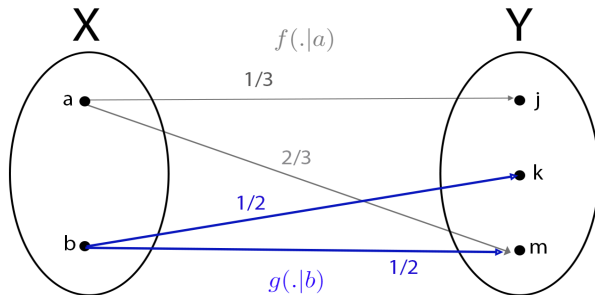


Markov Category : A 10 Minute Introduction

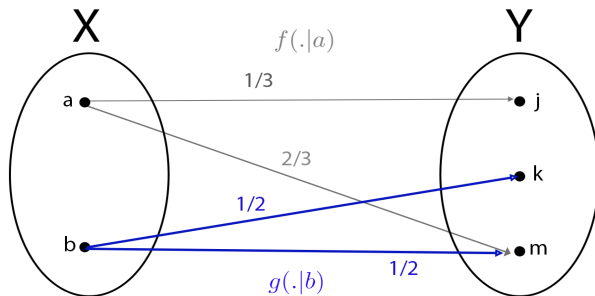
Mahdi Zamani

Nov 2023

Random Morphisms



Random Morphisms

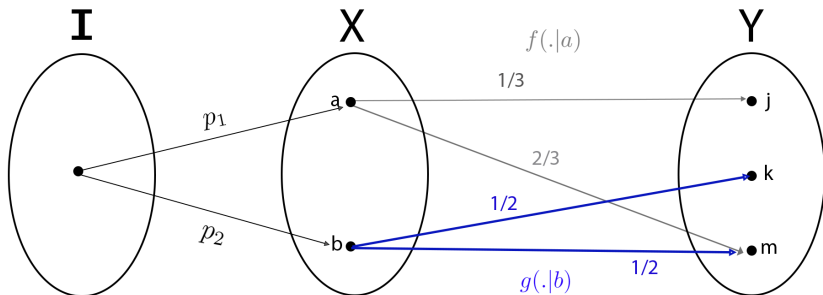


Example

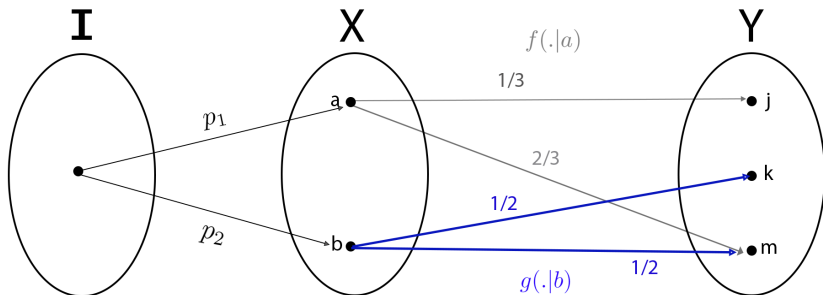
FinStoch Category, with measurable sets as objects, and random morphisms (Markov kernels for Stoch).

$$\begin{array}{c|cc} \swarrow & a & b \\ j & \left(\begin{array}{cc} 1/3 & 0 \end{array} \right) & f(j|a) = 1/3 \\ k & \begin{array}{cc} 0 & 1/2 \end{array} \\ m & \begin{array}{cc} 2/3 & 1/2 \end{array} \end{array}$$

Probability Distribution



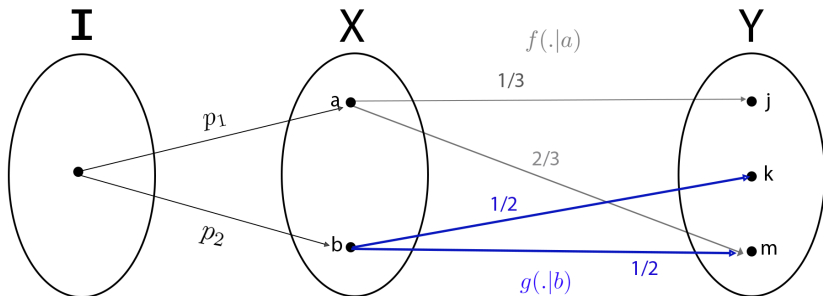
Probability Distribution



Notation

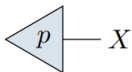
$$I \xrightarrow{P} X$$

Probability Distribution



Notation

$$I \xrightarrow{P} X$$

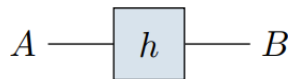
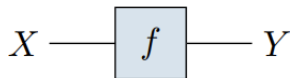


