Introduction cours n°2 du 18/09/2020

1. Qu'est-ce qu'un bit d'information ?
2. Qu'est ce que l'entropie d'une variable aléatoire réelle ?
3. Qu'elle est l'entropie d'une variable aléatoire réelle suivant la loi de Bernouilli de paramètre $p \in [0, 1]$?
4. Pour une variable aléatoire réelle de support fini de cardinal N par quelle quelle valeur est majorée son entropie ?
5. Si la contrainte de support est levée pour une contrainte de moyenne le résultat est-il toujours valable ? (Lois uniformes à support inclus dans $\mathbb{N}$ dont l'espérance vaut 1)
6. Comment minimiser le nombre de question d'un questionnaire ?

Entropie d'une loi géométrique $\mathcal{G}(p)$

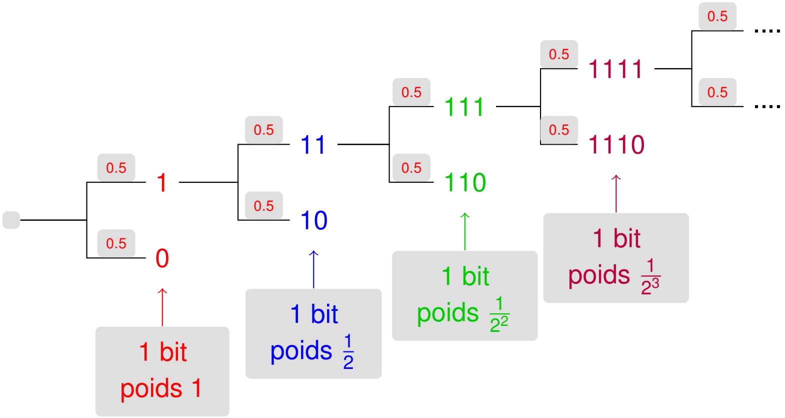
1. rappel $(\forall n \in \mathbb{N}^*)\, p(X=n) = q^{n-1}p$
2. Calcul de l'entropie

$$
\begin{aligned}
H(X) &= -\sum_{n=1}^{+\infty} q^{n-1} p log_2(q^{n-1}p) \\
&= -p\left[\sum_{n=1}^{+\infty} q^{n-1} log_2(q^{n-1}) + \sum_{n=1}^{+\infty} q^{n-1} log_2(p)\right] \\
&= -p\left[\sum_{n=1}^{+\infty} (n-1) q^{n-1} log_2(q) + \frac{1}{1-q} log_2(p)\right] \\
&= -p\left[\frac{q}{(1-q)^2} log_2(q) + \frac{1}{p} log_2(p)\right] \\
&= -p\left[\frac{q}{p^2} log_2(q) + \frac{1}{p} log_2(p)\right] \\
&= \frac{-q log_2(q) - p log_2(p)}{p}
\end{aligned}
$$

$$\boxed{H(\mathcal{G}(p)) = \frac{H(\mathcal{B}(p))}{p}}$$

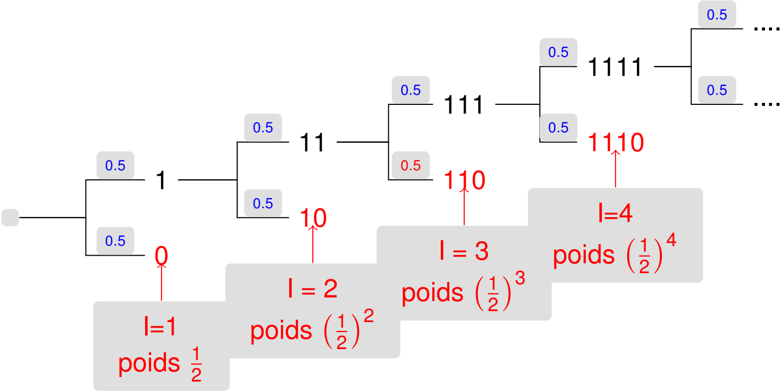
Entropie d'une loi géométrique $\mathcal{G}\left(\frac{1}{2}\right)$

$$H\left(\mathcal{G}\left(\tfrac{1}{2}\right)\right) = \left\{ \begin{array}{c} 2 \times H\left(\mathcal{B}\left(\tfrac{1}{2}\right)\right) = 2 \\ 1 + \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \cdots = \sum_{n\in\mathbb{N}} \frac{1}{2^n} = 2 \end{array} \right.$$



