spaces?*

Corollary 6.3 (Hewitt-savage zero–one law)**.** *Suppose Ω is a standard Borel space with probability measure P, and let $(f_i : \Omega \to X)_{i \in \mathbb{N}}$ be a sequence of independent, identically distributed random variables taking values in a standard Borel space X. Suppose $S \subseteq \prod_{i \in \mathbb{N}} X = X_{\mathbb{N}}$ is measurable in the product σ-algebra, and suppose moreover that, for each finite permutation $\sigma : \mathbb{N} \to \mathbb{N}$, we have $\hat{\sigma}(S) = S$ (in other words, S is invariant under finite permutations of the variables). Then $P(f_{\mathbb{N}}^{-1}(S)) \in \{0,1\}$, where $f_{\mathbb{N}}$ again denotes the induced map $\Omega \to X_{\mathbb{N}}$.*

Proof. The proof is similar to the previous proof. This time we want to apply Theorem 5.4 with $\mathsf{C} = \mathsf{BorelStoch}$, $J = \mathbb{N}$ and $A = I$. Again, we have $P : I \to \Omega$. As before, we let $s : X_{\mathbb{N}} \to \{0,1\}$ be the indicator function of S, and we let $p = f_{\mathbb{N}}P$. The i.i.d. assumption on the f_i implies that $\hat{\sigma}p = p$, and the assumption on S implies $s\hat{\sigma} = s$ for finite permutations σ. Hence we can apply Theorem 5.4. This again shows that the map $sp : I \to \{0,1\}$ is deterministic, which, just as before, means $P(f_{\mathbb{N}}^{-1}(S)) \in \{0,1\}$. □

By interpreting these results in different Markov categories, we can obtain other results. In particular, the following example is of a similar flavour as $\mathsf{SetMulti}$, but improves on its lack of infinite tensor products (Example 4.4).

With Top the category of topological spaces and continuous maps, we consider the *lower Vietoris monad* or *hyperspace monad H* on Top, introduced by Schalk for T_0 spaces [16] and investigated generally in [7, Section 2]. Let $\mathsf{Kl}(H)$ be its Kleisli category. Using the symmetric monoidal structure of H guaranteed by [7, Corollary 2.53], the general construction of Markov categories from symmetric monoidal monads of [6, Proposition 3.1] turns $\mathsf{Kl}(H)$ into a Markov category. Concretely, this category has the following structure:

1. The objects are topological spaces.

2. A morphism $f : X \to Y$ is a continuous function $X \to HY$, where HY denotes the set of closed subsets of the space Y, equipped with the topology generated by subbasic opens of the form
$$\mathrm{Hit}(U) := \{C \subseteq Y \text{ closed} \mid C \cap U \neq \emptyset\},$$

where $U \subseteq Y$ ranges over all opens.

3. Composition is defined as in SetMulti via (7), but taking the closure of the union in addition.

4. The monoidal product is the Cartesian product of topological spaces, and on morphisms

$$(f \otimes g)(a,b) := f(a) \times g(b),$$

exactly as in SetMulti.

5. The comultiplication is defined by $\mathsf{copy}_X(x) = \mathrm{cl}(\{(x,x)\})$

Proposition 6.4. $\mathsf{Kl}(H)$ *has Kolmogorov products of any cardinality, given by the usual infinite product of topological spaces with the product topology.*

Proof. First we observe that $H(\prod_{i \in J} X_i)$ is the cofiltered limit of the spaces $H(\prod_{i \in F} X_i)$ for $F \subseteq J$ finite. To see this, note that the marginalizations induce a continuous map

$$\alpha \; : \; H\left(\prod_{i \in J} X_i\right) \longrightarrow \varprojlim_{F \subseteq J} H\left(\prod_{i \in J} X_i\right)$$

We prove that this is a homeomorphism, showing first that it is an injection. Thus we start with distinct closed subsets $C, D \subseteq \prod X_i$. We assume without loss of generality that $C \nsubseteq D$, and choose $x \in C \setminus D$. Then let $U \ni x$ be an open neighborhood of x not intersecting D. By making U smaller if necessary, we can assume without loss of generality that U is a basic open in the product topology, meaning that there is a finite F so that

$$y = (y_i) \in U \quad \Longleftrightarrow \quad y_i \in \pi_i(U) \;\; \forall i \in F. \tag{11}$$

Then $\pi_F(x) \in \pi_F(U)$, which is disjoint from $\pi_F(D)$. Hence $\alpha(C) \neq \alpha(D)$.

