Informatique II - Série 9

Exercice 9-1: Entropie

Classer par ordre croissant d'entropie les distributions de probabilité associées aux évènements suivants:

1. Le jet d'un dé 6 faces non pipé
2. Le jet d'une pièce de monnaie
3. Le jet d'un couple de dés (un rouge un bleu) 6 faces
4. Le tirage au sort d'une valeur tirée de manière équiprobable entre 2 et 12
5. La valeur du jet d'un dé multipliée par deux

Si l'on considère qu'un dé truqué tombe plus souvent sur certaines faces, comment se comporte l'entropie par rapport aux tirages décrit ci-dessus ?

Solution

1. $H(X) = \log_2 6$
2. $H(X) = \log_2 2$
3. $H(X, X') = \log_2 36$
4. $H(X) = \log_2 11$
5. $H(X) = \log_2 6$

Le comportement d'un dé truqué dépend de la manière dont il est truqué. Par exemple, il peut toujours donner la mme valeur,il est alors déterministe, d'entropie nulle. Dans tous les cas, son entropie est strictement plus petite que celle d'un dé non pipé, $\log_2 6$.

Exercice 9-2: Stratégie de questionner: jeu de trésor

Il y a un trésor caché dans une de 5 boites, appelées A,B,C,D, et E. Des connaissances préablables vous permettent de savoir que la probabilité que le trésor se trouve en A est de $p_A = 0.20$, celles des autres: $p_B = 0.25$, $p_C = 0.25$, $p_D = 0.15$, $p_E = 0.15$. Vous pouvez poser des questions à un oracle qui répondra avec oui ou non.

1. Quelle est la meilleure stratégie?
2. Combien de questions est-ce qu'il faut poser (en moyenne) pour localiser le trésor, si le jeu est répéter 100 fois.?

Solution Avec les arbres de Huffman, on aura besoin de 2 questions si il se trouve en A,B ou C, et 3 questions si il se trouve en D ou E. Dans 100 jeu on aura besoin de 230 questions environ.

Exercice 9-3: Stratégie de questionner: jeu de trésor

a) Reprendre l'exercice 9-2, mais avec $p_A = 0.45$, $p_B = 0.2$, $p_C = 0.15$, $p_D = 0.1$, $p_E = 0.1$.

Montrer que les abres de Huffmann donnent 2 stratégies équivalentes, la première utilisant jusqu'à 4 questions, la deuxième 3 questions au maximum. Montrer que tous les deux donnent la même longueur moyenne, c.a.d., 210 questions pour 100 jeu.

Calculer l'entropie H et comparer avec le nombre moyen de questions.

Solution.

Stratégie 1: Séparer une possibilité après l'autre: A vs. B,C,D,E. Ensuite B vs. C,D,E,; ensuite C vs. D,E; finalement D vs E. La longueur moyenne du jeu est alors 0.45*1 + 0.2*2 + 0.15*3 + 0.2*4 = 2.1.

Stratégie 2: A vs. B,C,D,E. Ensuite B,C vs. D, E. Finalement B vs. C et D vs. E. La longueur moyenne du jeu est alors 0.45*1 + 0.55*3 = 2.1.

$H = 2.05 < 2.1$.

b) Refaire la même chose, mais avec $p_A = 1/2$, $p_B = 1/4$, $p_C = 1/8$, $p_D = 1/16$, $p_E = 1/16$. Calculer l'entropie H et comparer avec le nombre moyen de questions.

Solution.

Stratégie 1 - séparation successive est optimale.

Le nombre moyen de questions est $15/8 = 1.87$ ...

L'entropie H = 1.87. Au cas où toutes les probabilités sont $p = (1/2)^n$ avec n entier, le nombre moyen de questions est égal à l'entropie. L'entropie mesure l'incertitude donc le nombre moyen de questions nécessaires pour arriver à la certitude.