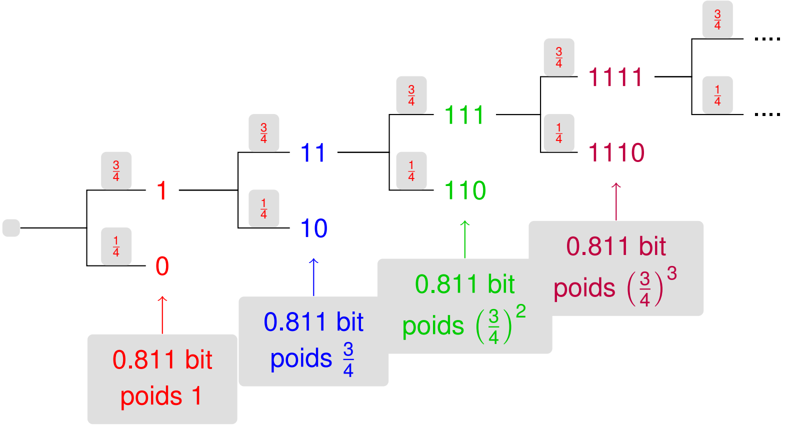
Loi géométrique $\mathcal{G}\left(\frac{1}{2}\right)$: longueur moyenne des mots

$$\begin{aligned} L &= 1 \times \frac{1}{2} + 2 \times \frac{1}{2^2} + 3 \times \frac{1}{2^3} + \cdots \\ &= \sum_{n \in \mathbb{N}} n \times \frac{1}{2^n} \\ &= \;= \frac{\frac{1}{2}}{\left(1 - \frac{1}{2}\right)^2} \\ &= 2 \end{aligned}$$



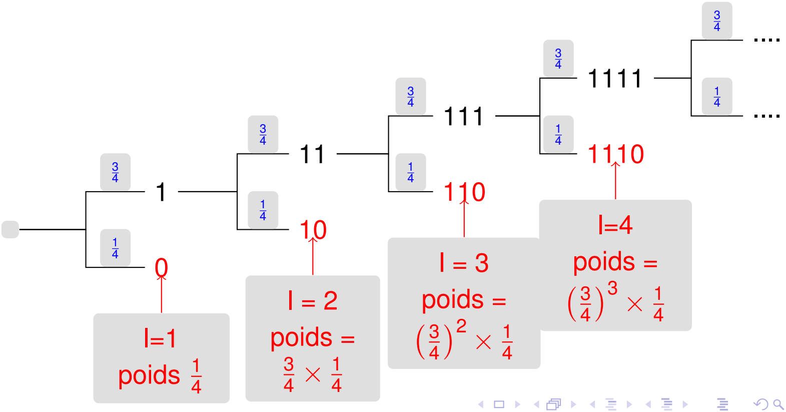
Entropie d'une loi géométrique $\mathcal{G}\left(\frac{1}{4}\right)$

$$H\left(\mathcal{G}\left(\tfrac{1}{4}\right)\right) = \begin{cases} 4 \times H\left(\mathcal{B}\left(\tfrac{1}{4}\right)\right) = 4 \times 0.811 \\ 0.811\left(1 + \tfrac{3}{4} + \left(\tfrac{3}{4}\right)^2 + \left(\tfrac{3}{4}\right)^3 + \cdots\right) = 0.811 \sum_{n\in\mathbb{N}} \left(\tfrac{3}{4}\right)^n \end{cases}$$



$\mathcal{G}\left(\frac{1}{4}\right)$: longueur moyenne des mots

$$
\begin{aligned}
L &= 1\times\frac{1}{4}+2\times\frac{3}{4}\times\frac{1}{4}+3\times\left(\frac{3}{4}\right)^2\times\frac{1}{4}+\cdots\\
&= \frac{1}{4}\sum_{n\in\mathbb{N}^*} n\times\left(\frac{3}{4}\right)^{n-1}\\
&= \frac{1}{4}\times\frac{1}{\left(1-\frac{3}{4}\right)^2}\\
&= 4 \qquad \left[= E\left(\mathcal{G}\left(\frac{1}{4}\right)\right)\right]
\end{aligned}
$$



Inégalité de Gibbs

Soit $P = \{p_k; k \in [[1, n]]\}$ et $Q = \{q_k; k \in [[1, n]]\}$ deux lois de probabilité sur un même support de cardinal k

$$\boxed{\sum_{k=1}^{n}, p_k log_2 \left(\frac{q_k}{p_k}\right) \leq 0}$$

couple de V.A.R à support fini : loi de probabilité, loi marginales

Soit $Z = (X, Y)$ un couple de VAR de support $\mathcal{A}$.
loi de $Z = (X, Y)$:

$$\{p((X, Y) = (x, y)) : (x, y) \in \mathcal{A}\}$$

X et Y sont les **VAR marginales** du couple.
Leurs support respectifs sont notés $\mathcal{A}_X$ et $\mathcal{A}_Y$.

