On the gain of information from independent samples

Student
Hauri Sébastien

Professor
Pr. Telatar Emre
Assistant
Quinton Pierre

A thesis presented for the
Bachelor semester project

Ecole Polytechnique Fédérale de Lausanne

June 11, 2021

1 Introduction

The following work describes an operator useful to measure the gain of information obtained from independent samples. First we need to define this operator and that is what we do in part 2. We also found some interesting properties that we claimed and proved. Then we discuss and try to find an inverse operator. In the end, we try to describe a decomposition on random variables in order to apply this operator not only on independent samples but any set of samples. We finish our study by giving some leads on what can be done to continue some research about this operator.

2 Definition of the $\dot{\otimes}$ operator

Let X be a random variable taking values in alphabet $\mathcal{X} = \{x_1, \dots, x_n\}$. We denote by P_X the n dimensional vector $[\mathbb{P}[X = x_1], \dots, \mathbb{P}[X = x_n]]^T$ that belongs to the simplex $S = \{\mathbf{p} \in [0,1]^n : \sum_{i=1}^n p_i = 1, \forall i, p_i \geq 0\}$.
Fix $P_X = \mathbf{p} \in S$. Let U be a random variable taking value in $\{u_1, \dots, u_{m_u}\}$ with probabilities $\mathbb{P}[U = u_i] = \alpha_i$, $\sum_{i=1}^{m_u} \alpha_i = 1$. Let V be defined in the same way as U taking values in $\{v_1, \dots, v_{m_v}\}$ and probabilities β_j, $1 \leq j \leq m_v$. Let us define $P_{X|U} \sim F$ with support $\{\mathbf{f}_1, \dots, \mathbf{f}_{m_u}\}$ where $\mathbf{f}_i \in S$ and $\sum_{i=1}^{m_u} \mathbf{f}_i \alpha_i = \mathbf{p}$. In the same way, let $P_{X|V} \sim G$ with support $\{\mathbf{g}_1, \dots, \mathbf{g}_{m_v}\}$ where $\mathbf{g}_j \in S$ and $\sum_{j=1}^{m_v} \mathbf{g}_j \beta_j = \mathbf{p}$. Here we assumed that $|S_U| = |S_F|$ (and same for V and G) for simplicity but the map from symbols of U to $P_{X|U}$ (respectively V to $P_{X|V}$) is not, in general, injective. A bit of intuition about $P_{X|U}$: it is a random vector and conditioned on $U = u_i$, then $P_{X|U} = P_{X|U=u_i} = \mathbf{f}_i$. For this reason, $P_{X|U} = \mathbb{E}[\delta_X | U]$ with $\delta_X = [0\ 0\ \dots\ 1\ \dots\ 0]^T$ with the 1 at position x.
The main goal is now to define an operator $\dot{\otimes}$ s.t.

$$P_{X|U,V} \sim F \dot{\otimes} G$$

In order to do that, we describe the support of $F \dot{\otimes} G$ and its probability distribution. We make here an important assumption : U and V are independent knowing X.

Support of $F \dot{\otimes} G$:
We start by computing

$$\mathbb{P}[X = x_i | U = u_j, V = v_k] = \frac{\mathbb{P}[X = x_i, U = u_j, V = v_k]}{\sum_{l=1}^n \mathbb{P}[X = x_l, U = u_j, V = v_k]}$$

Moreover,

$$\begin{aligned}
&\mathbb{P}[X = x_i, U = u_j, V = v_k] \\
&= \mathbb{P}[X = x_i | U = u_j] \cdot \mathbb{P}[U = u_j] \cdot \mathbb{P}[V = v_k | X = x_i, U = u_j] \\
&= \mathbb{P}[X = x_i | U = u_j] \cdot \mathbb{P}[U = u_j] \cdot \mathbb{P}[V = v_k | X = x_i] \\
&= \mathbb{P}[U = u_j] \cdot \mathbb{P}[V = v_k] \cdot \frac{\mathbb{P}[X = x_i | U = u_j] \cdot \mathbb{P}[X = x_i | V = v_k]}{\mathbb{P}[X = x_i]} \\
&= \mathbb{P}[U = u_j] \cdot \mathbb{P}[V = v_k] \cdot \frac{\mathbb{P}[X = x_i | U = u_j] \cdot \mathbb{P}[X = x_i | V = v_k]}{\mathbb{P}[X = x_i]} \\
&= \alpha_j \beta_k \cdot \frac{f_{j,i} \cdot g_{k,i}}{p_i}
\end{aligned}$$

