

20 →

16,67/20

IUT de Paris Descartes, Dpt INFO

DST de Modélisation Mathématique - S3 - Janv.2019

Durée : 2H - Sans Documents, calculatrice fournie

NOM-Prénom : LAPOSTOLLE Arsène ... Groupe : 205

Préambule :

- On prendra soin de répondre aux questions *dans l'espace imparti à cet effet au sein même du sujet* : si cet espace est minime, une réponse brève suffira.
- On prendra soin de répondre à ces questions *sans utiliser de crayon gris* : stylo ou bic à encre de couleur noire ou bleue SVP.
- La plupart des questions posées font seulement appel à vos capacités de réflexion et à votre connaissance du cours, sans nécessité de poser un calcul numérique précis. Pour les quelques questions nécessitant de poser un calcul, on donne les valeurs ci-dessous

$\log(0,05) \approx -4,32$	$\log(0,1) \approx -3,32$	$\log(0,15) \approx -2,74$	$\log(0,2) \approx -2,32$	$\log(0,25) = -2$
$\log(0,3) \approx -1,74$	$\log(0,35) \approx -1,51$	$\log(0,4) \approx -1,32$	$\log(0,45) \approx -1,15$	$\log(0,5) = -1$
$\log(0,55) \approx -0,86$	$\log(0,6) \approx -0,74$	$\log(0,65) \approx -0,62$	$\log(0,7) \approx -0,51$	$\log(0,75) \approx -0,42$
$\log(0,8) \approx -0,32$	$\log(0,85) \approx -0,23$	$\log(0,9) \approx -0,15$	$\log(0,95) \approx -0,07$	$\log(1) = 0$
$\log(2) = 1$	$\log(3) \approx 1,58$	$\log(4) = 2$	$\log(5) \approx 2,32$	$\log(6) \approx 2,58$
$\log(7) \approx 2,81$	$\log(8) = 3$	$\log(9) \approx 3,17$	$\log(10) \approx 3,32$	$\log(11) \approx 3,46$

ainsi que

$\log(0,01) \approx -6.6438562$, $\log(0,02) \approx -5.6438562$, $\log(0,98) \approx -0.0291463$, $\log(0,99) \approx -0.0144996$,

tous ces logarithmes étant bien entendu pris *en base 2*.

A. Questions sur l'Entropie et l'Information de Shannon :

1. Comment définir *intuitivement* (en français) l'Entropie $H(X)$ d'une variable aléatoire X à valeurs dans un ensemble fini $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$?

0,5 - L'entropie de X correspond à l'incertitude qui porte sur la valeur v_i que va prendre X .

2. Comment définir cette Entropie $H(X)$ *mathématiquement*, en ayant pris soin d'introduire les probabilités $p_1, p_2, \dots, p_n$ d'observer $X = v_1, X = v_2, \dots$ ou $X = v_n$?

0,5 - Soient $p_1 \dots p_n$ les probabilités respectives $X = v_1 \dots X = v_n$, on a :

-
$$H(X) = -\left(\sum_{i=1}^n p_i \log_2(p_i)\right)$$

3. En supposant toujours que $X(\Omega) = \{v_1, v_2, \dots, v_n\}$, quelle est la valeur maximale envisageable pour l'entropie $H(X)$? Et à quelle condition sur X verra-t-on $H(X)$ prendre cette valeur maximale ?

- La valeur maximale de $H(X)$ est $\log_2(n)$

1 - $H(X)$ atteint sa valeur maximale lorsque X suit la loi uniforme, c'est à dire que les valeurs $v_1 \dots v_n$ sont équiprobables. Elle augmente d'autant plus que n est grand

4. Comment définit-on mathématiquement l'Information $I(X;Y)$ portée par une variable aléatoire Y sur la var. al. X ?

$I(X;Y) = H(X) - H(X|Y)$

ou bien, d'après la règle d'enchaînement:

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

5. En quoi peut-on affirmer que cette notion d'Information permet de détecter précisément l'éventuelle indépendance des variables X et Y ?

$I(X;Y)$ correspond à l'incertitude sur X qui est levée par la connaissance de Y . Si I est nulle c'est que Y ne porte pas d'information sur X et par conséquent X et Y sont indépendantes.

