Corrigés des exercices

2 Communications numériques

Exercice 2.1 $D = 1/T_b$ où T_b est l'intervalle de temps entre les émissions de deux bits consécutifs. $R = 1/T$ où T désigne la vitesse de modulation. Si M désigne la taille de l'alphabet de modulation, on a $R = D_b/\log_2(M)$.

Exercice 2.2 $x(t) = \sum_k a_k h(t - kT)$ où $a_k \in \{-3, -1, 1, 3\}$ et $h(t) = \text{rect}_T(t)$.

Exercice 2.3 De deux manières en modifiant la forme de $g(t)$ ou en introduisant des corrélations entre les symboles a_k.

Exercice 2.4 L'interférence entre symboles (aux instants d'échantillonnage) est la présence dans l'observation en sortie de l'échantillonnage (en sortie du filtre adapté) de termes associés à *plus* d'un seul symbole. Pour s'en affranchir on impose le critère de Nyquist : si $p(t)$ désigne la réponse impulsionnelle totale de puis l'entrée du filtre d'émission jusqu'à la sortie du filtre de réception, on doit avoir :

$$\exists \tau \text{ tel que } \forall k \neq 0 \; p(kT + \tau) = 0$$

Exercice 2.5 Si $p(t)$ désigne la réponse impulsionnelle totale depuis l'entrée du filtre d'émission jusqu'à la sortie du filtre de réception, on doit avoir :

$$\exists \tau \text{ tel que } \forall k \neq 0 \; p(kT + \tau) = 0$$

La condition est $R = 1/T > 2/B$.

Exercice 2.6 (Fonctions vérifiant le critère de Nyquist)

1. $r(t) = C \sin(\pi t/T)/\pi t$ et $y(t) = x(t)r(t)$. Comme $r(t)$ s'annule en $t = kT$ pour tout $k \neq 0$, il en va de même pour $y(t)$ qui vérifie donc le critère de Nyquist.
2. Dans ce cas :
$$X(f) = e^{-2j\pi f/4b}\mathbb{1}_{(-b,b)}(f) + e^{+2j\pi f/4b}\mathbb{1}_{(-b,b)}(f)$$
et donc :
$$\begin{aligned} x(t) &= \frac{\sin(2\pi b(t - 1/4b))}{\pi(t - 1/4b)} + \frac{\sin(2\pi b(t + 1/4b))}{\pi(t + 1/4b)} \\ &= \frac{1}{2\pi b}\frac{\cos(2\pi bt)}{1/16b^2 - t^2} \end{aligned}$$
3. $y(t) = Cx(t)\sin(\pi t/T)/\pi t$. Avec $b = \alpha/2T$ et $C = \pi T/8b$, on obtient :
$$y(t) = \frac{\sin(\pi t/T)}{\pi t/T}\frac{\cos(\pi\alpha t/T)}{1 - 4\alpha^2 t^2/T^2}$$
Un calcul long (mais sans difficulté) de $Y(f) = X(f) \star R(f)$ donne pour $Y(f)$ l'expression (**??**) de la famille $C_\alpha(f)$ des fonctions en cosinus surélevé.

Remarque : cet exercice montre une façon simple d'engendrer des fonctions à bande limitée qui satisfont le critère de Nyquist.

Exercice 2.7 Le diagramme de l'oeil est la superposition des trajectoires du processus aléatoire $y(t)$ en sortie du filtre de réception (associée aux trajectoires de la suite aléatoire a_k des symboles). Cette superposition est observée sur 2 fois la durée symbole T.

En absence de bruit et dans la cas où l'IES est nulle, le diagramme de l'œil présente des trajectoires qui concourent en des instants espacés de T et en des ordonnées correspondant aux symboles de modulation.

Exercice 2.8 voir cours. Plus le "rolloff" est voisin de 1, plus le diagramme de l'oeil s'ouvre horizontalement ?