Thus,

$$\mathbb{P}[X = x_i | U = u_j, V = v_k] = \frac{1}{\sum_{l=1}^{n} \frac{f_{j,l} g_{k,l}}{p_l}} \cdot \frac{f_{j,i} g_{k,i}}{p_i}$$

This being computed, we can now state that:

$$P_{X|U=u_j, V=v_k} = \frac{1}{\sum_{l=1}^{n} \frac{f_{j,l} g_{k,l}}{p_l}} \cdot \begin{bmatrix} \frac{f_{j,1} g_{k,1}}{p_1} \\ \dots \\ \frac{f_{j,n} g_{k,n}}{p_n} \end{bmatrix}$$

With this being done, we have computed the support of $F \dot{\otimes} G$ which namely is :

$$S_{F \dot{\otimes} G} = \left\{ \frac{1}{\sum_{l=1}^{n} \frac{f_{j,l} g_{k,l}}{p_l}} \cdot \begin{bmatrix} \frac{f_{j,1} g_{k,1}}{p_1} \\ \dots \\ \frac{f_{j,n} g_{k,n}}{p_n} \end{bmatrix} \right\}_{1 \le j \le m_u, 1 \le k \le m_v}$$

Probabilities associated to the support :

It remains now to compute the probabilities associated to each of those vectors. In order to do so, we compute :

$$\begin{aligned}
&\mathbb{P}[U = u_j, V = v_k] \\
&= \mathbb{P}[V = v_k | U = u_j] \cdot \mathbb{P}[U = u_j] \\
&= \sum_{i=1}^{n} \mathbb{P}[V = v_k | X = x_i] \mathbb{P}[X = x_i | U = u_j] \cdot \mathbb{P}[U = u_j] \\
&= \sum_{i=1}^{n} \frac{\mathbb{P}[X = x_i | U = u_j] \mathbb{P}[X = x_i | V = v_k]}{\mathbb{P}[X = x_i]} \cdot \mathbb{P}[U = u_j] \mathbb{P}[V = v_k] \\
&= \alpha_j \beta_k \sum_{i=1}^{n} \frac{f_{j,i} g_{k,i}}{p_i}
\end{aligned}$$

Note : it is in fact possible that during the computation of $F \dot{\otimes} G$ several points collapse to the same point. In this case, we just take the sum of the probabilities associated to that point.

Conclusion of what we have done so far :

Given $P_X = \mathbf{p}, P_{X|U} \sim F$ and $P_{X|V} \sim G$ both averaging to $\mathbf{p}$ such that U and V are independent knowing X, we defined an operator $\dot{\otimes}$:

- $P_{X|U,V} \sim F \dot{\otimes} G$
- $S_{F \dot{\otimes} G} = \left\{ \frac{1}{\sum_{l=1}^{n} \frac{f_{j,l} g_{k,l}}{p_l}} \cdot \begin{bmatrix} \frac{f_{j,1} g_{k,1}}{p_1} \\ \dots \\ \frac{f_{j,n} g_{k,n}}{p_n} \end{bmatrix} \right\}_{1 \leq j \leq m_u, 1 \leq k \leq m_v}$
- $\forall \mathbf{y}_{jk} \in S_{F \dot{\otimes} G}, \mathbb{P}[P_{X|U,V} = \mathbf{y}_{jk}] = \alpha_j \beta_k \sum_{i=1}^{n} \frac{f_{j,i} g_{k,i}}{p_i}$

which will give us the behaviour $P_{X|U,V}$ in terms of $P_{X|U}$ and $P_{X|V}$ when U and V are independent knowing X.
In addition of that, we can state the following claims :