$$\mathcal{A} = \mathcal{A}_X \times \mathcal{A}_Y$$

Les lois marginales :

1. $(\forall x \in \mathcal{A}_X) p(X = x) = \sum_{y \in \mathcal{A}_Y} P((X, Y) = (x, y))$
2. $(\forall y \in \mathcal{A}_y) p(Y = y) = \sum_{x \in \mathcal{A}_X} P((X, Y) = (x, y))$

notations : $p(X = x)$ ou $p_X(x)$ sera noté $p(x)$ lorsqu'il n'y aura pas de confusion possible , de même $p(X = x, Y = y)$ ou $p_{X,Y}(x, y)$ sera noté $p(x, y)$

couple de V.A.R à support fini : lois de probabilité conditionnelles

$$\boxed{p(X = x | Y = y) = \frac{p(X=x, Y=y)}{p(Y=y)}}$$

1. $p(X = x | Y = y) = \frac{p(X=x, Y=y)}{\sum_{x \in \mathcal{A}_X} p(x,y)}$
2. $p(x, y) = p(X = x | Y = y) p(Y = y))$

Indépendance

1. $(\forall (x, y) \in \mathcal{A}) p(X = x | Y = y) = p(X = x)$
2. $(\forall (x, y) \in \mathcal{A}) p(x, y) = p(X = x) p(Y = y)$

Entropie conjointe et Entropies marginales

Entropie conjointe :

$$H(X,Y) = -\sum_{x \in \mathcal{A}_X} \sum_{y \in \mathcal{A}_Y} p(x,y) log_2 (p(x,y))$$

Entropies marginales

$$\begin{aligned} H(X) &= -\textstyle\sum_{x \in \mathcal{A}_X} p(x) log_2 (p(x)) \\ H(Y) &= -\textstyle\sum_{y \in \mathcal{A}_y} p(y) log_2 (p(y)) \end{aligned}$$

Relation entre Entropie conjointe et Entropies marginales

Entropie conjointe et entropies marginales lorsque X et Y sont indépendantes

$$H(X,Y) = H(X) + H(Y)$$

Entropie conjointe et entropies marginales

$$\boxed{H(X,Y) \leq H(X) + H(Y)}$$

$$H(X,Y) = H(X) + H(Y) + \underbrace{\sum_{x \in \mathcal{A}_X} \sum_{y \in \mathcal{A}_Y} p(x,y) log_2 \left(\frac{p(x)p(y)}{p(x,y)} \right)}_{<0 \ (cf \ \ ingalité \ \ de \ \ Gibbs)}$$

Divergence de Kullback- Liebler ou entropie relative

Soit deux loi $\mathcal{P}$ et $\mathcal{Q}$ définies sur un même support $\mathcal{A}$, la divergence de Kullback-Liebler est donnée par

$$\boxed{\mathcal{D}(\mathcal{P}||\mathcal{Q}) = \sum_{x\in\mathcal{A}} p(x) log_2 \left(\frac{p(x)}{q(x)}\right)}$$

avec les conventions
$0 log_2 \left(\frac{0}{0}\right) = 0, \; 0 log_2 \left(\frac{0}{q}\right) = 0 \; et \; p log_2 \left(\frac{p}{0}\right) = \infty$

1. $\mathcal{D}(\mathcal{P}||\mathcal{Q}) \geq 0$
2. $\mathcal{D}(\mathcal{P}||\mathcal{Q}) = 0$ si et seulement si $\mathcal{P}$ et $\mathcal{Q}$ sont identiques

Information mutuelle

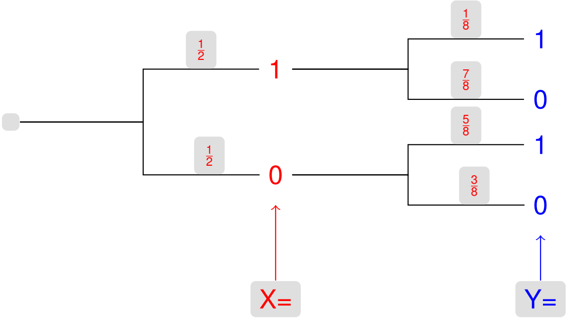
L'information mutuelle $I(X, Y)$ est la divergence de Kullback entre la loi conjointe et le produit de ces marginales.

$$\boxed{I(X, Y) = \sum_{x \in \mathcal{A}_X} \sum_{y \in \mathcal{A}_Y} p(x, y) log_2 \left(\frac{p(x,y)}{p(x)p(y)} \right)}$$

