sketch of pieces of

A synthetic ~~introduction to~~ probability and statistics

Tobias Fritz

November 2019

I will give a very short introduction to the **mathematical foundations of probability and statistics**.

This will be based on measure theory, as usual.

I will give a very short introduction to the **mathematical foundations of probability and statistics**.

This will be based on ~~measure theory, as usual~~ category theory.

⇒ An unconventional introduction!

References

- Tobias Fritz,
  A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics,
  arXiv:1908.07021.

- Tobias Fritz and Eigil Fjeldgren Rischel,
  The zero–one laws of Kolmogorov and Hewitt–Savage in categorical probability, in preparation.

- Peter V. Golubtsov,
  Axiomatic description of categories of information converters,
  Problemy Peredachi Informatsii, 35(3):80–98 (1999).
  (And other similar papers by Golubtsov.)

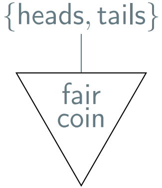
- Kenta Cho and Bart Jacobs,
  Disintegration and Bayesian inversion via string diagrams,
  Math. Struct. Comp. Sci. 29:938–971 (2019). arXiv:1709.00322.

Probability theory is about *processes*

which can be composed in series and in parallel.

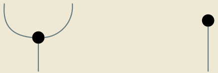
Intuition: process = function with random output.

A *probability distribution* is a process with no input, like this:

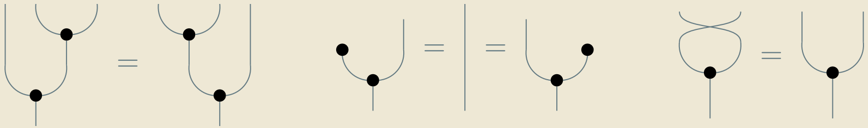


Definition

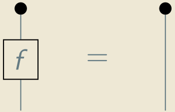
A **Markov category** C is a symmetric monoidal category with **copying** and **deleting** operations on every object,



giving commutative comonoid structures

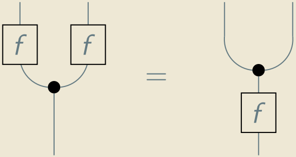


which interact well with the monoidal structure, and such that



Definition

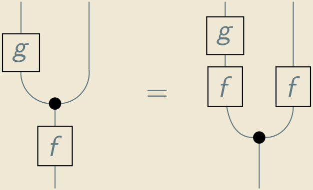
A morphism $f : X \to Y$ is **deterministic** if it commutes with copying,



- **Intuition:** Applying f to copies of input = copying the output of f.
- The deterministic morphisms form a cartesian monoidal subcategory.

Definition

C is **positive** if whenever gf is deterministic for composable f and g, then also



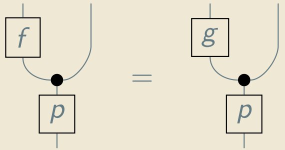
- **Intuition:** If a deterministic process has a random intermediate result, then that result can be computed independently from the process.

- Not every Markov category is positive.

Definition

Let $p : A \to X$ and $f, g : X \to Y$.

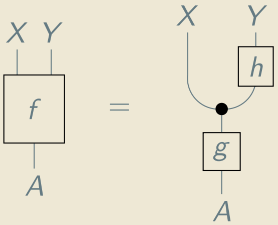
f and g are **equal p-almost surely**, $f =_{p\text{-a.s.}} g$, if



- **Intuition:** f and g behave the same on all inputs produced by p.
- Other concepts (besides equality) also relativize with respect to p-almost surely.
- In particular, C is **strictly positive** if it satisfies a relativized positivity axiom.

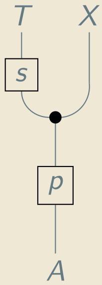
Definition

$f : A \to X \otimes Y$ **displays the conditional independence** $A \perp Y \mid X$ if there are g and h such that



Definition

- A **statistical model** on X is a morphism $p : A \to X$.
- A **statistic** for p is a deterministic morphism $s : X \to T$.
- The statistic is **sufficient** if



displays $A \perp X \mid T$.