6. Soient $p, p' \in]0; 1[$ tels que : $0 < p < \frac{1}{2}$ et $p' = (1 - p)$, puis U, V, W des variables de Bernoulli t. q. $U \hookrightarrow B(p)$, $V \hookrightarrow B(p')$ et $W \hookrightarrow B(\frac{1}{2})$. Parmi les affirmations ci-dessous, cocher celles qui vous paraissent correctes :

- ☒ $H(U) < H(V) < H(W)$ ☐ $H(U) = H(V) < H(W)$ ☐ $H(W) < H(V) = H(U)$
☐ $H(W) < 1$ ☒ $H(W) = 1$ ☐ $H(W) > 1$

7. Soient $p \in]0; \frac{1}{2}[$ et $n, n' \in \mathbb{N}^*$ tels que $n' > n > 1$, puis U, V, W des variables Binomiales telles que $U \hookrightarrow B(n; p)$, $V \hookrightarrow B(n'; p)$ et $W \hookrightarrow B(n'; \frac{1}{2})$. Parmi les affirmations ci-dessous, cocher celles qui vous paraissent justes :

- ☒ $H(U) < H(V) < H(W)$ ☐ $H(W) < H(V) < H(U)$ ☐ $H(U) = H(V) = H(W)$
☐ $H(W) = \log(n' + 1)$ ☒ $H(W) < \log(n' + 1)$ ☐ $H(W) > \log(n' + 1)$

B. Algorithme de Huffman et 1er Thm de Shannon :

Dans cet exercice, on dispose de deux dés tétraédriques (4 faces chacun) et équilibrés, ces dés sont lancés au hasard et X_1 désigne le résultat produit par le premier dé, X_2 le résultat produit par le 2ème. On introduit aussi la variable $Y = (X_1 + X_2)$.

1. Parmi les affirmations ci-dessous, cocher celles qui vous paraissent correctes (aucun calcul n'est ici nécessaire, il vous est surtout demandé de réfléchir soigneusement) :

- 825
- | | | |
|---|--|--|
| ☒ $H(X_1, X_2) < H(X_1) + H(X_2)$ | ☐ $H(X_1, X_2) = H(X_1) + H(X_2)$ | ☐ $H(X_1, X_2) > H(X_1) + H(X_2)$ |
| ☐ $H(X_1) = \log 4$ | ☐ $H(X_2) < 2$ | ☐ $H(X_2) = 2$ |
| ☒ $H(X_2, X_1) = H(X_1, X_2)$ | ☐ $H(X_2, X_1) = H(X_2)$ | ☐ $H(X_2 X_1) = H(X_2)$ |
| ☒ $I(X_2, X_1) = I(X_1, X_2)$ | ☒ $I(X_1, X_2) = 0$ | ☐ $I(X_1, X_2) > 0$ |
| ☐ $H(Y) = H(X_1, X_2)$ | ☒ $H(Y) < H(X_1, X_2)$ | ☐ $H(X_1, Y) = H(X_1, X_2)$ |
| ☐ $I(X_1, Y) = I(X_1, X_2)$ | ☒ $I(X_1, Y) > 0$ | ☐ $I(X_1, Y) = 0$ |

2. Manifestement, $Y = (X_1 + X_2)$ a pour valeurs possibles tous les entiers k tels que $2 \leq k \leq 8$. Justifier brièvement le fait que la loi de Y peut être décrite à travers le vecteur de probabilité

$$p = (p_2, p_3, p_4, p_5, p_6, p_7, p_8) = \left(\frac{1}{16}, \frac{2}{16}, \frac{3}{16}, \frac{4}{16}, \frac{3}{16}, \frac{2}{16}, \frac{1}{16}\right) = \left(\frac{1}{16}, \frac{1}{8}, \frac{3}{16}, \frac{1}{4}, \frac{3}{16}, \frac{1}{8}, \frac{1}{16}\right).$$

On a $2^4 = 16$ combinaisons possibles. On a une combinaison qui permet de tomber sur 2 (1 et 1), 2 combinaisons qui permettent de tomber sur 3 (1 et 2, 2 et 1), 3 qui permettent de tomber sur 4 (2 et 2, 3 et 1, 1 et 3) et 4 combinaisons qui permettent de tomber sur 5 (4 et 1, 1 et 4, 3 et 2, 2 et 3). Il en va de même pour les valeurs respectivement opposées à 5 : 3 combinaisons pour 6, 2 pour 7 et 1 pour 8.