Exercice 2.9 La distorsion maximale est le maximum d'interférences destructives :

$$D_{\max} = \frac{(M-1)\sum_k |p_k| - |p_{\max}|}{|p_{\max}|}$$

Si l'IES est nulle, $D_{\max} = 0$ pour un certain retard τ. Pour les impulsions en cosinus-surélevé, qui vérifie le critère de Nyquist, la distorsion maximale est nulle. Si on s'écarte des instants d'IES nulle, la distorsion maximale est une fonction décroissante du “rolloff”. Si le “rolloff” est nul, la distorsion maximale est infinie.

Exercice 2.10 Non corrigé.

Exercice 2.11 voir cours.

Exercice 2.12 Dans le cas d'un grand rapport signal sur bruit et d'un codage de Gray, on a :

$$\text{TEEB} \approx \frac{P_s}{\log_2(M)}$$

Exercice 2.13 L'efficacité spectrale est le rapport D_b/B. Elle se mesure en bits/s/Hz. Pour une MIA avec une impulsion en cosinus sur-élevé et un “rolloff”, l'expression de l'efficacité spectrale :

$$\eta = \frac{2\log_2(M)}{1+\alpha}$$

Exercice 2.14 $y_k = a_k p(0) + w_k$ où w_k est une variable aléatoire gaussienne, centrée de variance :

$$\mathbb{E}\left\{w_k^2\right\} = \mathbb{E}\left\{w_a^2(kT)\right\} = \int S_W(f)df = \frac{N_0}{2}\int |H_r(f)|^2 df$$

Mais $p(t) = \int h(u)h_r(t-u)du$. Comme $h_r(u) = h^*(-u)$, on a :

$$p(0) = \int h(u)h^*(u)du = \int |H_r(f)|^2 df$$

et donc $\mathbb{E}\left\{w_k^2\right\} = N_0 p(0)/2$.

Exercice 2.15 Pour annuler l'IES, on prend une fonction qui satisfait le critère de Nyquist, par exemple le cosinus surélevé. Comme il faut répartir entre émission et réception on a :

- à l'émission : $H(f) = \sqrt{C_\alpha(f)}$. Comme on part d'une impulsion NRZ dont la transformée de Fourier est $\sin(\pi fT)/\pi f$, il faut compenser à l'émission par un filtre :

$$H_E(f) = \sqrt{C_\alpha(f)} \times \frac{\pi f}{\sin(\pi fT)}$$

- à la réception : $H_R(f) = \sqrt{C_\alpha(f)}$.

Exercice 2.16 Pour transmettre un débit binaire de 2 Mbits/s, on utilise une modulation MIA à 4 états. La probabilité d'erreurs par bits est de 10^{-5}. On désire, en conservant une modulation de type MIA, doubler le débit binaire, sur ce même canal, sans toucher à la probabilité d'erreurs par bits.

1. Pour doubler le débit binaire $D_b = R\log_2(M)$ sans modifier la bande et donc R, il faut multiplier par $2\log_2(M)$ et donc passer de M à M^2 soit $M' = 16$.
2. D'après le tableau il faut augmenter E_b/N_0 de 13.6 à 22.9 dB soit, à bruit constant, E_b de 9.3 dB. Comme :

$$P_s = E_b D_b$$

on a en dB :

$$P'_{sdB} - P_{sdB} = E'_{bdB} - E_{bdB} + 10\log_{10}(\frac{D'_b}{D_b}) = 9.3 + 3$$

Exercice 2.17 (Transmission en bande de base) On considère une source binaire, indépendante, équirépartie, dont le débit est de 10000 bits/s. Le signal émis a pour expression :

$$s(t) = \sum_k a_k h(t - kT)$$

où les a_k sont une suite de symboles à valeurs dans l'alphabet $\{-3, -1, +1, +3\}$. T désigne l'intervalle de temps entre deux symboles consécutifs. $h(t)$ est un signal réel, d'énergie finie, dont la transformée de Fourier $H(f)$ est nulle à l'extérieur de la bande de fréquence $(-B, B)$. On donne $W = 3000\text{Hz}$. Le bruit est une réalisation d'un processus blanc, gaussien, centré, de densité spectrale de puissance $N_0/2$. On note $z(t) = s(t) + b(t)$ le signal reçu.