Claim 1 $F \dot{\otimes} G$ is a probability distribution

$$\begin{aligned}
\sum_{j=1}^{m_u} \sum_{k=1}^{m_v} \alpha_j \beta_k \sum_{i=1}^{n} \frac{f_{j,i} g_{k,i}}{p_i} &= \sum_{i=1}^{n} \frac{1}{p_i} \sum_{j=1}^{m_u} \sum_{k=1}^{m_v} \alpha_j \beta_k f_{j,i} g_{k,i} \\
&= \sum_{i=1}^{n} \frac{1}{p_i} \sum_{j=1}^{m_u} \alpha_j f_{j,i} \sum_{k=1}^{m_v} \beta_k g_{k,i} \\
&= \sum_{i=1}^{n} \frac{1}{p_i} \cdot p_i \cdot p_i \\
&= \sum_{i=1}^{n} p_i \\
&= 1
\end{aligned}$$

Claim 2 $F \dot{\otimes} G$ averages to $\mathbf{p}$

$$\sum_{j=1}^{m_u}\sum_{k=1}^{m_v}\alpha_j\beta_k\sum_{i=1}^{n}\frac{f_{j,i}g_{k,i}}{p_i}\cdot\frac{1}{\sum_{l=1}^{n}\frac{f_{j,l}g_{k,l}}{p_l}}\cdot\begin{bmatrix}\frac{f_{j,1}g_{k,1}}{p_1}\\ \dots\\ \frac{f_{j,n}g_{k,n}}{p_n}\end{bmatrix}$$

$$=\sum_{j=1}^{m_u}\sum_{k=1}^{m_v}\alpha_j\beta_k\begin{bmatrix}\frac{f_{j,1}g_{k,1}}{p_1}\\ \dots\\ \frac{f_{j,n}g_{k,n}}{p_n}\end{bmatrix}$$

$$=\begin{bmatrix}\sum_{j=1}^{m_u}\sum_{k=1}^{m_v}\alpha_j\beta_k\frac{f_{j,1}g_{k,1}}{p_1}\\ \dots\\ \sum_{j=1}^{m_u}\sum_{k=1}^{m_v}\alpha_j\beta_k\frac{f_{j,n}g_{k,n}}{p_n}\end{bmatrix}$$

$$=\begin{bmatrix}\frac{1}{p_1}\sum_{j=1}^{m_u}\alpha_j f_{j,1}\sum_{k=1}^{m_v}\beta_k g_{k,1}\\ \dots\\ \frac{1}{p_n}\sum_{j=1}^{m_u}\alpha_j f_{j,n}\sum_{k=1}^{m_v}\beta_k g_{k,n}\end{bmatrix}$$

$$=\begin{bmatrix}\frac{1}{p_1}\cdot p_1\cdot p_1\\ \dots\\ \frac{1}{p_n}\cdot p_n\cdot p_n\end{bmatrix}$$

$$=\begin{bmatrix}p_1\\ \dots\\ p_n\end{bmatrix}$$

$$=\mathbf{p}$$

Claim 3 $F \dot{\otimes} G = G \dot{\otimes} F$ (commutative)

Intuitively we have $P_{X|U,V} = P_{X|V,U}$. One can check that applying the definition the other way leads in fact to the same result.

Claim 4 $(F \dot{\otimes} G) \dot{\otimes} H = F \dot{\otimes} (G \dot{\otimes} H)$ (associative)

Let $P_{X|U} \sim F$, $P_{X|V} \sim G$, $P_{X|W} \sim H$ having supports $\{\mathbf{f}_1, \dots, \mathbf{f}_r\}$, $\{\mathbf{g}_1, \dots, \mathbf{g}_s\}$, $\{\mathbf{h}_1, \dots, \mathbf{h}_t\}$, with their associated probabilities $\alpha_1, \dots, \alpha_r$, $\beta_1, \dots, \beta_s$, $\gamma_1, \dots, \gamma_t$ respectively, all centered in P_X and following the assumption that U, V, W are fully independent knowing X.
By the previous definition, $F \dot{\otimes} G$ exists with support :

$$\left\{\frac{1}{\sum_{l=1}^{n}\frac{f_{i,l}g_{j,l}}{p_l}}\cdot\begin{bmatrix}\frac{f_{i,1}g_{j,1}}{p_1}\\ \dots\\ \frac{f_{i,n}g_{j,n}}{p_n}\end{bmatrix}\right\}_{1\le i\le r,1\le j\le s} = \{\mathbf{f}_{ij}\}_{1\le i\le r,1\le j\le s}$$

