20.5/23

Ayoub Echchahed
(111 274 558)

Théorie de l'information
GEL-7062

Devoir 3
Résolution de Problèmes

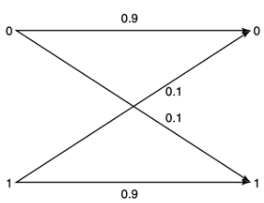
Travail présenté à
Mr. Jean-Yves Chouinard

Faculté des sciences et de génie
Université Laval
Hiver 2022

- **Problème 1: Séquences Conjointement Typiques** 4/4

Jointly typical sequences. As we did in Problem 3.13 for the typical set for a single random variable, we will calculate the jointly typical set for a pair of random variables connected by a binary symmetric

channel, and the probability of error for jointly typical decoding for such a channel.



We consider a binary symmetric channel with crossover probability 0.1. The input distribution that achieves capacity is the uniform distribution [i.e., $p(x) = (\frac{1}{2}, \frac{1}{2})$], which yields the joint distribution $p(x, y)$ for this channel is given by

$X \backslash Y$	0	1
0	0.45	0.05
1	0.05	0.45

The marginal distribution of Y is also $(\frac{1}{2}, \frac{1}{2})$.

(a) Calculate $H(X)$, $H(Y)$, $H(X, Y)$, and $I(X; Y)$ for the joint distribution above.

(b) Let $X_1, X_2, \ldots, X_n$ be drawn i.i.d. according the Bernoulli($\frac{1}{2}$) distribution. Of the 2^n possible input sequences of length n, which of them are typical [i.e., member of $A_\epsilon^{(n)}(X)$ for $\epsilon = 0.2$]? Which are the typical sequences in $A_\epsilon^{(n)}(Y)$?

(c) The jointly typical set $A_\epsilon^{(n)}(X, Y)$ is defined as the set of sequences that satisfy equations (7.35-7.37). The first two equations correspond to the conditions that x^n and y^n are in $A_\epsilon^{(n)}(X)$ and $A_\epsilon^{(n)}(Y)$, respectively. Consider the last condition, which can be rewritten to state that $-\frac{1}{n}\log p(x^n, y^n) \in (H(X, Y) - \epsilon, H(X, Y) + \epsilon)$. Let k be the number of places in which the sequence x^n differs from y^n (k is a function of the two sequences). Then we can write

$$p(x^n, y^n) = \prod_{i=1}^{n} p(x_i, y_i) \quad (7.156)$$

$$= (0.45)^{n-k}(0.05)^k \quad (7.157)$$

$$= \left(\frac{1}{2}\right)^n (1-p)^{n-k} p^k. \quad (7.158)$$

An alternative way at looking at this probability is to look at the binary symmetric channel as in additive channel $Y = X \oplus Z$, where Z is a binary random variable that is equal to 1 with probability p, and is independent of X. In this case,

$$p(x^n, y^n) = p(x^n)p(y^n|x^n) \quad (7.159)$$

$$= p(x^n)p(z^n|x^n) \quad (7.160)$$

$$= p(x^n)p(z^n) \quad (7.161)$$

$$= \left(\frac{1}{2}\right)^n (1-p)^{n-k} p^k. \quad (7.162)$$

Show that the condition that (x^n, y^n) being jointly typical is equivalent to the condition that x^n is typical and $z^n = y^n - x^n$ is typical.

(d) We now calculate the size of $A_\epsilon^{(n)}(Z)$ for $n = 25$ and $\epsilon = 0.2$. As in Problem 3.13, here is a table of the probabilities and numbers of sequences with k ones:

k	$\binom{n}{k}$	$\binom{n}{k}p^k(1-p)^{n-k}$	$-\frac{1}{n}\log p(x^n)$
0	1	0.071790	0.152003
1	25	0.199416	0.278800
2	300	0.265888	0.405597
3	2300	0.226497	0.532394
4	12650	0.138415	0.659191
5	53130	0.064594	0.785988
6	177100	0.023924	0.912785
7	480700	0.007215	1.039582
8	1081575	0.001804	1.166379
9	2042975	0.000379	1.293176
10	3268760	0.000067	1.419973
11	4457400	0.000010	1.546770
12	5200300	0.000001	1.673567

[Sequences with more than 12 ones are omitted since their total probability is negligible (and they are not in the typical set).] What is the size of the set $A_\epsilon^{(n)}(Z)$?

(e) Now consider random coding for the channel, as in the proof of the channel coding theorem. Assume that 2^{nR} codewords $X^n(1), X^n(2), \ldots, X^n(2^{nR})$ are chosen uniformly over the 2^n possible binary sequences of length n. One of these codewords is chosen and sent over the channel. The receiver looks at the received sequence and tries to find a codeword in the code that is jointly typical with the received sequence. As argued above, this corresponds to finding a codeword $X^n(i)$ such that $Y^n - X^n(i) \in A_\epsilon^{(n)}(Z)$. For a fixed codeword $x^n(i)$, what is the probability that the received sequence Y^n is such that $(x^n(i), Y^n)$ is jointly typical?

(f) Now consider a particular received sequence $y^n = 000000\ldots0$, say. Assume that we choose a sequence X^n at random, uniformly distributed among all the 2^n possible binary n-sequences. What is the probability that the chosen sequence is jointly typical with this y^n? [*Hint*: This is the probability of all sequences x^n such that $y^n - x^n \in A_\epsilon^{(n)}(Z)$.]

(g) Now consider a code with $2^9 = 512$ codewords of length 12 chosen at random, uniformly distributed among all the 2^n sequences of length $n = 25$. One of these codewords, say the one corresponding to $i = 1$, is chosen and sent over the channel. As calculated in part (e), the received sequence, with high probability, is jointly typical with the codeword that was sent. What is the probability that one or more of the other codewords (which were chosen at random, independent of the sent codeword) is jointly typical with the received sequence? [*Hint*: You could use the union bound, but you could also calculate this probability exactly, using the result of part (f) and the independence of the codewords.]