1. $R = 1/T = D_b/\log_2(M) = 5000$ bauds.
2. Codage de Gray : $00 : -3$, $01 : -1$, $11 : +1$ et $10 : +3$.
3. Le gain complexe du filtre de réception qui minimise la probabilité d'erreur est $H_r(f) = H^*(f)$. On note $P(f)$ la cascade constituée de le filtre d'émission $H(f)$ et du filtre de réception et $p(t)$ sa transformée de Fourier inverse.
4. $$\exists \tau \text{ tel que } \forall k \neq 0 \; p(kT + \tau) = 0$$
5. $W = 3000 = R(1 + \alpha)/2 = 2500(1 + \alpha)$ et donc $\alpha = 0.2$.
6. Comme $E_s = \frac{1}{4}(2 \times 9 \int |h(t)|^2 dt + 2 \times \int |h(t)|^2 dt) = 5p(0)$.
7. On a $E_b = E_s/2$ et donc $p(0) = 2E_b/5$.
8. voir réponse de l'exercice 2.14.
9. On compare l'amplitude $y_k = a_p p(0) + w_k$ aux seuils $-2p(0)$, 0 et $2p(0)$.
10. On pose
$$p = \int_{2p(0)}^{\infty} \frac{1}{\sigma_w \sqrt{2\pi}} e^{-\frac{1}{2\sigma_w^2}(u - 3p(0))^2} du = Q(\sqrt{2p(0)/N_0})$$
La probabilité de décision correcte est :
$$\begin{aligned} P_c &= \frac{1}{4} \int_{-\infty}^{-2p(0)} p_{Y|-3}(y) dy + \frac{1}{4} \int_{-2p(0)}^{0} p_{Y|-1}(y) dy \\ &\quad + \frac{1}{4} \int_{0}^{+2p(0)} p_{Y|+1}(y) dy + \frac{1}{4} \int_{p(0)}^{\infty} p_{Y|+3}(y) dy \\ &= (1 - q + (1 - 2q) + (1 - 2q) + 1 - q)/4 \\ &= 1 - \frac{3}{2} Q\left(\sqrt{\frac{2p(0)}{N_0}}\right) \end{aligned}$$
En remplaçant il vient :
$$P_e = \frac{3}{2} Q\left(\sqrt{\frac{2p(0)}{N_0}}\right) = \frac{3}{2} Q\left(\sqrt{\frac{2}{5}\frac{2E_b}{N_0}}\right)$$
11. TEEB $= P_e/\log_2(M) = P_e/2$.
12. Le spectre ne contient pas de raies car la suite a_k est centrée.

Exercice 2.18 (Variance de l'IES)

1. En effet on a par identification :
$$\epsilon_n = \sum_k a_k p(nT) - a_n p(0) = \sum_k a_k g(n - k)$$
avec
$$g(k) = p(kT) - p(0)\delta_k$$

où $\delta_k = 0$ pour $k \neq 0$ et 1 pour $k = 0$. En utilisant la formule de Poisson on en déduit le gain en fréquence :

$$\begin{aligned} G(u) &= \sum_k g(k) e^{2j\pi k u} = \sum_k p(kT) e^{2j\pi k (u/T) T} - p(0) \\ &= \frac{1}{T}\left(\sum_n P((u-n)/T) - p(0)T\right) \end{aligned}$$

2. D'après les formules de filtrage, on a $\mathbb{E}\{\epsilon_n\} = 0$ puisque $\mathbb{E}\{a_k\} = 0$.
3. D'après les formules de filtrage, on a $S_\epsilon(f) = S_a(f)|G(f)|^2$ avec $S_a(f) = \sum_k R_a(k) e^{-2j\pi k f}$. On en déduit :

$$\begin{aligned} \mathrm{var}(\epsilon_n) &= \mathbb{E}\{\epsilon_n^2\} = \int_{-1/2}^{1/2} S_\epsilon(f) df \\ &= \int_{-1/2}^{1/2} \frac{1}{T^2} \left| \sum_n P((f-n)/T) - p(0)T \right|^2 S_a(f) df \end{aligned}$$

Plus le terme entre les barres de module est faible, plus la variance de l'IES est faible. Il s'annule si $p(t)$ satisfait le critère de Nyquist. Dans le cas pratique où le canal ne laisse pas passer les fréquences autour de 0 (voir figure 1), le critère de Nyquist n'est plus satisfait. Toutefois un codage approprié donnant un spectre $S_a(f)$ nulle à l'extérieur de ce trou en fréquence, donnera une variance faible de l'IES. Par conséquent un codage approprié peut réduire l'IES sur un canal possédant un trou en 0.