As $F \mathbin{\dot\otimes} G$ is also centered in P_X, we can compute $(F \mathbin{\dot\otimes} G) \mathbin{\dot\otimes} H$. We find its support :

$$S_{(F\dot\otimes G)\dot\otimes H} = \left\{ \frac{1}{\sum_{l=1}^{n} \frac{f_{ij,l}h_{k,l}}{p_l}} \cdot \begin{bmatrix} \frac{f_{ij,1}h_{k,1}}{p_1} \\ \dots \\ \frac{f_{ij,n}h_{k,n}}{p_n} \end{bmatrix} \right\}_{1\le i\le r, 1\le j\le s, 1\le k\le t}$$

Similarly, we compute the supports of $G \mathbin{\dot\otimes} H$ and $F \mathbin{\dot\otimes} (G \mathbin{\dot\otimes} H)$:

$$S_{G\dot\otimes H} = \left\{ \frac{1}{\sum_{l=1}^{n} \frac{g_{j,l}h_{k,l}}{p_l}} \cdot \begin{bmatrix} \frac{g_{j,1}h_{k,1}}{p_1} \\ \dots \\ \frac{g_{j,n}h_{k,n}}{p_n} \end{bmatrix} \right\}_{1\le j\le s, 1\le k\le t} = \{\mathbf{g}_{jk}\}_{1\le j\le s, 1\le k\le t}$$

$$\begin{aligned}
S_{F\dot\otimes(G\dot\otimes H)} &= \left\{ \frac{1}{\sum_{l=1}^{n} \frac{g_{jk,l}f_{i,l}}{p_l}} \cdot \begin{bmatrix} \frac{g_{jk,1}f_{i,1}}{p_1} \\ \dots \\ \frac{g_{jk,n}f_{i,n}}{p_n} \end{bmatrix} \right\}_{1\le i\le r, 1\le j\le s, 1\le k\le t} \\
&= \left\{ \frac{\sum_{m_1=1}^{n} \frac{g_{j,m_1}h_{k,m_1}}{p_{m_1}}}{\sum_{l=1}^{n} \sum_{m_2=1}^{n} \frac{g_{j,m_2}h_{k,m_2}}{p_{m_2}} \cdot \frac{g_{j,l}h_{k,l}}{p_l} \cdot \frac{f_{i,l}}{p_l}} \begin{bmatrix} \frac{g_{j,1}h_{k,1}}{p_1} \cdot \frac{f_{i,1}}{p_1} \\ \dots \\ \frac{g_{j,n}h_{k,n}}{p_n} \cdot \frac{f_{i,n}}{p_n} \end{bmatrix} \right\} \\
&= \left\{ \frac{1}{\sum_{l=1}^{n} \frac{f_{i,l}g_{j,l}h_{k,l}}{p_l^2}} \begin{bmatrix} \frac{f_{i,1}g_{j,1}h_{k,1}}{p_1^2} \\ \dots \\ \frac{f_{i,n}g_{j,n}h_{k,n}}{p_n^2} \end{bmatrix} \right\}_{1\le i\le r, 1\le j\le s, 1\le k\le t} \\
&= \left\{ \frac{\sum_{m_1=1}^{n} \frac{f_{i,m_1}g_{j,m_1}}{p_{m_1}}}{\sum_{l=1}^{n} \sum_{m_2=1}^{n} \frac{f_{i,m_2}g_{j,m_2}}{p_{m_2}} \cdot \frac{f_{i,l}g_{j,l}}{p_l} \cdot \frac{h_{k,l}}{p_l}} \begin{bmatrix} \frac{f_{i,1}g_{j,1}}{p_1} \cdot \frac{h_{k,1}}{p_1} \\ \dots \\ \frac{f_{i,n}g_{j,n}}{p_n} \cdot \frac{h_{k,n}}{p_n} \end{bmatrix} \right\} \\
&= \left\{ \frac{1}{\sum_{l=1}^{n} \frac{f_{ij,l}h_{k,l}}{p_l}} \cdot \begin{bmatrix} \frac{f_{ij,1}h_{k,1}}{p_1} \\ \dots \\ \frac{f_{ij,n}h_{k,n}}{p_n} \end{bmatrix} \right\}_{1\le i\le r, 1\le j\le s, 1\le k\le t} \\
&= S_{(F\dot\otimes G)\dot\otimes H}
\end{aligned}$$