(h) Given that a particular codeword was sent, the probability of error (averaged over the probability distribution of the channel and over the random choice of other codewords) can be written as

$$\Pr(\text{Error}|x^n(1) \text{ sent}) = \sum_{y^n: y^n \text{causes error}} p(y^n|x^n(1)). \quad (7.163)$$

There are two kinds of error: the first occurs if the received sequence y^n is not jointly typical with the transmitted codeword, and the second occurs if there is another codeword jointly typical with the received sequence. Using the result of the preceding parts, calculate this probability of error. By

the symmetry of the random coding argument, this does not depend on which codeword was sent.

The calculations above show that average probability of error for a random code with 512 codewords of length 25 over the binary symmetric channel of crossover probability 0.1 is about 0.34. This seems quite high, but the reason for this is that the value of ϵ that we have chosen is too large. By choosing a smaller ϵ and a larger n in the definitions of $A_\epsilon^{(n)}$, we can get the probability of error to be as small as we want as long as the rate of the code is less than $I(X; Y) - 3\epsilon$.

Also note that the decoding procedure described in the problem is not optimal. The optimal decoding procedure is maximum likelihood (i.e., to choose the codeword that is closest to the received sequence). It is possible to calculate the average probability of error for a random code for which the decoding is based on an approximation to maximum likelihood decoding, where we decode a received sequence to the unique codeword that differs from the received sequence in ≤ 4 bits, and declare an error otherwise. The only difference with the jointly typical decoding described above is that in the case when the codeword is equal to the received sequence! The average probability of error for this decoding scheme can be shown to be about 0.285.

e) For a fixed codeword $x^n(i)$, **what is the probability** that the **received sequence Y^n is such that $(x^n(i), Y^n)$ is jointly typical?**

Correspond to finding a codeword $X^n(i)$ such that $Y^n - X^n(i) \in A_\varepsilon^{(n)}(Z)$.

Hence Probability = Probability that the noise sequence is typical in $A_\varepsilon^{(n)}(Z)$

Probability sequence typical = 0.8302 (TP2) 1/1

Probability received sequence not jointly typical with transmitted codeword = 1 - 0.8302 = 0.1698

f) What is the probability that the chosen sequence is jointly typical with this y^n?

= Probability of all sequences x^n such that $y^n - x^n \in A_\varepsilon^{(n)}(Z)$

- Probability of choosing x^n sequences = $(1 / 2)^n$
- Hence probability chosen sequence is jointly typical with this y^n = Number jointly typical $(x^n, y^n) * (1 / 2)^n$
- Number jointly typical $(x^n, y^n) = | A_\varepsilon^{(n)}(Z) |$
- Thus:

$P = | A_\varepsilon^{(n)}(Z) | * (1 / 2)^n$

$P = 15\,275 * 2^{-25}$ 1/1

$P = 4.552305 \times 10^{-4}$

g) What is probability that one or more of the other codewords (which were chosen at random, independently of the sent codeword) is jointly typical with the received sequence?

- Code with $2^9 = 512$ codewords
- Length 12
- Chosen at random
- Uniformly distributed among all the 2^n sequences of length n = 25

P (Other codewords jointly typical with received sequence) = 4.552305×10^{-4} (See above)

P (No left codewords jointly typical with received sequence) = $1 - (4.552305 \times 10^{-4})^{511}$
= 0.7924

P (Minimum 1 jointly typical with received sequence) = 1 − 0.7924 1/1
= 0.2076

h) Calculate the probability of error given that a particular codeword was sent.

- Error 1: received sequence y^n is not jointly typical with the transmitted codeword
- Error 2: if there is another codeword jointly typical with the received sequence.
- Two errors are conditionally independent given received sequence
- From the previous calculations:

P (Error 1) = 0.1698

P (Error 2) = 0.2076

(As by the symmetry of the random coding argument, this does not depend on which codeword was sent)

P (Error) ≤ (0.1698 + 0.2076)

P (Error) ≤ 0.3774 1/1

- **Problème 2: Distorsion Moyenne et Information Mutuelle** 3/5

Une source d'information sans mémoire X génère des symboles quaternaires avec la distribution : $p(x_1) = \frac{3}{7}$, $p(x_2) = \frac{1}{7}$, $p(x_3) = \frac{1}{7}$ et $p(x_4) = \frac{2}{7}$. Un encodeur de source avec pertes quaternaire de mesure de distorsion **d** est utilisé pour réduire le débit d'information à transmettre. Écrivez un programme pour effectuer l'algorithme de Blahut-Arimoto pour le calcul de la fonction de débit-distorsion. Donnez le listing de votre programme.

$$\mathbf{d} = \begin{bmatrix} d(x_1,\hat{x}_1) & d(x_2,\hat{x}_1) & d(x_3,\hat{x}_1) & d(x_4,\hat{x}_1) \\ d(x_1,\hat{x}_2) & d(x_2,\hat{x}_2) & d(x_3,\hat{x}_2) & d(x_4,\hat{x}_2) \\ d(x_1,\hat{x}_3) & d(x_2,\hat{x}_3) & d(x_3,\hat{x}_3) & d(x_4,\hat{x}_3) \\ d(x_1,\hat{x}_4) & d(x_2,\hat{x}_4) & d(x_3,\hat{x}_4) & d(x_4,\hat{x}_4) \end{bmatrix} = \begin{bmatrix} 0.7 & 2.3 & 3.2 & 5.4 \\ 6.3 & 1.4 & 3.9 & 4.3 \\ 2.5 & 4.3 & 0.7 & 9.2 \\ 2.7 & 4.4 & 8.5 & 1.1 \end{bmatrix}$$

a) Donnez l'entropie de la source, $H(X)$.