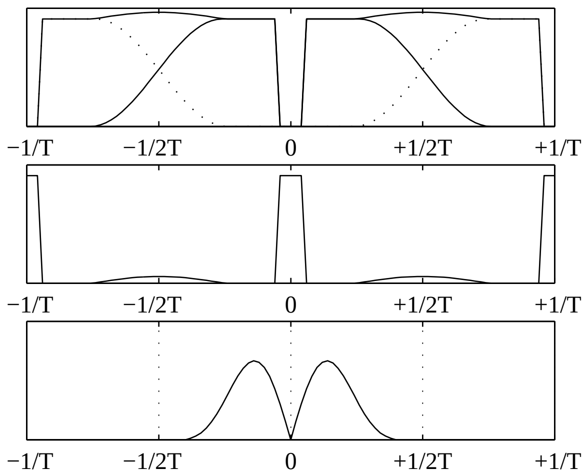

FIG. 1 – *Figure du haut : $\sum_k P(f - k/T)$ qui serait constant si $p(t)$ vérifiait le critère de Nyquist. Figure du centre : fonction de filtrage de l'expression donnant l'IES. Figure du bas : exemple de forme de corrélation sur les symboles induisant peu d'IES.*

Exercice 2.19 (Transmission sur porteuse) pas au programme

Exercice 2.20 (Choix de la modulation MDP-M) pas au programme

Exercice 2.21 pas au programme

3 Introduction aux codes correcteurs d'erreur

Exercice 3.1 Le code $\mathcal{C}(100, 50, 10)$ est préférable au code $\mathcal{C}(50, 25, 5)$: en effet ils ont même rendement $k/n = 1/2$ et le premier a une distance minimale plus grande et donc une meilleure capacité de correction.

Il n'existe pas de code linéaire $\mathcal{C}(24, 16, 10)$ car on doit avoir $d \leq n - k + 1$.

Exercice 3.2 (Canal à effacement) Si le canal introduit $d_{\min}-1$ effacements, il existe un unique mot-code à la distance $d_{\min}-1$. En effet, si il y en avait deux, il y aurait alors deux mots-code qui varient par moins de $d_{\min}-1$ bits. Ce qui est impossible puisque $d_{\min}$ est la distance minimale.

Exercice 3.3 Soit le code linéaire de matrice génératrice :

1.
$$H = \left[\begin{array}{ccc|c} 1 & 1 & 1 & \\ 0 & 1 & 1 & \\ 1 & 0 & 1 & I_4 \\ 1 & 1 & 0 & \end{array}\right]$$

2. Le syndrome correspondant est la i-ème colonne de H. Comme toutes les colonnes sont différentes, ce syndrome est caractéristique de la position de l'erreur.
3. Le syndrome correspondant est la somme des colonnes en position i et j. On ne peut pas corriger deux erreurs car la somme des colonnes 1 et 4 est la même que la somme des colonnes 2 et 3.

   On peut détecter deux erreurs en positions quelconques, puisque tout couple de colonnes a une somme différente de 0 ($\forall i \neq j, \mathrm{col}(i) \oplus \mathrm{col}(j) \neq 0$).