We show that α is also surjective. Given a family $(C_F) \in \lim_F H(\prod_{i \in F} X_i)$, simply set

$$C := \left\{ x \in \prod_{j \in J} X_j \;\middle|\; \pi_F(x) \in C_F \;\; \forall F \right\}$$

This is an intersection of closed subsets, hence closed, and the finite marginalizations clearly send it to the given family (C_F).

Finally, we must verify that α is an open map. It suffices to show that it carries a subbasic open set $\mathrm{Hit}(U) = \{C \subseteq \prod X_i \text{ closed} \mid C \cap U \neq \emptyset\}$ to an open subset of the limit. For a given point $C \in \mathrm{Hit}(U)$, it is enough to find an open neighborhood contained in $\mathrm{Hit}(U)$. But then again by choosing smaller U if necessary, so that $C \cap U \neq \emptyset$ still holds, we can assume (11) without loss of generality. Then

$$\alpha(A_U) = \pi_F^{-1}\left(\left\{ D \subseteq \prod_{j \in F} X_F \text{ closed} \;\middle|\; D \cap \pi_F(U) \neq \emptyset \right\}\right),$$

which is indeed open in the limit.

It follows that $\prod_i X_i$ is an infinite tensor product in $\mathsf{Kl}(H)$; the preservation condition is easily verified since adding another factor to this product again gives the infinite product. The Kolmogorov product condition is similarly immediate to verify. $\qquad\square$

We can therefore instantiate our Theorem 5.3 in $\mathsf{Kl}(H)$, which gives the following result as a special case.

Corollary 6.5. *Let $(X_i)_{i \in J}$ be a family of topological spaces, Y a Hausdorff space, and let $f : \prod_i X_i \to Y$ be a continuous function which is independent of any finite subset of the input. Then f is constant.*

Proof. These results follow from Theorem 5.3 applied to $\mathsf{Kl}(H)$, using the morphism $p : I \to X_J$ corresponding to the trivial closed subset $X_J \subseteq X_J$. The conclusion is that the closure of its image, call it A, satisfies

$$\mathrm{cl}(\{(a,a) \mid a \in A\}) = A \times A$$

as subsets of $Y \times Y$. But since the diagonal is closed in Y, this clearly implies that A must be a singleton. $\qquad\square$

However, we cannot instantiate the Hewitt–Savage zero–one law in $\mathsf{Kl}(H)$, since this Markov category is in fact not causal, as the following example shows.

Example 6.6. Let $A = \{*\}$ and $X = Y = Z = [0,1]$ with the usual topology. Let $f : A \to HX$ send $*$ to X, and define

$$g(x) := \begin{cases} [0, 1/2] & x < \frac{1}{3} \\ \{1/2\} & \frac{1}{3} \leq x \leq \frac{2}{3} \\ [1/2, 1] & x > \frac{1}{3} \end{cases}$$

$$h_1(y) := [0,1], \qquad h_2(y) = \begin{cases} [0,1] & y \neq \frac{1}{2} \\ \{0\} & y = \frac{1}{2} \end{cases}$$

Using the definition of the hyperspace topology, it is straightforward to see that these maps are continuous. To check the hypothesis of the statement of causality, we must verify that the resulting two maps $A \to Y \times Z$ agree; these are just closed subsets of $[0,1]^2$. From the definition of the composition, we see that any point (y, z) with $y \in g(x)$ for some x and $z \in h_1(y)$ must lie in the set corresponding to h_1, and similarly for h_2. Hence the set corresponding to h_1 contains all points, while the set corresponding to h_2 contains as a subset all points (y, z) with $y \neq 1/2$, as well as $(1/2, 0)$. But the closure of $\{(y, z) \mid y \neq 1/2\}$ in $[0,1]^2$ is the whole space, so that both sets are indeed equal to the whole space.

Hence the hypothesis is satisfied. But one can verify that the point $(x, y, z) = (1/2, 1/2, 1)$ is in the subset of $X \times Y \times Z$ corresponding to the h_1 side of (4), but not so for the h_2 side.