3. Donner une valeur approchée (aussi précise que possible) de $H(Y)$.

1

$$H(Y) = \left(\frac{1}{16}, \frac{1}{8}, \frac{3}{16}, \frac{1}{4}, \frac{3}{16}, \frac{1}{8}, \frac{1}{16}\right)$$

$$= -\left[\frac{1}{16}(\log 1 - \log 16) + \frac{1}{8}(\log 1 - \log 8) + \frac{3}{16}(\log 3 - \log 16) + \frac{1}{4}(\log 1 - \log 4)\right] = 2,6575$$

1

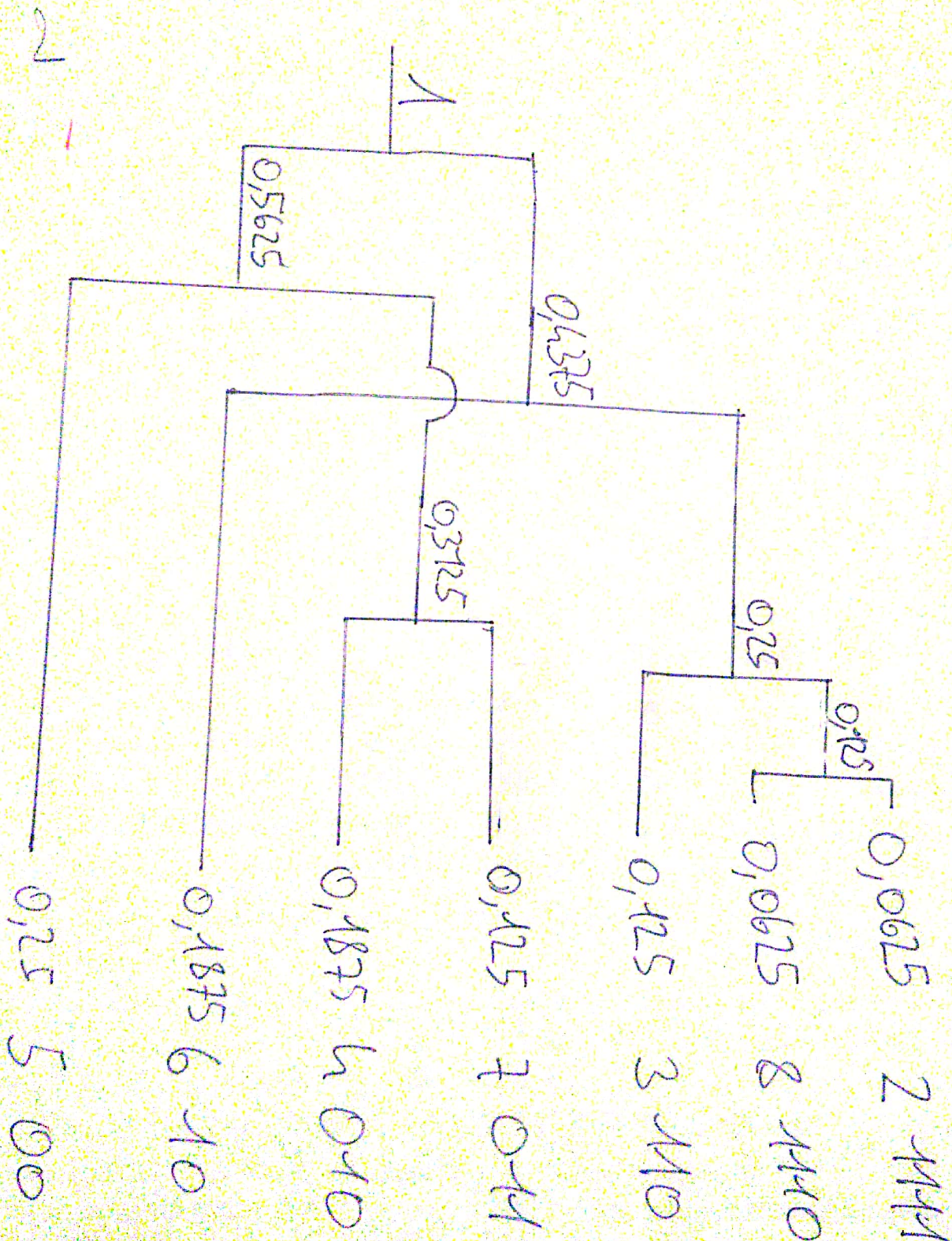
4. Pourquoi a-t-on $H(Y) < \log 7$?