One can check that the probabilities given to each points of both supports match.
Another way to look at this is if U and V are independent knowing X and if W independent of $P_{X|U}$ and $P_{X|V}$, then U, V, W are independent knowing X. Thus we have

$$\begin{aligned}
&P_{X|(U,V),W} \sim (F \mathbin{\dot\otimes} G) \mathbin{\dot\otimes} H \\
&= P_{X|U,(V,W)} \sim F \mathbin{\dot\otimes} (G \mathbin{\dot\otimes} H)
\end{aligned}$$

Claim 5 $(\lambda F_1 + (1-\lambda)F_2) \dot{\otimes} G = \lambda F_1 \dot{\otimes} G + (1-\lambda) F_2 \dot{\otimes} G$, $\lambda \in [0,1]$

Let $P_{X|U_i} \sim F_i$, $i = 1, 2$, $P_{X|V} \sim G$ having supports $\{\mathbf{f}_1, \ldots, \mathbf{f}_r\}$, $\{\mathbf{g}_1, \ldots, \mathbf{g}_s\}$, $\{\mathbf{h}_1, \ldots, \mathbf{h}_t\}$, with their associated probabilities $\alpha_1, \ldots, \alpha_r$, $\beta_1, \ldots, \beta_s$, $\gamma_1, \ldots, \gamma_t$ respectively, all centered in P_X, following the assumption that U_1, U_2, V are independent knowing X and $\lambda \in [0,1]$.
Define $B \sim B(\lambda)$ so that :

$$P_{X|U_B,B} \sim \lambda F_1 + (1-\lambda) F_2$$

and B independent of U_1, U_2, V and X. Then we have the following :

$$\begin{aligned} &P_{X|(U_B,B),V} \sim (\lambda F_1 + (1-\lambda)F_2) \dot{\otimes} G \\ &= P_{X|(U_B,V),B} \sim \lambda F_1 \dot{\otimes} G + (1-\lambda) F_2 \dot{\otimes} G \end{aligned}$$

3 Finding an inverse operator

As we defined a kind of "product operator", it justifies the idea of finding an inverse operator, a "division operator". We mainly focused on the following problem :

Given P_X, $P_{X|U} \sim F_1$ and some distribution F, can we find F_2 s.t.
$F_1 \dot{\otimes} F_2 = F$?

Of course this is not always feasible.

Some first attempt by an interesting example

Let $X \sim B(\frac{1}{2})$, $U_1, U_2, U_3 \overset{i.i.d.}{\sim} B(p)$ and $Y_i = X \oplus U_i$, $i = 1, 2, 3$. We compute $P_{X|Y_1,Y_2,Y_3} \sim G = F \dot{\otimes} F \dot{\otimes} F$ and the question is : does it exists $F_1 \neq F \dot{\otimes} F$ such that $F_1 \dot{\otimes} F = G$. We only care about the case for $X = 0$ as the other case can be resolved from this one as the second coordinate from each vector will be $1 -$ "$first$". To reduce notation we only use the first coordinate in all the following supports. We have :

$$\begin{aligned} &S_F = \{1-p, p\} \\ &S_G = \left\{ \frac{(1-p)^3}{(1-p)^3 + p^3}, 1-p, p, \frac{p^3}{p^3 + (1-p)^3} \right\} \end{aligned}$$

and we have to find S_{F_1}. The first problem arises here. We can not know precisely on how many points F_1 will be distributed. Let us assume that :

$$S_{F_1} = \{a, b\}$$

for some a and b each having probability α and β respectively. With the previous formula, we find :

$$\begin{aligned}S_{F_1 \dot{\otimes} F} = \{ & \frac{ap}{ap+(1-a)(1-p)}, \frac{a(1-p)}{a(1-p)+(1-a)p}, \\ & \frac{bp}{bp+(1-b)(1-p)}, \frac{b(1-p)}{b(1-p)+(1-b)p} \}\end{aligned}$$