We end with an open problem.

Problem 6.7. *Find an interesting causal Markov category which has all Kolmogorov products.*

Strangely enough, the only examples that we know of so far are of two kinds: first, cartesian monoidal categories (meaning that $\mathsf{C} = \mathsf{C}_{\mathrm{det}}$) with all products; and second, the finitary complete monoids of Example 3.8 which are monoidal posets. For the former, the universal property of the monoidal structure implies that the causality axiom holds. For the latter, it holds trivially since any two parallel morphisms are equal.

References

[1] Erik Sparre Andersen and Børge Jessen. On the introduction of measures in infinite product sets. *Kongelige Danske Videnskabernes Selskab, Matematiske-Fysiske meddelelser*, 25(4):8, 1948. ↓9

[2] Bruce E. Blackadar. Infinite tensor products of C^*-algebras. *Pacific Journal of Mathematics*, 72(2):313–334, 1977. ↓8

[3] Kenta Cho and Bart Jacobs. Disintegration and Bayesian inversion via string diagrams. *Mathematical Structures in Computer Science*, 29:938–971, 2019. https://doi.org/10.1017/S0960129518000488. ↓1, 2, 4, 6

[4] Pierre Deligne, James S. Milne, Arthur Ogus, and Kuang-yen Shih. *Hodge cycles, motives, and Shimura varieties*, volume 900 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin-New York, 1982. https://doi.org/10.1007/978-3-540-38955-2. ↓7

[5] D. H. Fremlin. *Measure theory. Vol. 4.* Torres Fremlin, Colchester, 2006. Topological measure spaces. Part I, II, Corrected second printing of the 2003 original. https://www1.essex.ac.uk/maths/people/fremlin/mt.htm www1.essex.ac.uk/maths/people/fremlin/mt.htm. ↓9

[6] Tobias Fritz. A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics. https://doi.org/10.1016/j.aim.2020.107239. ↓1, 2, 4, 5, 6, 7, 11, 12, 14, 17

[7] Tobias Fritz, Paolo Perrone, and Sharwin Rezagholi. Probability, valuations, hyperspace: Three monads on top and the support as a morphism. *ArXiv*, abs/1910.03752, 2019. ↓17

[8] Malte Gerhold, Stephanie Lachs, and Michael Schürmann. Categorial Lévy processes. https://arxiv.org/abs/1612.05139 arXiv:1612.05139. ↓3

[9] Michèle Giry. A categorical approach to probability theory. In *Categorical aspects of topology and analysis (Ottawa, Ont., 1980)*, volume 915 of *Lecture Notes in Mathematics*, pages 68–85. Springer, 1982. https://doi.org/10.1007/BFb0092872. ↓4

[10] Peter V. Golubtsov. Axiomatic description of categories of information converters. *Problemy Peredachi Informatsii*, 35(3):80–98, 1999. In Russian. English translation in *Problems of Information Transmission* 35(3):259–274, 1999. ↓1, 4

[11] Edwin Hewitt and Leonard J. Savage. Symmetric measures on cartesian products. *Transactions of the American Mathematical Society*, 80(2):470–470, February 1955. https://doi.org/10.1090/s0002-9947-1955-0076206-8. ↓3

[12] Georg Karner. Continuous monoids and semirings. *Theoretical Computer Science*, 318(3):355–372, 2004. https://doi.org/10.1016/j.tcs.2004.01.020. ↓10

[13] Achim Klenke. *Probability theory, a comprehensive course.* Universitext. Springer, London, second edition, 2014. https://doi.org/10.1007/978-1-4471-5361-0. ↓2, 9

[14] A. N. Kolmogorov. *Grundbegriffe der Wahrscheinlichkeitsrechnung.* Springer, Berlin, 1933. ↓2

[15] D. Ramachandran. Perfect mixtures of perfect measures. *Annals of Probability*, 7(3):444–452, 1979. https://doi.org/10.1214/aop/1176995045. ↓9

[16] Andrea Schalk. *Algebras for generalized power constructions.* PhD thesis, University of Darmstadt, 1993. Available at www.cs.man.ac.uk/~schalk/publ/diss.ps.gz www.cs.man.ac.uk/ schalk/publ/diss.ps.gz. ↓17