b) Considérez la paire de codeur et décodeur de source avec perte (pseudocanal 1) caractérisée par la matrice de probabilité de transition :

$$\mathbf{P}_1 = \begin{bmatrix} p(\hat{x}_1|x_1) & p(\hat{x}_1|x_2) & p(\hat{x}_1|x_3) & p(\hat{x}_1|x_4) \\ p(\hat{x}_2|x_1) & p(\hat{x}_2|x_2) & p(\hat{x}_2|x_3) & p(\hat{x}_2|x_4) \\ p(\hat{x}_3|x_1) & p(\hat{x}_3|x_2) & p(\hat{x}_3|x_3) & p(\hat{x}_3|x_4) \\ p(\hat{x}_4|x_1) & p(\hat{x}_4|x_2) & p(\hat{x}_4|x_3) & p(\hat{x}_4|x_4) \end{bmatrix} = \begin{bmatrix} 0.7 & 0.1 & 0.1 & 0.1 \\ 0.1 & 0.7 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.7 & 0.1 \\ 0.1 & 0.1 & 0.1 & 0.7 \end{bmatrix}$$

Calculez la distortion moyenne $d(\mathbf{P}_1)$ et l'information mutuelle $I(\mathbf{P}_1) = I_1\left(X, \hat{X}\right)$.

c) Considérez maintenant la paire de codeur et décodeur de source avec perte (pseudocanal 2) caractérisée par la matrice de probabilité de transition :

$$\mathbf{P}_2 = \begin{bmatrix} p(\hat{x}_1|x_1) & p(\hat{x}_1|x_2) & p(\hat{x}_1|x_3) & p(\hat{x}_1|x_4) \\ p(\hat{x}_2|x_1) & p(\hat{x}_2|x_2) & p(\hat{x}_2|x_3) & p(\hat{x}_2|x_4) \\ p(\hat{x}_3|x_1) & p(\hat{x}_3|x_2) & p(\hat{x}_3|x_3) & p(\hat{x}_3|x_4) \\ p(\hat{x}_4|x_1) & p(\hat{x}_4|x_2) & p(\hat{x}_4|x_3) & p(\hat{x}_4|x_4) \end{bmatrix} = \begin{bmatrix} 0.4 & 0.3 & 0.2 & 0.1 \\ 0.1 & 0.4 & 0.3 & 0.2 \\ 0.2 & 0.1 & 0.4 & 0.3 \\ 0.3 & 0.2 & 0.1 & 0.4 \end{bmatrix}$$

Calculez la distortion moyenne $d(\mathbf{P}_2)$ et l'information mutuelle $I(\mathbf{P}_2) = I_1\left(X, \hat{X}\right)$.

a) $H(X) = -\sum p(x_k) \log_2 p(x_k)$

$= -[(3/7)\log_2(3/7) + (1/7)\log_2(1/7) + (1/7)\log_2(1/7) + (2/7)\log_2(2/7)]$

$= -[(-0.52388) + (-0.40105) + (-0.40105) + (-0.516387)]$

$=$ 1.8424 Sh 1/1

b) $d(P_1) = \sum\sum p(x_k)\, p(\hat{x}_j \mid x_k)\, d(x_k, \hat{x}_j)$ 1/2

$= (3/7)[(0.7 * 0.7) + (0.1 * 6.3) + (0.1 * 2.5) + (0.1 * 2.7)]$
$+ (1/7)[(0.1 * 2.3) + (0.7 * 1.4) + (0.1 * 4.3) + (0.1 * 4.4)]$
$+ (1/7)[(0.1 * 3.2) + (0.1 * 3.9) + (0.7 * 0.7) + (0.1 * 8.5)]$
$+ (2/7)[(0.1 * 5.4) + (0.1 * 4.3) + (0.1 * 9.2) + (0.7 * 1.1)]$

$=$ 2.0529 1/1

$$I(P_1) = \sum\sum p(x_k)\, p(\hat{x}_j \mid x_k) \log_2 [p(\hat{x}_j \mid x_k) / \sum p(x_l)\, p(\hat{x}_j \mid x_l)]$$

$$\begin{aligned}
= (3/7)\ [&(0.7) * (\log_2 [(0.7) / [(3/7)(0.7) + (1/7)(0.1) + (1/7)(0.1) + (2/7)(0.1)]] \\
&+ (0.1) * (\log_2 [(0.1) / [(3/7)(0.1) + (1/7)(0.7) + (1/7)(0.1) + (2/7)(0.1)]] \\
&+ (0.1) * (\log_2 [(0.1) / [(3/7)(0.1) + (1/7)(0.1) + (1/7)(0.7) + (2/7)(0.1)]] \\
&+ (0.1) * (\log_2 [(0.1) / [(3/7)(0.1) + (1/7)(0.1) + (1/7)(0.1) + (2/7)(0.7)]]] \\
+ (1/7)\ [&(0.1) * (\log_2 [(0.1) / [(3/7)(0.7) + (1/7)(0.1) + (1/7)(0.1) + (2/7)(0.1)]] \\
&+ (0.7) * (\log_2 [(0.7) / [(3/7)(0.1) + (1/7)(0.7) + (1/7)(0.1) + (2/7)(0.1)]] \\
&+ (0.1) * (\log_2 [(0.1) / [(3/7)(0.1) + (1/7)(0.1) + (1/7)(0.7) + (2/7)(0.1)]] \\
&+ (0.1) * (\log_2 [(0.1) / [(3/7)(0.1) + (1/7)(0.1) + (1/7)(0.1) + (2/7)(0.7)]]] \\
+ (1/7)\ [&(0.1) * (\log_2 [(0.1) / [(3/7)(0.7) + (1/7)(0.1) + (1/7)(0.1) + (2/7)(0.1)]] \\
&+ (0.1) * (\log_2 [(0.1) / [(3/7)(0.1) + (1/7)(0.7) + (1/7)(0.1) + (2/7)(0.1)]] \\
&+ (0.7) * (\log_2 [(0.7) / [(3/7)(0.1) + (1/7)(0.1) + (1/7)(0.7) + (2/7)(0.1)]] \\
&+ (0.1) * (\log_2 [(0.1) / [(3/7)(0.1) + (1/7)(0.1) + (1/7)(0.1) + (2/7)(0.7)]]] \\
+ (2/7)\ [&(0.1) * (\log_2 [(0.1) / [(3/7)(0.7) + (1/7)(0.1) + (1/7)(0.1) + (2/7)(0.1)]] \\
&+ (0.1) * (\log_2 [(0.1) / [(3/7)(0.1) + (1/7)(0.7) + (1/7)(0.1) + (2/7)(0.1)]] \\
&+ (0.1) * (\log_2 [(0.1) / [(3/7)(0.1) + (1/7)(0.1) + (1/7)(0.7) + (2/7)(0.1)]] \\
&+ (0.7) * (\log_2 [(0.7) / [(3/7)(0.1) + (1/7)(0.1) + (1/7)(0.1) + (2/7)(0.7)]]]
\end{aligned}$$