4 Eléments de théorie de l'information

Exercice 4.1 (Canal en Z) 1. On note $\beta = \mathbb{P}\{X=1\}$. On en déduit la loi conjointe de (X,Y) pour X et Y appartenant à $\{0,1\}$:

$X \setminus Y$	0	1
0	$(1-\beta)$	0
1	βp	βq

Par conséquent l'information mutuelle a pour expression :

$$\begin{aligned} I(X,Y) &= -(1-\beta)\log_2((1-\beta)+\beta p) + \beta p \log_2(\frac{p}{(1-\beta)+\beta p}) \\ &\quad -\beta q \log_2(\beta) \\ &= -(1-\beta)\log_2((1-\beta)+\beta p) + \beta p \log_2(p) \\ &\quad -\beta p \log_2((1-\beta)+\beta p) - \beta q \log_2(\beta) \\ &= -((1-\beta)+\beta p)\log_2((1-\beta)+\beta p) + \beta p \log_2(p) \\ &\quad +\beta q \log_2(q) - \beta q \log_2(q) - \beta q \log_2(\beta) \end{aligned}$$

En remarquant que $(1-\beta)+\beta p = 1-\beta q$ et en posant $h(x) = -x\log(x) - (1-x)\log(1-x)$, on a :

$$I(X,Y) = h(\beta q) - \beta h(q)$$

En annulant la dérivée par rapport à β, on obtient :

$$\beta = \frac{p^{p/q}}{1+qp^{p/q}}$$

On en déduit l'expression de la capacité :

$$C = h\left(\frac{qp^{p/q}}{1+qp^{p/q}}\right) - \frac{p^{p/q}}{1+qp^{p/q}} h(p)$$

2. Par suite du protocole utilisé, on a d'une part $\bar{\ell}_1 = (1-p)\times 1 + p \times (\bar{\ell}_0 + 1)$ et d'autre part $\bar{\ell}_0 = 1 + \bar{\ell}_1$. On en déduit que $\bar{\ell}_0 = 2/(1-p)$ et $\bar{\ell}_1 = (1+p)/(1-p)$. Par conséquent le nombre moyen d'utilisation du canal est :
$$\bar{\ell} = \frac{1}{2}\,\bar{\ell}_0 + \frac{1}{2}\,\bar{\ell}_1 = \frac{p+3}{2(1-p)}$$

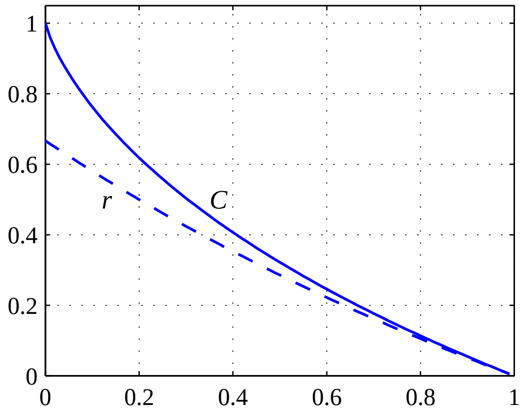

FIG. 2 – *Canal en Z. Courbe en trait plein : capacité en fonction de la probabilité d'erreur p. Courbe en pointillé : débit donnant une probabilité d'erreur nulle en fonction de p.*

3. Le nombre moyen de bits transmis par utilisation du canal est :

$$r = \frac{2(1-p)}{p+3}$$

Notons que le protocole utilisé conduit à une probabilité d'erreur moyenne qui tend vers 0. Notons que, pour $p = 1$, on trouve $r = 0$. Dans ce cas la capacité est nulle.

Exercice 4.2 On se limite au cas où $a \geq 0$. Si $a \notin \{0, 1\}$, Y peut prendre 4 valeurs différentes, à savoir $\{0, a, 1, 1+a\}$. Comme X et Z sont supposés indépendantes, on a :

$$\begin{array}{ll} \mathbb{P}\{Y=0|X=0\} = 1/2 & \mathbb{P}\{Y=0|X=1\} = 0 \\ \mathbb{P}\{Y=a|X=0\} = 1/2 & \mathbb{P}\{Y=a|X=1\} = 0 \\ \mathbb{P}\{Y=1|X=0\} = 0 & \mathbb{P}\{Y=1|X=1\} = 1/2 \\ \mathbb{P}\{Y=1+a|X=0\} = 0 & \mathbb{P}\{Y=1+a|X=1\} = 1/2 \end{array}$$

On en déduit que la capacité est égale à 1.