Products

$$X \times A \xrightarrow{\sim} X \otimes A$$

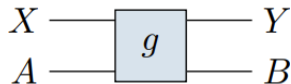
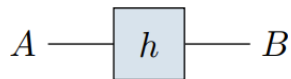
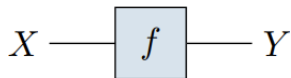
Products

$$X \times A \xrightarrow{\sim} X \otimes A$$



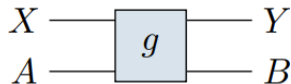
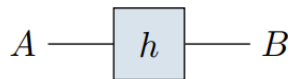
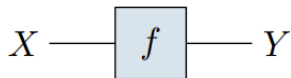
Products

$$X \times A \xrightarrow{\sim} X \otimes A$$



Products

$$X \otimes A \rightsquigarrow X \otimes A$$



$$g : X \otimes A \rightarrow Y \otimes B$$

Markov Category [2]

Markov Category [2]

Definition

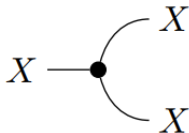
A Markov category C is a symmetric monoidal category supplied with copying and deleting morphisms on every object,

Markov Category [2]

Definition

A Markov category \mathcal{C} is a symmetric monoidal category supplied with copying and deleting morphisms on every object,

$\text{copy} : X \rightarrow X \otimes X$

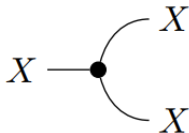


Markov Category [2]

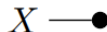
Definition

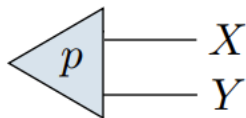
A Markov category \mathcal{C} is a symmetric monoidal category supplied with copying and deleting morphisms on every object,

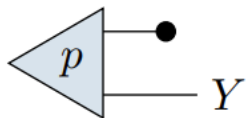
$\text{copy} : X \rightarrow X \otimes X$



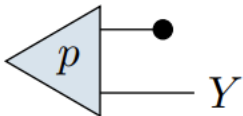
$\text{del} : X \rightarrow I$







$$\mathbb{P}_Y(y) = \sum_{x \in X} \mathbb{P}_{X,Y}(x, y)$$



Commutative comonoid axioms

Commutative comonoid axioms

$$X \text{ --- } \bullet \begin{array}{c} \curvearrowright \\ \bullet \end{array} = X \text{ --- } X = X \text{ --- } \bullet \begin{array}{c} \curvearrowleft \\ \bullet \end{array} X$$

The diagram illustrates the commutative comonoid axiom. It shows three expressions separated by equals signs. The first expression is a horizontal line labeled X ending in a black dot, from which a curved line goes up and then down to another black dot, with the label X at the bottom right. The second expression is a simple horizontal line labeled X at both ends. The third expression is a horizontal line labeled X starting from a black dot, from which a curved line goes up and then down to another black dot, with the label X at the top right.

Commutative comonoid axioms

$$\begin{array}{c} \bullet \\ \curvearrowright \\ X - \bullet \\ \curvearrowleft \\ X \end{array} = X \text{ --- } X = \begin{array}{c} X \\ \curvearrowleft \\ X - \bullet \\ \curvearrowright \\ \bullet \end{array}$$

$$\begin{array}{c} X \\ \curvearrowleft \\ \bullet \\ \curvearrowright \\ X \\ \curvearrowleft \\ X - \bullet \\ \curvearrowright \\ X \end{array} = \begin{array}{c} X \\ \curvearrowleft \\ X - \bullet \\ \curvearrowright \\ \bullet \\ \curvearrowleft \\ X \\ \curvearrowright \\ X \end{array}$$

Commutative comonoid axioms

$$X \text{ --- } \bullet \begin{array}{l} \nearrow \\ \searrow \end{array} \begin{array}{l} \bullet \\ X \end{array} = X \text{ --- } X = X \text{ --- } \bullet \begin{array}{l} \searrow \\ \nearrow \end{array} \begin{array}{l} X \\ \bullet \end{array}$$

$$X \text{ --- } \bullet \begin{array}{l} \nearrow \\ \searrow \end{array} \begin{array}{l} \bullet \\ \nearrow \\ \searrow \end{array} \begin{array}{l} X \\ X \\ X \end{array} = X \text{ --- } \bullet \begin{array}{l} \nearrow \\ \searrow \end{array} \begin{array}{l} X \\ \bullet \\ \nearrow \\ \searrow \end{array} \begin{array}{l} X \\ X \\ X \end{array}$$

$$X \text{ --- } \bullet \begin{array}{l} \nearrow \\ \searrow \end{array} \begin{array}{l} \bullet \\ \nearrow \\ \searrow \end{array} \begin{array}{l} X \\ X \end{array} = X \text{ --- } \bullet \begin{array}{l} \nearrow \\ \searrow \end{array} \begin{array}{l} X \\ X \end{array}$$