$$= 0.1824 + 0.0127 + 0.00258 + 0.00307$$
$$= 0.20075$$

0/1

c) 1/2

$$d(P_2) = \sum\sum p(x_k)\, p(\hat{x}_j \mid x_k)\, d(x_k, \hat{x}_j)$$

$$\begin{aligned}
&= (3/7) [(0.4 * 0.7) + (0.1 * 6.3) + (0.2 * 2.5) + (0.3 * 2.7)] \\
&\quad + (1/7) [(0.3 * 2.3) + (0.4 * 1.4) + (0.1 * 4.3) + (0.2 * 4.4)] \\
&\quad + (1/7) [(0.2 * 3.2) + (0.3 * 3.9) + (0.4 * 0.7) + (0.1 * 8.5)] \\
&\quad + (2/7) [(0.1 * 5.4) + (0.2 * 4.3) + (0.3 * 9.2) + (0.4 * 1.1)]
\end{aligned}$$

$$= 3.05$$

1/1

$$I(P_2) = \sum\sum p(x_k)\, p(\hat{x}_j \mid x_k) \log_2 [p(\hat{x}_j \mid x_k) / \sum p(x_l)\, p(\hat{x}_j \mid x_l)]$$

$$\begin{aligned}
= (3/7)\ [&(0.4) * (\log_2 [(0.4) / [(3/7)(0.4) + (1/7)(0.3) + (1/7)(0.2) + (2/7)(0.1)]] \\
&+ (0.1) * (\log_2 [(0.1) / [(3/7)(0.1) + (1/7)(0.4) + (1/7)(0.3) + (2/7)(0.2)]] \\
&+ (0.2) * (\log_2 [(0.2) / [(3/7)(0.2) + (1/7)(0.1) + (1/7)(0.4) + (2/7)(0.3)]] \\
&+ (0.3) * (\log_2 [(0.3) / [(3/7)(0.3) + (1/7)(0.2) + (1/7)(0.1) + (2/7)(0.4)]]] \\
+ (1/7)\ [&(0.3) * (\log_2 [(0.3) / [(3/7)(0.4) + (1/7)(0.3) + (1/7)(0.2) + (2/7)(0.1)]] \\
&+ (0.4) * (\log_2 [(0.4) / [(3/7)(0.1) + (1/7)(0.4) + (1/7)(0.3) + (2/7)(0.2)]] \\
&+ (0.1) * (\log_2 [(0.1) / [(3/7)(0.2) + (1/7)(0.1) + (1/7)(0.4) + (2/7)(0.3)]] \\
&+ (0.2) * (\log_2 [(0.2) / [(3/7)(0.3) + (1/7)(0.2) + (1/7)(0.1) + (2/7)(0.4)]]] \\
+ (1/7)\ [&(0.2) * (\log_2 [(0.2) / [(3/7)(0.4) + (1/7)(0.3) + (1/7)(0.2) + (2/7)(0.1)]] \\
&+ (0.3) * (\log_2 [(0.3) / [(3/7)(0.1) + (1/7)(0.4) + (1/7)(0.3) + (2/7)(0.2)]] \\
&+ (0.4) * (\log_2 [(0.4) / [(3/7)(0.2) + (1/7)(0.1) + (1/7)(0.4) + (2/7)(0.3)]] \\
&+ (0.1) * (\log_2 [(0.1) / [(3/7)(0.3) + (1/7)(0.2) + (1/7)(0.1) + (2/7)(0.4)]]] \\
+ (2/7)\ [&(0.1) * (\log_2 [(0.1) / [(3/7)(0.4) + (1/7)(0.3) + (1/7)(0.2) + (2/7)(0.1)]] \\
&+ (0.2) * (\log_2 [(0.2) / [(3/7)(0.1) + (1/7)(0.4) + (1/7)(0.3) + (2/7)(0.2)]] \\
&+ (0.3) * (\log_2 [(0.3) / [(3/7)(0.2) + (1/7)(0.1) + (1/7)(0.4) + (2/7)(0.3)]] \\
&+ (0.4) * (\log_2 [(0.4) / [(3/7)(0.3) + (1/7)(0.2) + (1/7)(0.1) + (2/7)(0.4)]]]
\end{aligned}$$

$$= 0.0798 + 0.024865 + 0.01194 + 0.005815$$
$$= 0.12242$$

0/1

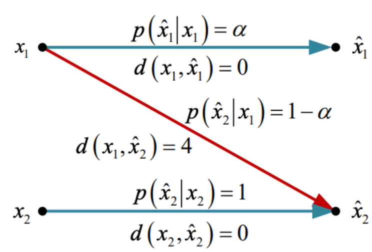
- **<u>Problème 3: Fonction de débit-distorsion R(D)</u>** 5.5/6

