EISTI. **Année scolaire 2008-2009**

DÉPARTEMENT "INFORMATIQUE"

THÉORIE DE L'INFORMATION

Série d'exercices N^{o}1

PARTIE I. REPRÉSENTATION MATHÉMATIQUE D'UN CANAL DE TRANSMISSION

Exercice 1 (Exemple)**.** Soient X et Y deux variables aléatoires prenant leurs valeurs respectivement dans $\Omega_X = \{x_1, x_2, x_3\}$ et $\Omega_Y = \{y_1, y_2\}$ et ayant la matrice de probabilités conjointes suivante

$$P(X,Y) = \begin{array}{c|cc} & y_1 & y_2 \\ \hline x_1 & 0.25 & 0 \\ x_2 & 0.1 & 0.3 \\ x_3 & 0.1 & 0.25 \end{array}$$

En déduire les distributions conditionnelles et marginales.

Solution de l'exercice 1

Distributions marginales de X et Y. On utilise les définitions

$$\forall i = 1, \ldots, 3,\ p(x_i) = \sum_{j=1}^{2} p(x_i, y_j), \quad \forall j = 1, \ldots, 2,\ p(y_j) = \sum_{i=1}^{3} p(x_i, y_j)$$

On peut résumer ces calculs sous forme d'un tableau

	y_1	y_2	P_X
x_1	0.25	0	0.25
x_2	0.1	0.3	0.4
x_3	0.1	0.25	0.35
P_Y	0.45	0.55	

La dernière colonne de ce tableau représente la distribution marginale de X et s'obtient en calculant les sommes des éléments de chaque ligne.

La dernière ligne du tableau représente la distribution marginale de Y et s'obtient en calculant les sommes des éléments de chaque colonne.

Distributions conditionnelles $P(X|Y)$ et $P(Y|X)$. On utilise la définition

$$\forall i = 1, \ldots, 3,\ \forall j = 1, \ldots, 2,\ p(x_i|y_j) = \frac{p(x_i, y_j)}{p(y_j)}, \quad p(y_j|x_i) = \frac{p(x_i, y_j)}{p(x_i)}$$

On trouve les matrices

$$P(X|Y) = \begin{pmatrix} 5/9 & 0 \\ 2/9 & 6/11 \\ 2/9 & 5/11 \end{pmatrix}, \qquad P(Y|X) = \begin{pmatrix} 1 & 0 \\ 1/4 & 3/4 \\ 2/7 & 5/7 \end{pmatrix}$$

Exercice 2 (Un modèle probabiliste de transmission)**.** Soit une source binaire qui émet des signaux binaires, composés de 0 et de 1. On associé à l'expérience d'envoi d'un seul symbole une variable aléatoire X qui prend donc des valeurs dans $\Omega_X = \{0, 1\}$. On suppose que la source émet 0 ou 1 avec équiprobabilité. Autrement dit, la distribution de X est connue :

$$p_X(0) = p_X(1) = 0.5$$

Les symboles émis sont ensuite envoyés via un canal de transmission ayant des perturbations aléatoires. On associe à l'expérience d'observation du symbole reçu la variable aléatoire Y qui peut prendre trois valeurs : $\Omega_Y = \{-1, 0, 1\}$. La valeur -1 correspond au cas où le système n'est pas capable d'identifier un 0 ou un 1 à la sortie. On a établi les probabilités de réception suivantes :

- Si le symbole envoyé X est 0 alors on reçoit :
  1. $Y = 1$ avec la probabilité 0.2
  2. $Y = 0$ avec la probabilité 0.7
  3. $Y = -1$ avec la probabilité 0.1
- Si le symbole envoyé X est 1 alors on reçoit :
  1. $Y = 1$ avec la probabilité 0.6
  2. $Y = 0$ avec la probabilité 0.3
  3. $Y = -1$ avec la probabilité 0.1

1. Dire si les probabilités de réception données ci-dessus définissent la distribution conjointe ou conditionnelle de X et Y ? Si conditionnelle, préciser de quelle variable ?
2. Pouvez-vous donner une représentation sous forme de graphe des probabilités de réception ?
3. Former une matrice à partir des probabilités de réception.
4. Calculer toutes les distributions manquantes.

PARTIE II. VERS LA NOTION D'ENTROPIE

Exercice 3. Soit l'algorithme suivant.

Entrée : $x_0 \in \{1, 2, 3, 4\}$ tiré aléatoirement avec équiprobabilité.