We solve for a and b and we find :

$$\begin{aligned}a &= \frac{p^2}{p^2+(1-p)^2} \\ b &= \frac{(1-p)^2}{p^2+(1-p)^2}\end{aligned}$$

We find the probabilities for α, β such that F_1 is centered in P_X. We find :

$$\begin{aligned}\alpha &= \frac{1}{2} \\ \beta &= \frac{1}{2}\end{aligned}$$

Here arises a second problem : we found points for F_1 such that $S_G = S_{F_1 \dot{\otimes} F}$ but there is a mismatch regarding the probabilities. In fact here are the probabilities associated to the points of G :

$$\begin{aligned}\frac{(1-p)^3}{(1-p)^3+p^3} &\to \frac{1}{2}[(1-p)^3+p^3] \\ 1-p &\to \frac{3}{2}(1-p)^2 p \\ p &\to \frac{3}{2}p^2(1-p) \\ \frac{p^3}{(1-p)^3+p^3} &\to \frac{1}{2}[(1-p)^3+p^3]\end{aligned}$$

and here are the probabilities associated to the points of $F_1 \dot{\otimes} F$:

$$\begin{aligned}\frac{(1-p)^3}{(1-p)^3+p^3} &\to \frac{1}{2}\frac{(1-p)^3+p^3}{(p^2+(1-p)^2)} \\ 1-p &\to \frac{1}{2}\frac{(1-p)^2p+p^2(1-p)}{(p^2+(1-p)^2)} \\ p &\to \frac{1}{2}\frac{(1-p)^2p+p^2(1-p)}{(p^2+(1-p)^2)} \\ \frac{p^3}{(1-p)^3+p^3} &\to \frac{1}{2}\frac{(1-p)^3+p^3}{(p^2+(1-p)^2)}\end{aligned}$$

where $l \to r$ stands for *"point l has associated probability r"*. As we can see, the two distributions are not equal. Hence $F_1 \dot{\otimes} F \neq G(= F \dot{\otimes} F \dot{\otimes} F)$. The main fact why those distributions are not equal is that we looked for F_1 having two points and $F \dot{\otimes} F$ has three. One thing to note is that $F \dot{\otimes} F$ contains the two points of F_1 and a third point being exactly $\frac{1}{2}$ which is the "trivial" point coming from P_X (as reminder $X \sim B(\frac{1}{2}) \implies P_X = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$). This may be a coincidence coming from the big symmetry of the problem.

Second try from the same example

We tried to solve our first problem of number of points. In order to do so, we started from the same example but this time not looking for a distribution F_1 but for two distributions F_a and F_b on two points such that

$$F_a \dot{\otimes} F_b = F \dot{\otimes} F \dot{\otimes} F = G.$$

The intuition behind this is that in most of the cases, taking the operator between two distributions on two points should give us a distribution of four points. In fact $F \dot{\otimes} F \dot{\otimes} F$ gave us four points because of the symmetry of the example. Let us assume the following distributions for F_a and F_b :

F_a :

$$\begin{aligned} a_1 &\to \alpha_1 = \frac{1}{2} \\ a_2 &\to \alpha_2 = \frac{1}{2} \implies a_2 = 1 - a_1 \end{aligned}$$

F_b :

$$\begin{aligned} b_1 &\to \beta_1 = \frac{1}{2} \\ b_2 &\to \beta_2 = \frac{1}{2} \implies b_2 = 1 - b_1 \end{aligned}$$

We compute $F_a \dot{\otimes} F_b$ and we find :

$$\begin{aligned} \frac{a_1 b_1}{a_1 b_1 + (1-a_1)(1-b_1)} &\to \frac{1}{2}[a_1 b_1 + (1-a_1)(1-b_1)] \\ \frac{a_1(1-b_1)}{a_1(1-b_1) + (1-a_1)b_1} &\to \frac{1}{2}[a_1(1-b_1) + (1-a_1)b_1] \\ \frac{(1-a_1)b_1}{(1-a_1)b_1 + a_1(1-b_1)} &\to \frac{1}{2}[(1-a_1)b_1 + a_1(1-b_1)] \\ \frac{(1-a_1)(1-b_1)}{(1-a_1)(1-b_1) + a_1 b_1} &\to \frac{1}{2}[(1-a_1)(1-b_1) + a_1 b_1] \end{aligned}$$