Naturality axiom

$$X \text{ --- } \boxed{f} \text{ --- }^Y \bullet = X \text{ --- } \bullet$$

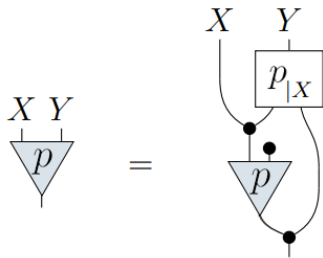
Conditional Distribution

Conditional Distribution

$p_{|X} : X \rightarrow Y$ is a conditional of p wrt to X

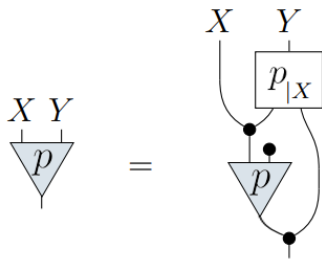
Conditional Distribution

$p_{|X} : X \rightarrow Y$ is a conditional of p wrt to X



Conditional Distribution

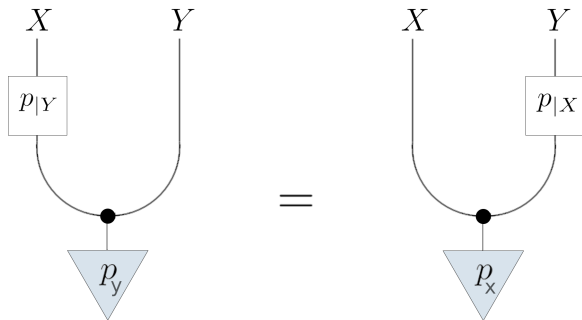
$p_{|X} : X \rightarrow Y$ is a conditional of p wrt to X



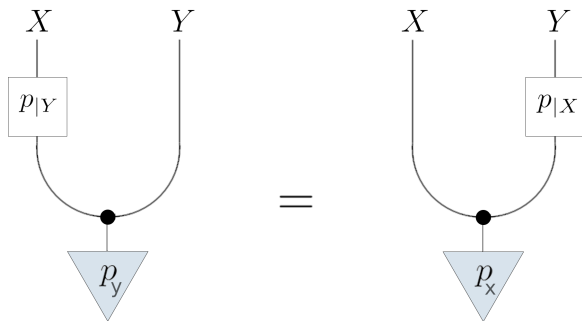
$$\mathbb{P}_{X,Y}(x,y) = \mathbb{P}_X(x)\mathbb{P}_{Y|X}(y|x)$$

Bayes

Bayes



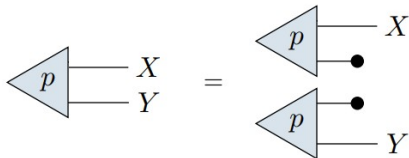
Bayes



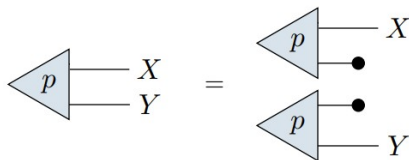
$$\mathbb{P}_Y(y)\mathbb{P}_{X|Y}(x|y) = \mathbb{P}_X(x)\mathbb{P}_{Y|X}(y|x)$$

(Conditional) Independence

(Conditional) Independence

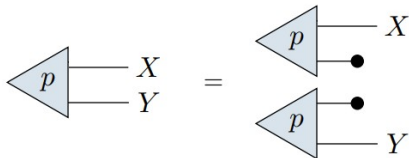


(Conditional) Independence

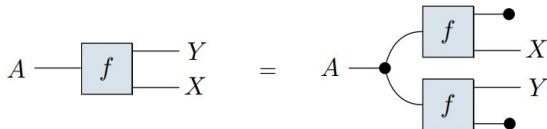


$$\mathbb{P}_{X,Y}(x,y) = \mathbb{P}_X(x)\mathbb{P}_Y(y)$$

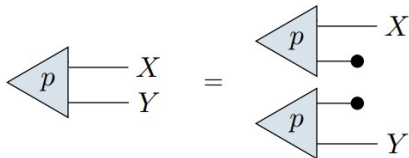
(Conditional) Independence



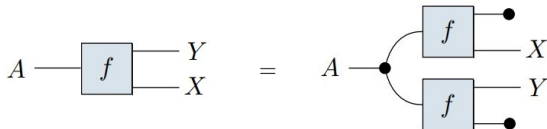
$$\mathbb{P}_{X,Y}(x,y) = \mathbb{P}_X(x)\mathbb{P}_Y(y)$$