Car les valeurs de X ne ne sont pas équiprobables, ainsi $H(Y)$ n'atteint pas son maximum

1

5. Produire un arbre permettant de visualiser graphiquement le déroulement de l'Algorithme de Huffman appliqué à un codage binaire des valeurs possibles de Y (Dessin soigné SVP!).

(Indication : on aura avantage à permuter l'ordre naturel de ces valeurs en posant $v_1 = 2$, $v_2 = 8$, $v_3 = 3$, $v_4 = 7$, $v_5 = 4$, $v_6 = 6$, $v_7 = 5$).



Probabilities code

conventions : Priorité au plus élevé, montée : 1 descend : 0

6. A partir de cet arbre, proposer un code binaire uniquement déchiffirable pour les valeurs $v_1, v_2, v_3, v_4, v_5, v_6, v_7$ de Y .

$$\begin{aligned}\varphi(2) &= 1111, \varphi(8) = 1110, \varphi(3) = 110 \\ \varphi(7) &= 011, \varphi(4) = 010, \\ \varphi(6) &= 10, \varphi(5) = 00\end{aligned}$$

7. Pourquoi dit-on de ce code-source obtenu par l'Alg. de Huffman qu'il constitue un code préfixe?

On dit qu'il constitue un code préfixe car aucun mot de code n'est le début d'un autre.

8. Calculer le nombre moyen de bits utilisés par valeur-source dans ce code. Soit $\ell(\varphi)$ le nb moyen de bits utilisés:

$$\begin{aligned}\ell(\varphi) &= 0,0625 \times 6 + 0,0625 \times 4 + 0,125 \times 3 \\ &\quad + 0,125 \times 3 + 0,1875 \times 3 + 0,1875 \times 2 \\ &\quad + 0,25 \times 2 \\ &= 2,6875\end{aligned}$$

9. En quoi peut-on affirmer que ce code-source obtenu par l'Alg. de Huffman est optimal?

Ce code-source est optimal car c'est un code uniquement déchiffirable qui utilise un nombre moyen de bit par mot source minimal.

10. Au vu des résultats qui précèdent, en quoi le 1er Thm de Shannon est-il confirmé?

D'après le 1^{er} THM de Shannon:

$$H(X) < \ell(\varphi) < H(X) + 1$$

ou on a bien:

$$2,6575 < 2,6875 < 3,6575$$

Donc le 1^{er} THM de Shannon est vérifié.

12. Sur le 2ème Thm de Shannon :

Une source discrète et sans mémoire d'entropie 50 bits/mot-source est connectée à un Canal Binaire Symétrique de paramètre d'erreur $\epsilon = 0,01$ qui peut transmettre (physiquement) jusqu'à 10^6 bits/seconde au Récepteur.

1. En utilisant les valeurs approchées données dans le préambule, vérifier que la Capacité de ce Canal Binaire Symétrique s'élève à

$$C \approx 0,9192069$$

Pour un canal binaire symétrique,

$$C(\epsilon) = 1 + \epsilon \log_2(\epsilon) + (1-\epsilon) \log_2(1-\epsilon)$$

$$C(\epsilon) = 1 + 0,01 \log_2(0,01) + 0,99 \log_2(0,99)$$

On a bien : $C(\epsilon) \approx 0,9192069$

2. Si la transmission doit être effectuée de manière fiable, combien de mots-source peuvent être émis par seconde ?

Soient D_v le nombre de mots-source émis de manière fiable, D_p le débit physique du canal et W la source. On a :

$$D_v = \frac{D_p \times C}{H(W)}$$

$$= \frac{10^6 \times 0,9192069}{50}$$

$$= 18\,384,12$$

On peut donc transmettre 18 384,12 mots-source de manière fiable chaque seconde.