Si $a = 1$, Y peut prendre 3 valeurs différentes, à savoir $\{0, 1, 2\}$. Comme X et Z sont supposés indépendantes, on a :

$$\begin{array}{ll} \mathbb{P}\{Y=0|X=0\} = 1/2 & \mathbb{P}\{Y=0|X=1\} = 0 \\ \mathbb{P}\{Y=1|X=0\} = 1/2 & \mathbb{P}\{Y=1|X=1\} = 1/2 \\ \mathbb{P}\{Y=2|X=0\} = 0 & \mathbb{P}\{Y=2|X=1\} = 1/2 \end{array}$$

Il s'agit du canal à effacements : sa capacité est égale à 1/2.

Si $a = 0$, Y peut prendre 2 valeurs différentes, à savoir $\{0, 1\}$ et on a $\mathbb{P}\{Y=0|X=0\} = 1$ et $\mathbb{P}\{Y=1|X=1\} = 1$. La capacité est égale à 1.

Exercice 4.3 On note $X = (X_1, X_2)$ et $Y = (Y_1, Y_2)$ les entrée et sortie respectives du canal considéré. L'information mutuelle s'écrit :

$$I(X,Y) = \sum_{\mathcal{X}_1 \times \mathcal{X}_2} \sum_{\mathcal{Y}_1 \times \mathcal{Y}_2} p(y_1|x_1)p(y_2|x_2)p(x_1,x_2)$$
$$\log\left(\frac{p(y_1|x_1)p(y_2|x_2)}{\sum_{u_1 \in \mathcal{X}_1 \times u_2 \in \mathcal{X}_2} p(y_1|u_1)p(y_2|u_2)p(u_1,u_2)}\right)$$

En prenant pour $p(x_1)$ et $p(x_2)$ les lois qui maximisent respectivement $I(X_1, Y_1)$ et $I(X_2, Y_2)$, on a :

$$I(X,Y) \geq C_1 + C_2$$

Et par conséquent $C \geq C_1 + C_2$.

Exercice 4.4 Avec des notations évidentes on a :

$$Y_n = X_1 \oplus B_1 \oplus \cdots \oplus B_n = X_1 \oplus B$$

où B est une variable aléatoire à valeurs dans $\{0,1\}$ avec :

$$p_n = \mathbb{P}\{B=0\} = \sum_{j \text{ pair}} p^j q^{n-j}$$

La capacité est alors donnée par $C_n = 1 + H_2(p_n)$ où $h_2(x) = -x\log(x) - (1-x)\log(1-x)$.

Exercice 4.5 On considère le canal dont les alphabets d'entrée et de sortie comportent 5 symboles notés $\{0,1,2,3,4\}$ et dont les probabilités de transition sont :

$$\mathbb{P}\{Y=i|X=j\} = \begin{cases} 1/2 & \text{si} \quad i = j \pm 1 \bmod 5 \\ 0 & \text{sinon} \end{cases}$$

1. $C = \log(5/2)$.
2. Soit l'ensemble des $k=2$ messages $\mathcal{M} = \{0,1\}$ et le codage de $\mathcal{M}$ dans l'alphabet d'entrée (à la puissance $n=1$) qui associe $0 \mapsto 0$ et $1 \mapsto 3$ et qui adopte comme fonction de décodage 0 et $1 \mapsto 0$, 2 et $3 \mapsto 1$. Le débit est alors de $k/n = 1$ bit par utilisation du canal et la probabilité d'erreur est nulle. La capacité est donc supérieure à 1.

Exercice 4.6
1. $Y = e^X$. Comme la fonction $y = e^x$ est strictement monotone, à deux valeurs différentes de x correspondent deux valeurs différentes de Y et donc $H(Y) = H(X)$.
2. On note $\mathcal{X}$ les valeurs de X. Si les valeurs $\{y = \cos(x), x \in \mathcal{X}\}$ sont toutes distinctes, alors $H(Y) = H(X)$. si, par contre on a $\cos(x_1) = \cos(x_2)$, on a, en posant $\mu = p_1/(p1+p_2)$:

$$\begin{aligned} H(X) - H(Y) &= (p_1+p_2)\log(p_1+p_2) - p_1\log(p_1) - p_2\log(p_2) \\ &= -(p_1+p_2)(\mu\log(\mu) + (1-\mu)\log(1-\mu)) \geq 0 \end{aligned}$$

d'après la convexité de la fonction logarithme.