Algorithme : $n \leftarrow 0$

TANT QUE $(x_n \neq 1)$

$$A_n \leftarrow \begin{cases} 0, \text{ avec probabilité } p_n = 1/x_n \\ 1, \text{ avec probabilité } q_n = 1 - 1/x_n \end{cases}$$

SI (x_n est pair)

ALORS $x_{n+1} \leftarrow \dfrac{x_n}{2}$

SINON $x_{n+1} \leftarrow A_n x_n + 1$

FIN SI

$n \leftarrow n + 1$

FIN TANT QUE

Sortie : Nombre d'itérations n.

Soient X, la variable aléatoire correspondante à l'entrée de l'algorithme, uniformément distribuée sur $\{1, 2, 3, 4, \}$ et N la variable aléatoire correspondante à la sortie de l'algorithme.

L'objectif de l'exercice *est de calculer le nombre moyen d'itérations effectuées par cet algorithme. Autrement dit, calculer $E[N]$.*

Voici le plan à suivre.

1. Etablir la matrice de probabilités conditionnelles $P(N|X)$, en utilisant les schémas sous forme d'arbre pour le déroulement de l'algorithme pour chaque valeur de X.Voir l'exemple sur la figure ci-dessous.
2. Sachant que X est uniformément distribué, déduire la matrice de probabilités conjointes de N et X.
3. Calculer la distribution de probabilités marginale de N
4. En déduire le nombre moyen d'itérations.

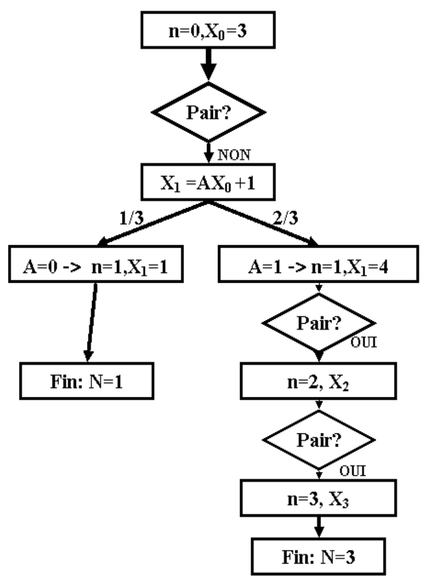

FIGURE 1 – $X_0 = 3$.

PARTIE III. POUR APPROFONDIR

Exercice 4 (Le jeu de la fausse pièce)**.** On considère un jeu de n pièces, toutes d'apparence identiques. On sait qu'une seule pièce est fausse. Elle a un poids différent des autres mais on ne sait pas si elle plus légère ou plus lourde.

On dispose d'une balance à deux plateaux. À chaque pesée la balance peut se trouver dans l'une des trois configurations :

G Elle penche vers la gauche.

D Elle penche vers la droite.

E Elle reste à l'équilibre.

L'objectif est de déterminer avec le moins de pesées possible la fausse pièce et si elle plus légère ou plus lourde.

1. Soit $n = 8$.
   (a) Combien de réponses possibles il y a dans ce problème ?
   (b) Combien de possibilités différentes peut on obtenir avec 2 pesées ? Avec 3 pesées ? Conclusion sur la faisabilité de détermination de la fausse pièce en 2 ou en 3 pesées.
   (c) Est-il intéressant de peser deux lots de 4 pièces chacun ? Pourquoi ? Combien d'information nous apporterait une telle pesée ?
   (d) A l'aide du programme java à l'adresse `http://nlvm.usu.edu/fr/nav/frames_asid_139_g_4_t_2.html` Élaborer une stratégie de pesée permettant de déterminer la fausse pièce en 3 pesées dans tous les cas. Voici quelques conseils qui peuvent être utiles
      i. Essayez, pour commencer, de diviser l'ensemble de vos pièces en trois lots, de taille à peu près égale. Pour 8 pièces vous pouvez tester les décompositions $8 = 3 + 3 + 2$ ou $8 = 2 + 2 + 4$;
      ii. Lorsque l'on compare le même lot de pièces $L1$ à deux lots différents $L2$ et $L3$ on peut interpréter les résultats de façon suivante : si les deux résultats de comparaison sont identiques, la fausse pièce est dans le lot 1 et nous avons le sens de la différence de poids ($+$ ou $-$) ; il n'est pas possible que l'un des résultats soit G et l'autre D.
   (e) Essayez de représenter votre stratégie sous forme d'arbre de décision. Les feuilles de l'arbre doivent représenter toutes les réponses possibles. La profondeur de cet arbre indique alors le nombre maximal de pesées.
2. A l'aide du même programme java élaborer une stratégie pour $n = 9, 10, 12$.