Problème 7.1 : Une source d'information sans mémoire X génère des bits avec la distribution : $p(x_1) = \frac{3}{4}$ et $p(x_2) = \frac{1}{4}$. Un encodeur de source avec pertes, avec un alphabet de reproduction binaire $\hat{X}$, est utilisé pour compresser ces données avant de les transmettre. La matrice des mesures de distortion **d** est donnée par :

$$\mathbf{d} = \begin{bmatrix} d(x_1,\hat{x}_1) & d(x_2,\hat{x}_1) \\ d(x_1,\hat{x}_2) & d(x_2,\hat{x}_2) \end{bmatrix} = \begin{bmatrix} 0 & \infty \\ 4 & 0 \end{bmatrix}$$

où une distortion infinie, i.e., $d(x_k,\hat{x}_j) = \infty$, indique qu'il n'y a pas de transition de x_k à $\hat{x}_j$. La matrice de probabilité de transition **P** est donnée par :

$$\mathbf{P} = \begin{bmatrix} p(\hat{x}_1|x_1) & p(\hat{x}_1|x_2) \\ p(\hat{x}_2|x_1) & p(\hat{x}_2|x_2) \end{bmatrix} = \begin{bmatrix} \alpha & 0 \\ (1-\alpha) & 1 \end{bmatrix}$$



a) Donnez l'entropie de la source, $H(X)$.
b) Déterminez les probabilités des symboles de reproduction en fonction de α.
c) Déterminez les valeurs de distortion minimale d_{min} et de distortion maximale d_{max}.
d) Dérivez l'expression de la distortion moyenne par symbole $d(X,\hat{X})$ en fonction de α.
e) Dérivez l'expression de l'information mutuelle du pseudocanal $I(X;\hat{X})$ en fonction de α.
f) Calculez la fonction de débit-distortion $R(D)$.
g) Dessinez la fonction de débit-distortion $R(D)$ de $(d_{min} - 2) \leq D \leq (d_{max} + 2)$.

a) $H(X) = -\sum p(x_k) \log_2 p(x_k)$

$= -[(3/4) \log_2 (3/4) + (1/4) \log_2 (1/4)]$

$= -[(-0.311278) + (-0.5)]$

$= 0.8113$ Sh 1/1

b) 1/1

$P(\hat{x}1) = [(3/4) * \alpha]$

$= 0.75\alpha$

$P(\hat{x}2) = [(3/4) * (1-\alpha)] + [(1/4) * (1)]$

$= 0.75(1-\alpha) + 0.25$

c) 1/1

$d_{min} = \sum p(x_k)\, d(x_k, \hat{x}_{j\,min})$

$= p(x_1)\, d(x_1, \hat{x}_1) + p(x_2)\, d(x_2, \hat{x}_2)$

$= [(3/4) * 0] + [(1/4) * 0]$

$= 0$

$d_{max} = {}_{(\min j = 1,2)} \sum p(x_k)\, d(x_k, \hat{x}_j)$
$= \text{Min}\, [p(x_1)\, d(x_1, \hat{x}_1) + p(x_2)\, d(x_2, \hat{x}_1)],$
$[p(x_1)\, d(x_1, \hat{x}_2) + p(x_2)\, d(x_2, \hat{x}_2)]$
$= \text{Min}\, [(0.75 * 0) + (0.25 * \infty)], [(0.75 * 4) + (0.25 * 0)]$
$= \text{Min}\, \infty, 3$
$= 3$

d) $E\,[d(X, \hat{X})] = \sum\sum p(x_k)\, p(\hat{x}_j \mid x_k)\, d(x_k, \hat{x}_j)$
$= 0.75\, [(\alpha * 0) + ((1 - \alpha) * 4)] + 0.25\, [(0 * \infty) + (1 * 0)]$
$= 0.75\, [(1 - \alpha) * 4)]$
$= 3\,(1 - \alpha)$

1/1

e) $I(X ; \hat{X}) = \sum\sum p(x_k)\, p(\hat{x}_j \mid x_k) \log_2 [p(\hat{x}_j \mid x_k) / \sum p(x_l)\, p(\hat{x}_j \mid x_l)]$

$= 0.75\, [\alpha \log_2 [\alpha / [(0.75 * \alpha) + (0.25 * 0)]] + (1 - \alpha) \log 2\, [(1 - \alpha) / [(0.75 * (1 - \alpha)) + (0.25 * 1)]]]$
$+ 0.25\, [\log_2 [1 / [(0.75 * (1 - \alpha) + (0.25 * 1)]]]$

1/1

f)

$R(D) = \min \{(0.75)\, [(\alpha) \log_2 (\alpha / 0.75\alpha) + (1-\alpha) \log_2 ((1-\alpha) / (0.75\,(1-\alpha) + 0.25)]$
$+ (0.25)\, [(1) \log_2 (1 / (0.75\,(1-\alpha) + 0.25)]\}$

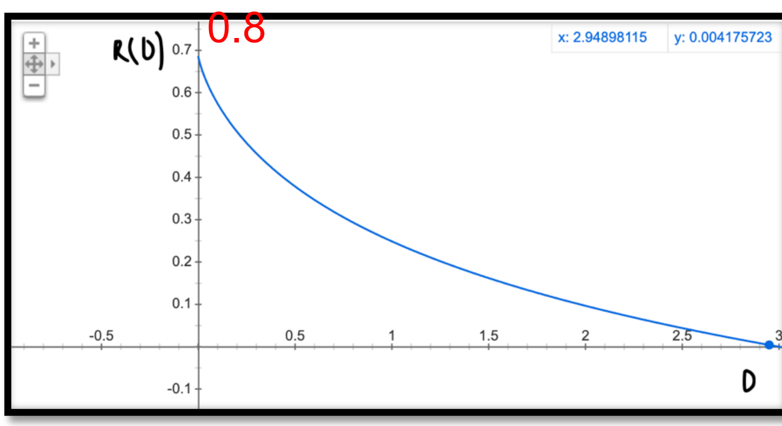
$= (0.75)\, [(\alpha) \log_2 (4/3) + (1-\alpha) \log_2 ((1-\alpha) / (1 - 0.75\alpha)]$
$+ (0.25)\, [\log_2 (1 / (1 - 0.75\alpha)]\}$

g) Draw function R (D) 0.5/1

$(d_{min} - 2) \le D \le (d_{max} + 2)$
$-2 \le D \le 5$

With $0 \le \alpha < 1$
$D = 3\,(1 - \alpha)$
$\alpha = -(D/3) + 1$

$R(D) = (0.75)\, ((-(D/3) + 1) \log_2 (1/ 0.75) + (1-(-(D/3) + 1)) \log_2 ((1-(-(D/3) + 1)) / (1 - 0.75(-(D/3) + 1)) + (0.25)\, ((\log_2 (1 / (1 - 0.75(-(D/3) + 1))))$