It remains now to solve the system from the probabilities

$$\begin{aligned} a_1 b_1 + (1-a_1)(1-b_1) &= p^3 + (1-p)^3 \\ a_1(1-b_1) + (1-a_1)b_1 &= 3[p^2(1-p) + (1-p)^2 p] \end{aligned}$$

The solution to this system remains unclear because the two equations are linearly dependent and also we cannot find another equation to solve the system.

We couldn't find a way of describing this "divide" operator. It is also unclear if it exists or not.

4 Decomposition

We focused our research on another perspective. The goal was to be able to apply the operator with any distribution. The main concern was the assumption we made at the beginning of our research : U and V are independent knowing X. We made two claims that remain to be proven in order to achieve what we wanted. Let $P_X \in S$ and let $Y_1, \ldots, Y_m$ be m random variables.

Claim 6 Existence of the decomposition

$\exists U_1, \ldots, U_k$, $\forall i$ U_i has support on two points, is centered in P_X and $U_1, \ldots, U_k$ are fully independent knowing X, such that

$$P_{X|Y_1,\ldots,Y_n} = P_{X|U_1,\ldots,U_k}.$$

If we denote $P_{X|U_i} \sim F_i$, $i = 1, \ldots, k$, then

$$\begin{aligned} P_{X|Y_1,\ldots,Y_n} &= P_{X|U_1,\ldots,U_k} \\ &= F_1 \dot{\otimes} F_2 \dot{\otimes} \ldots \dot{\otimes} F_k \end{aligned}$$

and we call $F_1 \dot{\otimes} \ldots \dot{\otimes} F_k$ the decomposition of $P_{X|Y_1,\ldots,Y_n}$.

Claim 7 Uniqueness of the decomposition

The decomposition of $P_{X|Y_1,\ldots,Y_n}$ is unique.

As stated before, those claims remain not proven. Our intuition trying to prove them was that claim 6 could be true and claim 7 could be false.

5 Future directions

We state here some ideas on continuing the work about the operator that we created. For the sake of simplicity we state those ideas for P_X being unidimensionnal below, but they can be generalised to n dimensions.

Asymptotics
If we can compute $F \dot{\otimes} G$ in an efficient way, then we can do the following

procedure

$$F_0 = F$$
$$F_{n+1} = F_n \dot{\otimes} F_n.$$

This allows us to compute $F \dot{\otimes} F \dot{\otimes} \ldots \dot{\otimes} F$ with 2^n combinations. This could be efficient to study asymptotics.

Concavity

Let $p_x \in [0,1]$ be fixed. Let μ be a discrete probability distribution on support S_μ, averaging to p_x and let $p \in [0,1]$. We define $h_{p_x} : [0,1] \to [0,1]$ a function having the following properties :

- h_{p_x} is affine on $[0, p_x]$ and $[p_x, 1]$
- $h_{p_x}(0) = h_{p_x}(1) = 0$
- $h_{p_x}(p_x) = 1$

We also define for any distribution F the following :

$$F(h_{p_x}) = \sum_{q \in S_F} h_{p_x}(q) F(q)$$

Then we claim that

$$h^{'}(p) = (p \dot{\otimes} \mu)(h_{p_x})$$

has the same properties as h_{p_x}, where the operator between a point and a distribution follows exactly our above definition.

Inner product

Let μ and ν be two probability distributions averaging to p_x. Then we define the inner product of μ and ν as

$$\langle \mu, \nu \rangle = (\mu \dot{\otimes} \nu)(h_{p_x})$$

It is interesting to study this inner product as it gives an inner product structure to our measure space.

Last problem

Here is a last problem that we want to explore :
Given h, is there μ_h such that

$$(p \dot{\otimes} \mu_h)(h_{p_x}) = h(p), \ \forall p$$

6 Conclusion

In conclusion, we created an operator that reflects the gain of information from independent samples. This operator has some interesting properties and there still need some work to find new ones. It also can be easily generalised for some continuous distribution P_X.