Entropie, entropies marginales, information mutuelle

$$\boxed{H(X,Y) = H(X) + H(Y) - I(X,Y)}$$

Exemple $H(X, Y) = H(X) + H(Y) - I(X, Y)$



$$H(X, Y) = 1.748999$$

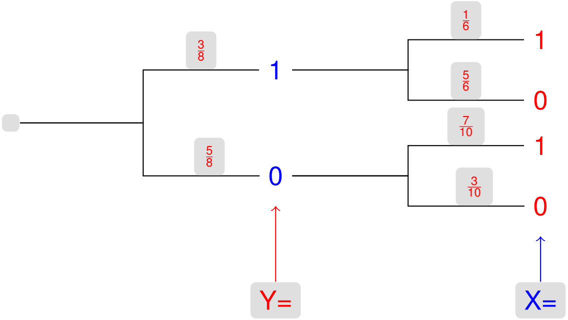
$$-\frac{1}{16} log_2\left(\frac{1}{16}\right) - \frac{7}{16} log_2\left(\frac{7}{16}\right) - \frac{5}{16} log_2\left(\frac{5}{16}\right) - \frac{3}{16} log_2\left(\frac{3}{16}\right)$$

$$4\left(\frac{1}{16} + \frac{7}{16} + \frac{5}{16} + \frac{3}{16}\right) - \frac{1}{16}\left(7 log_2(7) + 5 log_2(5) + 3 log_2(3)\right)$$

$$4 - \frac{7 log_2(7) + 5 log_2(5) + 3 log_2(3)}{16}$$

$$H(X) = 1$$

Exemple $H(X,Y) = H(X) + H(Y) - I(X,Y)$



$$
\begin{aligned}
H(Y) &= -\frac{3}{8} log_2\left(\frac{3}{8}\right) - \frac{5}{8} log_2\left(\frac{5}{8}\right) \\
&= 3\left(\frac{3}{8} + \frac{5}{8}\right) - \frac{3log_2(3)+5log_2(5)}{8} \\
&= 3 - \frac{3log_2(3)+5log_2(5)}{8} \\
&= 0.954434
\end{aligned}
$$

$$
\begin{aligned}
I(X,Y) &= 1.954434 - 1.748999 \\
&= 0.205435
\end{aligned}
$$

Entropie conditionnelle

$$H(Y|X=x) = -\sum_{y \in \mathcal{A}_Y} p(y|x) log_2\left(p(y|x)\right)$$

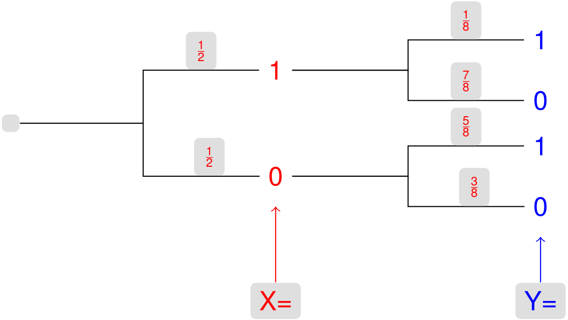
$$H(Y|X) = -\sum_{x \in \mathcal{A}_X} p(x) H(Y|X=x)$$

$$\boxed{H(Y|X) = -\sum_{x \in \mathcal{A}_X} \sum_{y \in \mathcal{A}_Y} p(x,y) log_2(y|x)}$$

Cas particulier

1. pour le couple (X,X) on a $H(X,X) = 0$
   $(\forall (x_i, x_j) \in \mathcal{A}_X^2) p(x_i|x_j) = \delta_{ij}$, et donc $(\forall x_j \in \mathcal{A}_X) H(X|x_j) = 0$
2. pour un couple (X,Y) indépendant $H(Y|X) = H(Y)$:
   $(\forall (x_i, y_j) \in \mathcal{A}_X \times \mathcal{A}_Y) p(y_j|x_i) = p(y_j)$, et donc
   $(\forall x_i \in \mathcal{A}_X) H(Y|x_i) = H(Y)$

Exemple $H(Y|X)$



$$
\begin{aligned}
H(Y|X=1) &= -\frac{1}{8}log_2\left(\frac{1}{8}\right) - \frac{7}{8}log_2\left(\frac{7}{8}\right)\\
&= 3 - \frac{7log_2(7)}{8}\\
&= 0.54356\\
H(Y|X=0) &= -\frac{5}{8}log_2\left(\frac{5}{8}\right) - \frac{3}{8}log_2\left(\frac{3}{8}\right)\\
&= 3 - \frac{5log_2(5)-3log_2(3)}{8}\\
&= 0.95443
\end{aligned}
$$

$$H(Y|X) = 0.748995$$