- **Problème 4: Algorithme de Blahut-Arimoto** 3/3

Problème 7.2 : Une source d'information sans mémoire X génère des symboles quaternaires avec la distribution : $p(x_1) = \frac{3}{7}$, $p(x_2) = \frac{1}{7}$, $p(x_3) = \frac{1}{7}$ et $p(x_4) = \frac{2}{7}$. Un encodeur de source avec pertes quaternaire de mesure de distorsion d est utilisé pour réduire le débit d'information à transmettre. Écrivez un programme pour effectuer l'algorithme de Blahut-Arimoto pour le calcul de la fonction de débit-distorsion. Donnez le listing de votre programme.

$$\mathbf{d} = \begin{bmatrix} d(x_1,\hat{x}_1) & d(x_2,\hat{x}_1) & d(x_3,\hat{x}_1) & d(x_4,\hat{x}_1) \\ d(x_1,\hat{x}_2) & d(x_2,\hat{x}_2) & d(x_3,\hat{x}_2) & d(x_4,\hat{x}_2) \\ d(x_1,\hat{x}_3) & d(x_2,\hat{x}_3) & d(x_3,\hat{x}_3) & d(x_4,\hat{x}_3) \\ d(x_1,\hat{x}_4) & d(x_2,\hat{x}_4) & d(x_3,\hat{x}_4) & d(x_4,\hat{x}_4) \end{bmatrix} = \begin{bmatrix} 0.7 & 2.3 & 3.2 & 5.4 \\ 6.3 & 1.4 & 3.9 & 4.3 \\ 2.5 & 4.3 & 0.7 & 9.2 \\ 2.7 & 4.4 & 8.5 & 1.1 \end{bmatrix}$$

a) Donnez l'entropie de la source, $H(X)$.

b) Déterminez les valeurs de distorsion minimale d_{min} et de distorsion maximale d_{max}.

c) Calculez et faites tracer la fonction de débit-distorsion $R(D)$.

a) $H(X) = -\Sigma\, p(x_k) \log_2 p(x_k)$

$= -[(3/7) \log_4 (3/7) + (1/7) \log_4 (1/7) + (1/7) \log_4 (1/7) + (2/7) \log_4 (2/7)]$

$= -[(-0.26194) + (-0.2005) + (-0.2005) + (-0.25819)]$

$= 0.9211$ 1/1 (attention : indiquez les unités (log_4) ...)

b) 1/1

$d_{min} = \sum p(x_k)\, d(x_k, \hat{x}_{j\,min})$

$= p(x_1)\, d(x_1, \hat{x}_1) + p(x_2)\, d(x_2, \hat{x}_2) + p(x_3)\, d(x_3, \hat{x}_3) + p(x_4)\, d(x_4, \hat{x}_4)$

$= (3/7)(0.7) + (1/7)(1.4) + (1/7)(0.7) + (2/7)(1.1)$

$= 0.9143$ (or 32/35)

$d_{max} = {}_{(\min j = 1,2\ldots)} \sum p(x_k)\, d(x_k, x_j)$

$= \text{Min}\ [p(x_1)\, d(x_1, \hat{x}_1) + p(x_2)\, d(x_2, \hat{x}_1) + p(x_3)\, d(x_3, \hat{x}_1) + p(x_4)\, d(x_4, \hat{x}_1)],$

$[p(x_1)\, d(x_1, \hat{x}_2) + p(x_2)\, d(x_2, \hat{x}_2) + p(x_3)\, d(x_3, \hat{x}_2) + p(x_4)\, d(x_4, \hat{x}_2)],$

$[p(x_1)\, d(x_1, \hat{x}_3) + p(x_2)\, d(x_2, \hat{x}_3) + p(x_3)\, d(x_3, \hat{x}_3) + p(x_4)\, d(x_4, \hat{x}_3)],$

$[p(x_1)\, d(x_1, \hat{x}_4) + p(x_2)\, d(x_2, \hat{x}_4) + p(x_3)\, d(x_3, \hat{x}_4) + p(x_4)\, d(x_4, \hat{x}_4)].$

$= \text{Min}\ [(3/7)(0.7) + (1/7)(2.3) + (1/7)(3.2) + (2/7)(5.4)],$

$[(3/7)(6.3) + (1/7)(1.4) + (1/7)(3.9) + (2/7)(4.3)],$

$[(3/7)(2.5) + (1/7)(4.3) + (1/7)(0.7) + (2/7)(9.2)],$

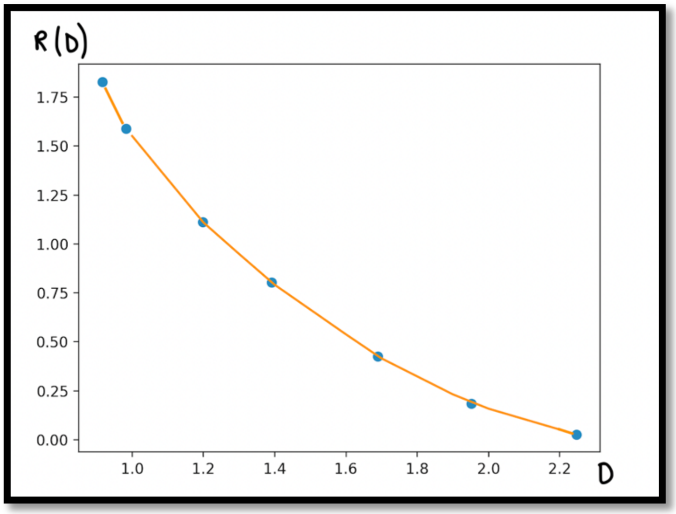
$[(3/7)(2.7) + (1/7)(4.4) + (1/7)(8.5) + (2/7)(1.1)].$