Exercice 4.7 On pose $q(n) = (1-\gamma)\gamma^n$ où $\gamma = A/(A+1)$. On vérifie que sa moyenne est A. En appliquant le lemme fondamental, il vient :

$$H(X) \leq -\sum_{n\in\mathbb{N}} p(n)\log((1-\gamma)\gamma^n) = \log((1-\gamma) + \log(\gamma)A$$

Cette borne indépendante de $p(n)$ peut être atteinte pour $q(n) = p(n)$. C'est donc le maximum.

Exercice 4.8 On a :

$$\begin{aligned} H(X) - H(Y) &= (p_i+p_j)\log(p_i+p_j) - p_i\log(p_i) - p_j\log(p_j) \\ &= = -(p_i+p_j)(\mu\log(\mu) + (1-\mu)\log(1-\mu)) \geq 0 \end{aligned}$$

d'après la convexité de la fonction logarithme.

Exercice 4.9 On considère les variables aléatoires X et Y à valeurs dans $\{0,1\}$ dont la loi conjointe est donnée par :

$$\mathbb{P}\{X=0,Y=0\} = 1/4,\ \ \mathbb{P}\{X=0,Y=1\} = 1/4,\ \ \mathbb{P}\{X=1,Y=0\} = 1/2$$

On en déduit $\mathbb{P}\{X=1,Y=1\} = 0$, $\mathbb{P}\{X=0\} = \mathbb{P}\{X=1\} = 1/2$, $\mathbb{P}\{Y=0\} = 3/4$ et $\mathbb{P}\{Y=1\} = 1/4$. $H(X) = 1$, $H(Y) = -3/4\log(3/4) - 1/4\log(1/4) = 2 - 3\log(3)$, $H(X,Y) = -1/2(\log(1/4) + \log(1/2)) = 3/2$, $H(Y|X) = H(X,Y) - H(X) = 1/2$, $H(X|Y) = H(X,Y) - H(Y) = 3\log(3) - 1/2$, $I(X,Y) = H(X,Y) - H(X) - H(Y) = 3\log(3) - 1$.

Exercice 4.10 (Inégalité de Fano) On considère deux variables aléatoires discrètes X et Y. On note $M = |\mathcal{X}|$ la cardinalité de l'ensemble des valeurs de X. On considère la variable aléatoire $\hat{X} = g(Y)$ obtenue à partir de Y par la fonction (mesurable) g. On pose $E = \mathbb{1}(\hat{X} \neq X)$ et on note $P_e = \mathbb{P}\{E=1\}$.

1. Comme $E = \mathbb{1}(g(Y) \neq X)$ est une fonction de X et de Y, d'après l'exemple **??**, on a $H(E|X,Y) = 0$.
2. On a d'une part $H(X,Y,E) = H(X,Y) + H(E|X,Y) = H(X,Y)$ et d'autre part $H(X,Y,E) = H(E,Y) + H(X|E,Y)$ et donc $H(X,Y) = H(E,Y) + H(X|E,Y)$. On en déduit $H(X|Y) = H(E|Y) + H(X|E,Y)$.
3. On a $H(E|Y) \leq H(E)$ où $H(E) = -P_e \log(P_e) - (1 - P_e)\log(1 - P_e)$.
4. On a $H(X|Y,E) = H(X|Y,E = 0)(1 - P_e) + H(X|Y,E = 1)P_e \leq P_e \log(M - 1)$. Le terme $H(X|Y,E = 0) = 0$: en effet, si $E = 0$, alors $X = g(Y)$. Le terme $H(X|Y,E = 1) \leq \log(M - 1)$: en effet, si $E = 1$, $X \neq g(Y)$ et par conséquent X ne peut prendre que $(M - 1)$ valeurs autres que $g(Y)$. Par conséquent son entropie est inférieure ou égale à $\log(M - 1)$.

On en déduit (inégalité de Fano) $H(P_e) + P_e \log(M - 1) \geq H(X|Y)$.