(Conditional) Independence



$$\mathbb{P}_{X,Y}(x,y) = \mathbb{P}_X(x)\mathbb{P}_Y(y)$$



$$\mathbb{P}_{X,Y|A}(x,y|a) = \mathbb{P}_X(x|a)\mathbb{P}_Y(y|a)$$

Applications : Sufficient statistics [2]

Definition

Applications : Sufficient statistics [2]

Definition

A "(parametric) statistical model" is a map $p : \Theta \rightarrow X$

Applications : Sufficient statistics [2]

Definition

A "(parametric) statistical model" is a map $p : \Theta \rightarrow X$

A "statistic" is a deterministic map $s : X \rightarrow V$

Applications : Sufficient statistics [2]

Definition

A "(parametric) statistical model" is a map $p : \Theta \rightarrow X$

A "statistic" is a deterministic map $s : X \rightarrow V$

A statistic is *sufficient* if $X \perp \Theta \mid V$

Applications : Sufficient statistics [2]

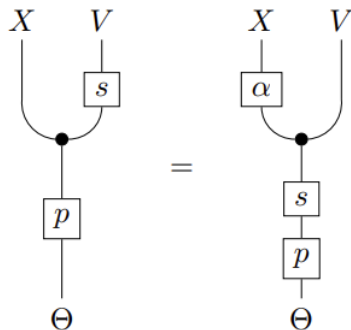
Definition

A "(parametric) statistical model" is a map $p : \Theta \rightarrow X$

A "statistic" is a deterministic map $s : X \rightarrow V$

A statistic is *sufficient* if $X \perp \Theta \mid V$

Which means that we have $\alpha : V \rightarrow X$, s.t



Where to go

Where to go

- ▶ Categorical Bayesian network [1]

Where to go

- ▶ Categorical Bayesian network [1]
- ▶ De Finetti's Theorem [3]

Where to go

- ▶ Categorical Bayesian network [1]
- ▶ De Finetti's Theorem [3]
- ▶ Kolmogorov 0/1 law [4]

Where to go

- ▶ Categorical Bayesian network [1]
- ▶ De Finetti's Theorem [3]
- ▶ Kolmogorov 0/1 law [4]
- ▶ Markov categories and entropy (generalized divergence) [5]

Where to go

- ▶ Categorical Bayesian network [1]
- ▶ De Finetti's Theorem [3]
- ▶ Kolmogorov 0/1 law [4]
- ▶ Markov categories and entropy (generalized divergence) [5]
- ▶ Generalized Optimization [6]

Where to go

- ▶ Categorical Bayesian network [1]
- ▶ De Finetti's Theorem [3]
- ▶ Kolmogorov 0/1 law [4]
- ▶ Markov categories and entropy (generalized divergence) [5]
- ▶ Generalized Optimization [6]
- ▶ and much more! [7]

References

- [1] Brendan Fong. *Causal Theories: A Categorical Perspective on Bayesian Networks*. 2013. eprint: [arXiv:1301.6201](#).
- [2] Tobias Fritz. “A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics”. In: (2019). doi: [10.1016/j.aim.2020.107239](#). eprint: [arXiv:1908.07021](#).
- [3] Tobias Fritz, Tomáš Gonda, and Paolo Perrone. “De Finetti’s Theorem in Categorical Probability”. In: (2021). doi: [10.31390/josa.2.4.06](#). eprint: [arXiv:2105.02639](#).
- [4] Tobias Fritz and Eigil Fjeldgren Rischel. “Infinite products and zero-one laws in categorical probability”. In: (2019). doi: [10.32408/compositionality-2-3](#). eprint: [arXiv:1912.02769](#).
- [5] Paolo Perrone. *Markov Categories and Entropy*. 2022. eprint: [arXiv:2212.11719](#).
- [6] Dan Shiebler. *Generalized Optimization: A First Step Towards Category Theoretic Learning Theory*. 2021. eprint: [arXiv:2109.10262](#).
- [7] Dan Shiebler, Bruno Gavranović, and Paul Wilson. *Category Theory in Machine Learning*. 2021. eprint: [arXiv:2106.07032](#).

Thank You

Questions