$= \text{Min}\ 2.6286;\ 4.6857;\ 4.4143;\ 3.3143$

$= 2.6286$

1/1

c)



```
b_rd.py > ...
import numpy as np

# Vecteur Probability source
p_x = np.array([[3/7, 1/7, 1/7, 2/7]])

# Distorsion matrix (k x j)
dist_mat = np.array([[0.7, 2.3, 3.2, 5.4],
[6.3, 1.4, 3.9, 4.3],
[2.5, 4.3, 0.7, 9.2],
[2.7, 4.4, 8.5, 1.1]])

# Pente
beta = -1.25



def BlahutArimato(dist_mat, p_x, beta ,max_it=60,eps=1e-3) :

    l,l_hat = dist_mat.shape
    p_cond = np.tile(p_x, (l_hat,1)).T      # Equiprobable Distribution (start)

    p_x = p_x / np.sum(p_x)
    p_cond /= np.sum(p_cond,1,keepdims=True)

    it = 0
    Du_prev = 0
    Du = 2*eps

    while it < max_it and np.abs(Du-Du_prev)> eps :

        it+=1
        Du_prev = Du
        p_hat = np.matmul(p_x,p_cond)

        p_cond = np.exp(beta * dist_mat) * p_hat
        p_cond /= np.expand_dims(np.sum(p_cond,1),1)

        Iu = np.matmul(p_x,p_cond*np.log(p_cond / np.expand_dims(p_hat,0))).sum()
        Du = np.matmul(p_x,(p_cond * dist_mat)).sum()

    return Iu/np.log(2), Du



R, D = BlahutArimato(dist_mat, p_x, beta)

print(R, D)
```

Beta = -1	**R = 0.8029957018183123**	**D = 1.391039709151122**
Beta = -1/2	**R = 0.18502164248885006**	**D = 1.951300637175144**
Beta = -1/4	**R = 0.027494506083299786**	**D = 2.246557238373074**
Beta = -2	**R = 1.5902665378150278**	**D = 0.9823774929939508**
Beta = -4	**R = 1.8281955455768917**	**D = 0.9164211121910948**
Beta = -0.75	**R = 0.42502890729143944**	**D = 1.689644467942369**
Beta = -1.25	**R = 1.1120229384212583**	**D = 1.1986049841396558**

- **<u>Problème 5: Région de Capacité de Canaux Discrets</u>** 2/2

Problème 8.1 : Déterminez et tracez la région de capacité des canaux suivants :

a) Canal à accès multiple avec deux sources binaires $X_1 \in \{0,1\}$ et $X_2 \in \{0,1\}$ et

$$Y = X_1 + X_2 \bmod 2.$$

b) Canal à accès multiple avec deux sources $X_1 \in \{0,1,2\}$ et $X_2 \in \{0,1\}$ et

$$Y = X_1 + X_2 \bmod 3.$$

c) Canal à accès multiple multiplicatif avec $X_1 \in \{0,1\}$ et $X_2 \in \{1,2,3\}$ et

$$Y = X_1 \cdot X_2.$$

Théorème *(capacité des canaux à accès multiple)* : La région de capacité d'un canal à accès multiple sans mémoire est définie par la région convexe fermée des débits R_1 et R_2 tel que :

$$\begin{aligned} R_1 &\leq I(X_1;Y|X_2), \\ R_2 &\leq I(X_2;Y|X_1), \qquad \text{et} \\ R_1 + R_2 &\leq I(X_1,X_2;Y) \end{aligned}$$

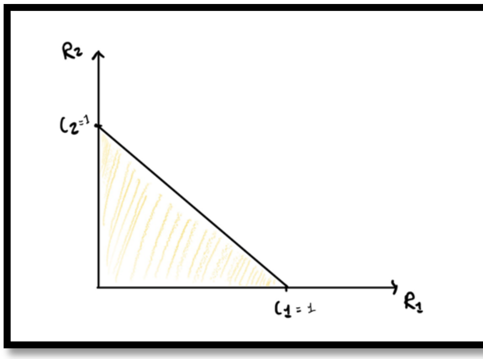
pour une distribution produit $\{p_1(x_{1,k}), p_2(x_{2,j})\}$ de la paire de sources (X_1, X_2).

a) 1/1

$X_1 = \{0,1\}$
$X_2 = \{0,1\}$

If Y = 0 or Y = 2, X1 et X2 are known
H (Y) ≤ 1
Max rate possible = 1 bit
Rate 1 bit from X_1 if X2 = 0
Rate 1 bit from X_2 if X1 = 0

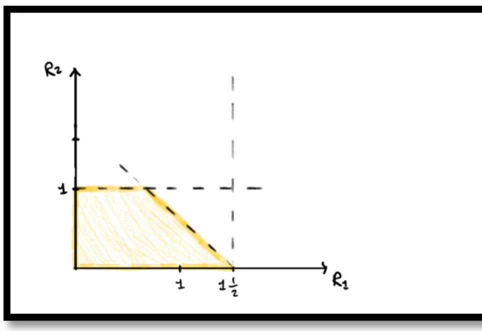
R1 ≤ I (X1; Y |X2) = H(X1)
R2 ≤ I (X2; Y |X1) = H (X2)
R1 + R2 ≤ I (X1, X2; Y) = H (Y)
...
R1 ≤ 1
R2 ≤ 1
R1 + R2 ≤ 1