There is a version of the **Fisher–Neyman factorization theorem**.

Theorem

Suppose that C is strictly positive.

A statistic $s : X \to T$ is sufficient for $p : A \to X$ if and only if there is $\alpha : T \to X$ with $\alpha s p = p$.

There are versions of other classical theorems of statistics.

Basu's theorem

A complete sufficient statistic for p is independent of any ancillary statistic.

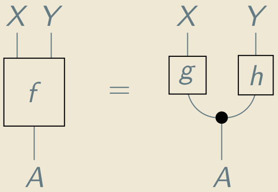
Bahadur's theorem

If a minimal sufficient statistic exists, then a complete sufficient statistic is minimal sufficient.

Explaining these would first require stating the relevant additional definitions, for which I don't have time.

Definition

$f : A \to X \otimes Y$ **displays the conditional independence** $X \perp Y \,||\, A$ if there are g and h such that



- **Intuition:** The outputs X and Y can be produced independently.

- Note the difference from the earlier definition of conditional independence!

Definition

Let $(X_i)_{i \in I}$ be a family of objects. The **infinite tensor product**

$$X_I := \bigotimes_{i \in I} X_i$$

is the cofiltered limit of the finite tensor products $X_F := \bigotimes_{i \in F} X_i$, if this limit exists and is preserved by every $- \otimes Y$.

Definition

An infinite tensor product X_I is a **Kolmogorov product** if the limit projections $\pi_F : X_I \to X_F$ are deterministic.

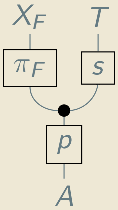
- This additional condition fixes the comonoid structure on X_I.

Theorem (Kolmogorov zero–one law)

Let X_I be a Kolmogorov product of a family $(X_i)_{i \in I}$.

If

- $p : A \to X_I$ makes the X_i independent and identically distributed, and
- $s : X_I \to T$ is such that

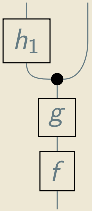


  displays $X_F \perp T \,||\, A$ for every finite $F \subseteq I$,

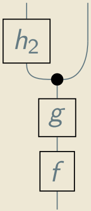
then ps is deterministic.

Definition

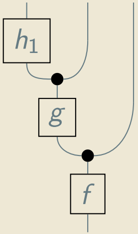
C is **causal** if



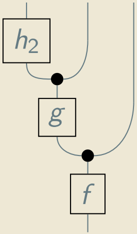
=



implies



=



- **Intuition:** The choice between h_1 and h_2 in the "future" of g does not influence the "past" of g.

- Not every Markov category is causal.

Theorem (Hewitt–Savage zero–one law)

Suppose that C is causal, I infinite and $X_I := \bigotimes_{i \in I} X$ a Kolmogorov product of the same X with itself.

If

- ▷ $p : A \to X_I$ makes the X_i independent and identically distributed, and
- ▷ $s : X_I \to T$ is deterministic and invariant under finite permutations,

then ps is deterministic.

Example

If $\prod_{i \in I} X$ is an infinite product of the same topological space, Y a Hausdorff space and $f : \prod_i X \to Y$ continuous and invariant under finite permutations, then f is constant.

The End

Please join us for

“Categorical Probability and Statistics”

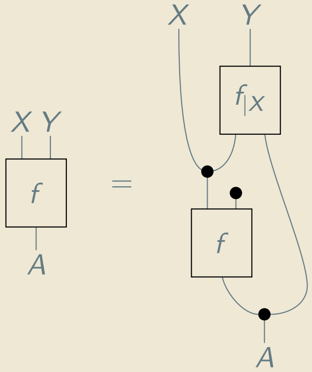
at 75th Anniversay Summer Meeting, Canadian Mathematical Society,

5–8 June 2020, University of Ottawa.

Organizers: Rory Lucyshyn-Wright and myself.

Definition

C **has conditionals** if for $f : A \to X \otimes Y$ there is $f_{|X} : X \otimes A \to Y$ with



- If C has conditionals, then it is both strictly positive and causal.
- The positivity and causality axioms (partly?) eliminate the relevance of conditionals!