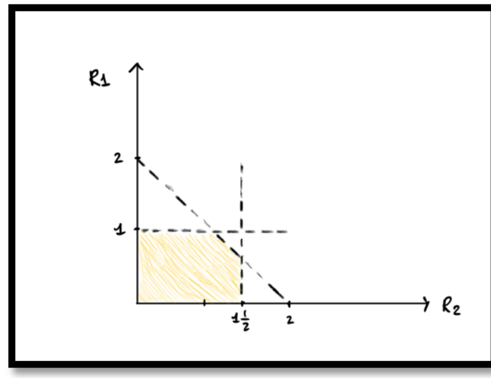
b) 1/1

If Y = 0 or Y = 3, X1 et X2 are known
Only Y = 1 & Y = 2 carry uncertainty

R1 ≤ I (X1; Y |X2) = H (X1)
R2 ≤ I (X2; Y |X1) = H (X2)
R1 + R2 ≤ I (X1, X2; Y) = H (Y)
…
R1 < 1.5 log2(3)
R2 < 1
R1 + R2 < 1.5 log2(3)



c) vu



If X1 = 0, then Y = 0 and X2 carries no information
If X1 = 1, then Y = X2

$X_1 = \{0,1\}$
$X_2 = \{1,2,3\}$

$Y = \{0,1,2,3\}$

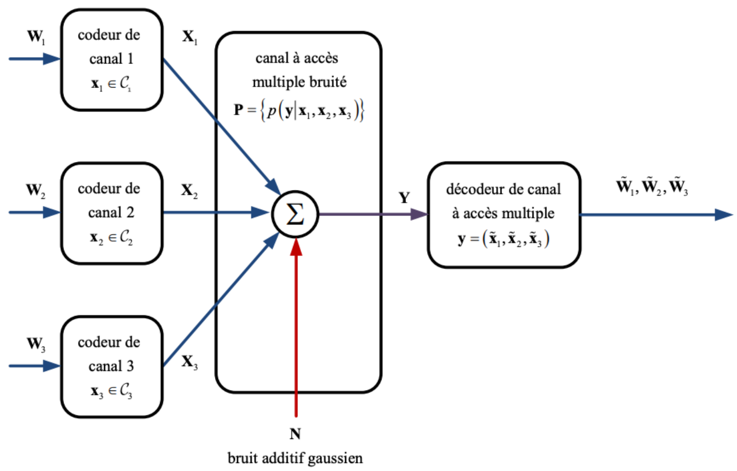
- Constraints when we play with probabilities of $X_1 = 1$:

$R1 \leq I\ (X1; Y\ |X2) = H\ (X_1)$
$R2 \leq I\ (X2; Y\ |X1) = H\ (X_2) * P\ (X_1{=}1)$
$R1 + R2 \leq I\ (X1, X2; Y) = H\ (Y) = (H\ (P\ (X_1{=}1)) + (H\ (X_2) * (P\ (X_1{=}1))$
...
$R1 \leq 1$ *(maximum when p=0.5)*
$R2 \leq 1.5$ *(maximum when p=1)*
$R1 + R2 \leq 2$ *(maximum when p=0.75) (0.81 + 1.19 = 2)*

3/3

- **<u>Problème 6: Région Capacité Canal à Accès Multiple Gaussien</u>**

Problème 8.3 : Soit le canal à accès multiple à bruit additif gaussien (AWGN) avec 3 utilisateurs illustré ci-dessous. Les trois sources transmettent les signaux X_1, X_2 et X_3 avec des puissances $P_1 = 10$, $P_2 = 8$ et $P_3 = 6$, respectivement. Le canal est affecté par du bruit blanc additif gaussien N de moyenne μ_N nulle et de variance $\sigma_N^2 = 5$.



a) Définissez la région de capacité de ce canal à accès multiple en donnant l'expression des débits ci-dessous en fonction des rapports signal-à-bruit, R_1, R_2, R_3, $R_1 + R_2$, $R_1 + R_3$, $R_2 + R_3$ et $R_1 + R_2 + R_3$

b) Calculez les valeurs numériques de chacun de ces débits.

c) À l'aide d'un logiciel, faites un graphique de la région de capacité de ce canal à accès multiple gaussien à 3 utilisateurs.

a) & b) 2/2

$$\boxed{\sum_{i \in \mathcal{S}} R_i \leq \frac{1}{2} \log_2 \left[1 + \frac{\sum_{i \in \mathcal{S}} P_i}{\sigma_Z^2}\right]}$$

$R_1 \leq 0.5 \log_2 [1 + (P1 / \sigma^2 z)]$
$\leq 0.5 \log_2 [1 + (10 / 5)]$
≤ 0.7925

$R_2 \leq 0.5 \log_2 [1 + (P2 / \sigma^2 z)]$
$\leq 0.5 \log_2 [1 + (8 / 5)]$
≤ 0.6893

$R_3 \leq 0.5 \log_2 [1 + (P3 / \sigma^2 z)]$
$\leq 0.5 \log_2 [1 + (6 / 5)]$
≤ 0.5688

$R_1 + R_2 \leq 0.5 \log_2 [1 + (P1 + P2 / \sigma^2 z)]$
$\leq 0.5 \log_2 [1 + ((10 + 8) / 5)]$
≤ 1.101

$R_1 + R_3 \leq 0.5 \log_2 [1 + (P1 + P3 / \sigma^2 z)]$
$\leq 0.5 \log_2 [1 + ((10 + 6) / 5)]$
≤ 1.0352

$R_2 + R_3 \leq 0.5 \log_2 [1 + (P2 + P3 / \sigma^2 z)]$
$\leq 0.5 \log_2 [1 + ((8 + 6) / 5)]$
≤ 0.963

$R_1 + R_2 + R_3 \leq 0.5 \log_2 [1 + (P1 + P2 + P3 / \sigma^2 z)]$
$\leq 0.5 \log_2 [1 + ((10 + 8 + 6) / 5)]$
≤ 1.2680

c) (Dessin)

- Difficultés avec logiciel
- Passé plus de 2h à essayer

O.K. 1/1